MASTERARBEIT | MASTER'S THESIS

Titel | Title

Semi-synthetic strategies for the synthesis of post-translationally modified Hsp90 variants

verfasst von | submitted by

Moritz Neuhauser BSc

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

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 862
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Chemie
Betreut von | Supervisor:	Univ.-Prof. Dr. Christian Friedrich Wilhelm Becker

Abstract

This thesis comprises two sub-projects, focusing on semi-synthetic methods that integrate recombinant DNA technology and solid-phase peptide synthesis (SPPS) to generate specifically modified protein domains. Semi-synthetic access to modified Heat shock protein 90 (Hsp90) variants will provide new insights into the possible influences of specific modifications on the chaperones structure and function. As Hsp90 is subject to various post-translational modifications (PTMs), the great challenge is to decode patterns of PTMs that modulate protein function. Here we rely on two synthetic strategies based on the principles of native chemical ligation (NCL). The first strategy is focusing on the synthesis of post-translationally modified variants of the Hsp90 N-terminal domain (NTD). This includes the recombinant expression of a polypeptide with a N-terminal cysteine residue and SPPS of a 19 amino acids long N-terminal peptide hydrazide, carrying PTMs at specific positions. The synthetic peptide hydrazide is ligated to the recombinantly expressed segment using *in situ* hydrazide to thioester conversion, followed by a novel *in situ* post-ligation-desulfurization approach to yield the native alanine residue. The second semi-synthetic strategy aims to generate full length Hsp90, using a sequential multiple segment semi-synthesis strategy from C- to N-terminus. To this end, residues 1-528 are produced via recombinant expression and ligated to the remaining residues 529-732. Preparation of segment 529-732 is achieved via ligation of three different segments in total. In a first diselenide selenoester ligation (DSL) the recombinant segment 572-710 is ligated to peptide 711-732 that is prepared by SPPS. In a second ligation, the semisynthetic construct 572-732 is ligated to another peptide 529-571 prepared by SPPS to generate the semisynthetic middle- and C-terminal domain (MD/CTD). This strategy allows for the simultaneous introduction of modifications in the two different domains. Consequently, this ligation scheme yields modified 529-732 building blocks, that can be used to synthesize the first full-length modified semi-synthetic Hsp90 and thus give information about the influence of post-translational modifications that have not been studied yet. Both strategies highlight the versatility and adaptability of protein semi-synthesis, enabling the production of homogeneous protein with specific natural or unnatural modifications at either the N- or C-termini or the middle domain.

Zusammenfassung

Diese Arbeit umfasst zwei Teilprojekte, die sich auf semi-synthetische Methoden konzentrieren, die rekombinante DNA-Technologie und die Festphasenpeptidsynthese (SPPS) beinhalten, um speziell modifizierte Proteindomänen zu erzeugen. Semi-synthetischer Zugang zu modifizierten Varianten des Hitzeschockproteins 90 (Hsp90) gewährt neue Einblicke in die möglichen Einflüsse spezifischer Modifikationen auf die Struktur und Funktion der Chaperone. Da Hsp90 verschiedenen posttranslationalen Modifikationen (PTMs) unterliegt, besteht die große Herausforderung darin, das Muster der PTMs zu entschlüsseln, die die Funktion des Proteins modulieren. Hierbei verlassen wir uns auf zwei synthetische Strategien, die auf den Prinzipien der nativen chemischen Ligation (NCL) basieren. Die erste Strategie konzentriert sich auf die Synthese posttranslational modifizierter Varianten der Hsp90 N-terminalen Domäne (NTD). Dies beinhaltet die rekombinante Expression eines Polypeptids und die Festphasensynthese eines 19 Aminosäuren langen Peptids, das PTMs an spezifischen Positionen trägt. Nach erfolgreicher Ligation soll ein neuartiger *in situ* post-ligations-desulfurierungsansatz implementiert werden. Die zweite semi-synthetische Strategie zielt darauf ab, Hsp90 in voller Länge zu erzeugen, indem eine sequenzielle Mehrsegment-Semi-Synthese-Strategie von C- nach N-Terminus angewendet wird. Zu diesem Zweck werden die Aminosäuren 1-528 durch rekombinante Expression produziert und an die verbleibenden Aminosäuren 529-732 ligiert. Die Herstellung des Fragments 529-732 erfolgt durch Ligation von insgesamt drei verschiedenen Segmenten. In einer ersten Diselenid-Selenoester-Ligation (DSL) wird das rekombinante Segment 572-710 an das durch SPPS hergestellte Peptid 711-732 ligiert. In einer zweiten Ligation wird das semi-synthetische Konstrukt 572-732 an ein weiteres durch SPPS hergestelltes Peptid 529-571 ligiert. Diese Strategie ermöglicht die simultane Einführung von Modifikationen in zwei Domänen. Die modifizierten 529-732-Bausteine können verwendet werden, um das erste modifizierte Hsp90 in voller Länge herzustellen und somit Informationen über den Einfluss posttranslationaler Modifikationen zu liefern, die bisher noch nicht untersucht wurden. Beide Strategien unterstreichen die Vielseitigkeit und Anpassungsfähigkeit der Protein-Semi-Synthese, die die Produktion homogener Proteine mit spezifischen natürlichen oder unnatürlichen Modifikationen an den N- oder C-Termini oder in der mittleren Domäne ermöglicht.

Table of contents

List of Abbreviations

ACN	Acetonitrile
ATP	Adenosine triphosphate
Boc	tert-Butyloxycarbonyl
tBu	tert-Butyl
CBD	Chitin binding domain
DCM	Dichloromethane
DMF	Dimethylformamide
DMS	Dimethyl sulfide
DIPEA	Diisopropylethylamine
EPL	Expressed protein ligation
Gdn·HCl	Guanidine hydrochloride
HATU	[O-(7-Azabenzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium-hexafluorphosphat]
MS	Mass Spectrometry
MPAA	4-Mercaptophenylacetic acid
MTG	Methyl thioglycolate
MesNa	Sodium methanethiolate
TFA	2,2,2-Trifluoroacetic acid
TIPS	Triisopropyl silane
TCEP	tris(2-carboxyethyl)phosphine

1. Introduction

1.1 Structure and function of heat shock protein 90 (Hsp90)

The heat shock protein 90 (Hsp90) serves as a crucial controller of proteostasis (i.e. protein homeostasis) in eukaryotic cells, maintaining protein homeostasis both under normal physiological conditions and during periods of cellular stress. It belongs to the family of molecular chaperones, which are essential to ensure that changes within cells that affect proteins can be controlled, and thus cellular homeostasis can be maintained under physiological or altered conditions. This is of special importance under stress conditions, such as increased temperature, hence the name “heat shock protein”.[1] There are various members of the heat shock protein family, categorized according to their molecular weights, including prominent groups such as large Hsp’s (Hsp90, Hsp70, Hsp60, Hsp40) and small Hsp’s.[2] Figure 1 gives a schematic overview of the proteins function and structure.

Chaperones are characterized as proteins that interact with and stabilize an otherwise unstable conformer of another protein, usually termed substrate or client. Through regulated binding and release, chaperones assist in ensuring the proper fate of the client *in vivo*, whether this involves folding, binding ligands to their receptors or targets, oligomeric assembly (as depicted in *fig. 1a*), facilitating transport to a specific subcellular compartment, or orchestrating disposal through degradation.[3] There are several hundred known clients of Hsp90, most of them being involved in signal transduction pathways, including transcription factors, protein kinases and E3 ubiquitin ligases.[4][5] However, a significant number of Hsp90-client proteins play crucial roles in establishing cancer hallmarks and the function and stability of many oncogenic mutant proteins rely heavily on Hsp90 and its co-chaperones.[6][7] The Chaperone is also one of the most abundant proteins in the cytosol under physiological conditions, and is involved in many other cellular processes beyond protein folding, such as DNA repair, the immune response, and neurodegenerative diseases.[1][8] Its abundance and multifunctionality make Hsp90 an appealing drug target for cancer treatment and diseases associated with protein misfolding (e.g., Alzheimer's, Parkinson's, Creutzfeldt-Jakob, Huntington's).[1] Consequently, a comprehensive understanding of Hsp90's complex ATPase conformation cycle, co-chaperones, and post-translation modifications is imperative.[9]

Hsp90 consists of three highly conserved domains (*Figure 1b)*). The amino-terminal domain (NTD) mediates binding to ATP.[10] The middle domain (MD) is of importance for ATP hydrolysis and client binding. Through the carboxy-terminal domain (CTD) dimerization is achieved, which is essential for the proteins function *in vivo*.[11-12] The CTD further contains a C-terminal motif with the amino acids Met-Glu-Glu-Val-Asp (MEEVD), that serves as an interaction site for co-chaperones.[13] NTD and MD are connected via a long, flexible, and charged linker, that affects Hsp90 function in closing the NTD among ATP binding.[14][15][16][17]

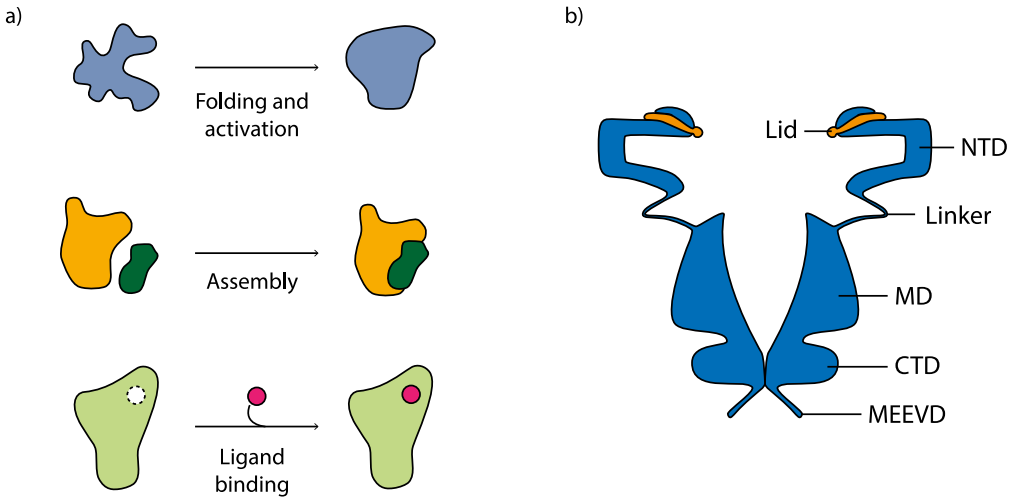


Figure 1. a) *Schematic representation of Hsp90s function, facilitating protein folding, the assembly of multiprotein complexes and the binding of a ligand to its receptor or target* ***b)*** *Schematic representation of Hsp90s structure, consisting of the three highly confined domains: The NTD , that is connected to the MD via a linker and the CTD carrying the MEEVD motif*

A schematic representation of Hsp90s conformational cycle is shown in figure 2. When there is no ATP present, the protein adopts a V-shaped, open confirmation.[18-19] This changes drastically when ATP is bound, leading to a N-terminally closed conformation *(Fig.2 "Closed 2 state")* with intermediate steps.[20] This conformational change affects every domain of the protein.[1] First, rearrangements take place in the NTD where a "lid" closes over the ATP-bound state *("Intermediate state")*. After dimerization of the NTDs *("Closed 1 state")* and association with the MDs *("Closed 2 state")* ATP hydrolysis can be achieved.[1][21] The NTDs then dissociate, ADP and inorganic phosphate P_i are released, and the protein adopts the open conformation again.[1]

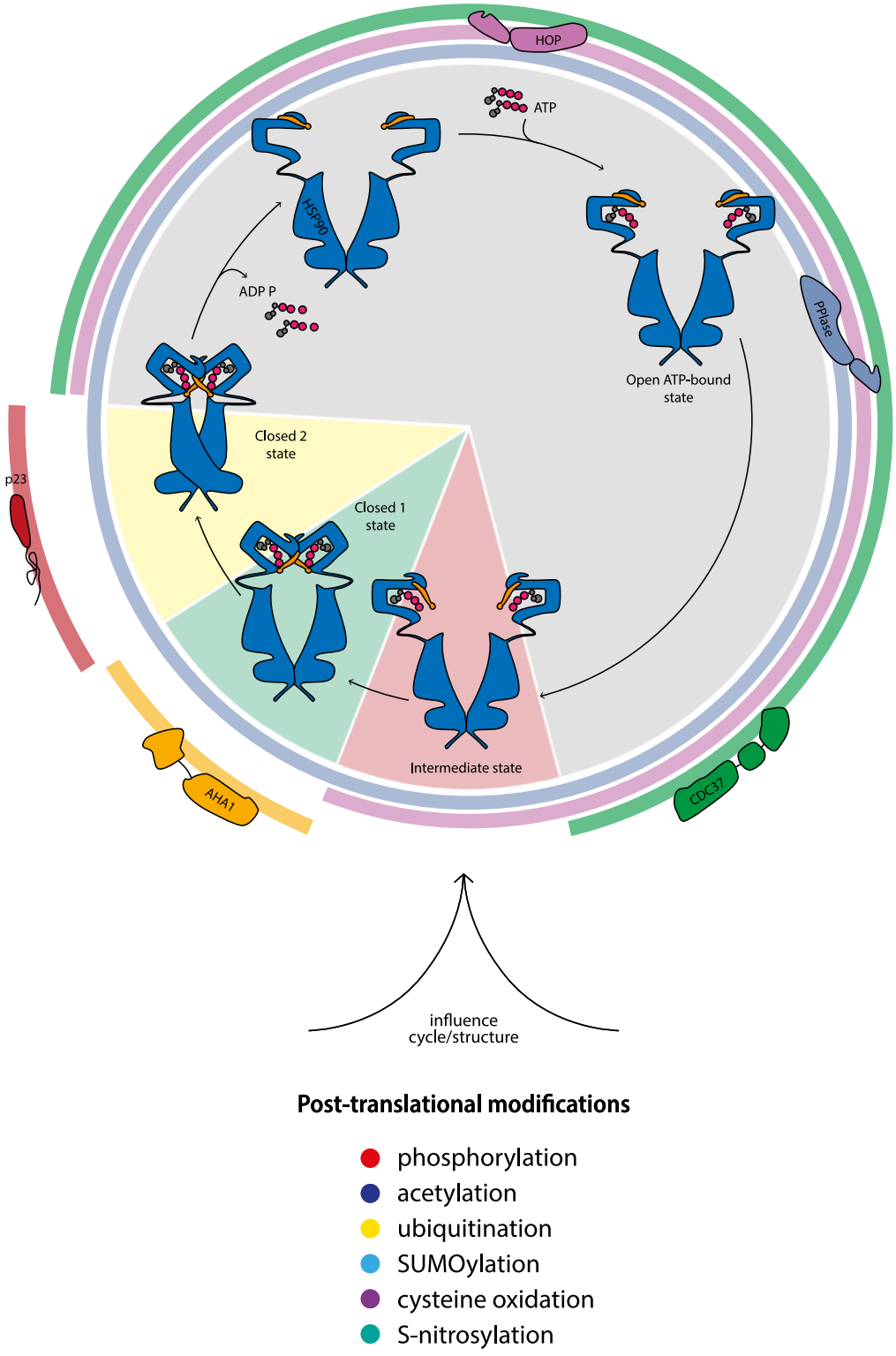


Figure 2. *Schematic representation of Hsp90s conformational cycle and the role of different co-chaperones during the cycle. Transition from the open ATP-bound state to the intermediate state with ATP bound and a closed lid. Interaction between the monomers NTDs (closed 1 state) and twisting of the Hsp90 monomers (closed 2 state). HSC70/HSP90-organizing protein (HOP) stabilizes the open conformation and inhibits the ATPase activity of Hsp90. Peptidyl-prolyl cis-trans isomerases (PPIases) are thought to be present at all conformational states. CDC37 binds early in the cycle and recruits*

clients' kinases. Activator of HSP90 ATPase homologue 1 (AHA1) promotes closed 1 state and accelerates the cycle. P23 stabilizes the closed 2 state and regulates the cycles progression by reducing the ATPase activity of Hsp90. The figure is adapted from Buchner et. al. and created with Adobe Illustrator.[1]

1.2 Regulation of Hsp90 chaperone activity

The function of Hsp90 can be regulated in many ways, including transcriptional regulations as well as post-translational modifications, the presence of co-chaperones and the clients themselves.[1]

The large cohort of co-chaperones that associate with eukaryotic Hsp90 make it one of the most complex chaperone machineries known to date.[22] Some of the co-chaperones associate to specific conformational states and lead to inhibition (e.g. HOP and p23) or activation (e.g. AHA1) of the ATPase activity, as shown in figure 2.[23] But they can also aid in the recruitment of specific client proteins.[24] The latter is confirmed by an example of early studies on the client protein steroid hormone receptor (SHR). These soluble cytosolic proteins are reliant on Hsp90 for their capacity to bind to their respective steroid hormone ligands.[25] In this context, SHR initially forms a complex with the molecular chaperone Hsp70 and its co-chaperone Hsp40. This complex, facilitated by the adaptor protein Hop/Sti1, subsequently interacts with Hsp90. Within this assembly, the client protein transitions from Hsp70 to Hsp90. Following this step, Hsp70 and Hsp40 dissociate from the Hsp90 complex. After dissociation, two additional co-chaperones, a member of the large prolyl isomerases (PPIases) and the small protein p23, associate with the Hsp90-client complex. For SHRs, this sequential progression through various complexes is crucial to attain the activatable state necessary for ligand binding.[22] Encrypting Hsp90 co-chaperone interactions and regulation remains a challenging issue because the analysis of hetero-oligomeric complexes requires an array of biophysical techniques like surface plasmon resonance spectroscopy, isothermal titration calorimetry, electron microscopy or novel approaches such as ultracentrifugation in combination with fluorescence labelling and detection.[14][26][27][28][29][30][31] However, according to Horwich *et al.*, the reconstitution of the chaperone cycle in vitro is within reach when previously outlined techniques are combined.[22]

1.3 Regulation trough post-translational modifications – the chaperone code

The possibilities of Hsp90 regulation can be readily extended to post translational modifications. Here, the challenge lies in deciphering the comprehensive and combinatorial

array of PTMs that modulate Hsp90 chaperone function, a phenomenon termed the “chaperone code”.[32]

The two most common post translational modifications for Hsp90 are phosphorylation and acetylation, accompanied by more rare modifications like s-nitrosylation, ubiquitination, SUMOylation and cysteine oxidation. Research aimed at understanding the influence of post-translational modifications on Hsp90 has revealed that these modifications significantly affect various aspects of Hsp90's chaperone function. The modified states lead to alterations in Hsp90 ATPase activity, interactions with co-chaperones and clients, client maturation, subcellular localization, degradation, and sensitivity to Hsp90 inhibitors.[33][34][35] Additional works further detail how Hsp90 modification affects downstream cellular processes such as cell cycle control, DNA repair, and steroid hormone signaling as well as many others via effects on the stability and activity of Hsp90 client proteins.[32] While a large number of modification sites have not been validated or linked to functional consequences, some have been more thoroughly studied and will be discussed in the next paragraph. Single modification sites are frequently studied using mimics such as glutamic acid to introduce a negative charge, though this is not equivalent to actual phosphorylation. Phosphomimetic mutations have limitations in accurately replicating the specific structural and functional impacts of phosphorylation. For instance, substituting phosphoserine or phosphothreonine with glutamic acid may not fully mimic the biological activity of the true phospho-residue due to variations in size, charge distribution, and hydrogen bonding potential.[36] Multiple modification sites on one target molecule are usually studied by using enzymes, that are known to process and modify the molecule of interest. However, enzymes often fail to produce homogeneous products and may only enable significant changes in function and structure upon modification introduction. This highlights the importance of a semi-synthetic approach to further deduce structure-activity relationships of Hsp90 in a more authentic, controlled and precise manner.

Phosphorylation of serine, threonine, and tyrosine residues is among the most well-studied of these Hsp90 modifications. Interestingly, phosphorylation can have various impacts, depending on the residues and domains that are modified, illustrating the complexity of the chaperone code. Studies about the impact of phosphorylation on the chaperone cycle show, that NTD phosphorylation of the conserved threonine residue T22 in yeast Hsp90 (human Hsp90a-T36) by the casein kinase 2 (CK2) disrupts binding of Hsp90 with the kinase-specific co-chaperone Cdc37.[37-38] Additionally, it compromises Hsp90 interaction with the activating co-chaperone Aha1 and subsequently results in decreased Hsp90 ATPase activity.[37] The introduction of a phosphomimetic mutation at Hsp90a-T90 enhanced its affinity for co-chaperones such as Aha1,

p23, PP5, and the C-terminus of Hsp70-interacting protein (CHIP), while concurrently reducing its interaction with Hsp70, Cdc37, and HOP.[39] On the contrary, when specific residues in the C-terminal domain (CTD) of Hsp90a undergo phosphorylation by kinases CK2, CK1, and glycogen synthase kinase 3b (GSK3b), there is an augmentation in the association between Hsp90a and HOP, coupled with a reduction in the interaction with CHIP. Studies on the phosphorylated C-terminal residues p725 and p726 showed diminished interaction with CHIP, which resulted in decreased ubiquitination and in turn degradation of Hsp90. These data point to Hsp90 ubiquitination as a critical cellular regulatory control module. Additionally, Hsp90a-T725 and -S726 and Hsp90b-S718 phosphomimetic mutations decreased the doubling time of HEK293 cells.[40] When examining the charged linker region of Hsp90, the phosphorylation of two distinct residues, S231 and S263, proved essential for maintaining telomerase activity by ensuring stability in the hTERT (human telomerase reverse transcriptase) telomerase catalytic subunit, known as an Hsp90 client. Woo et al. reported that cells expressing a truncated variant of the co-chaperone p23 exhibited a diminished phosphorylation status of these two residues, ultimately resulting in the downregulation of telomerase activity.[41] In another study conducted by Soroka *et al.*, the pivotal influence of phosphatases, specifically the co-chaperone Ppt1 (yeast ortholog of Ppt5), on the phosphorylation status of Hsp90 was examined. The researchers identified a total of 10 residues dephosphorylated by Ppt1, and each of these residues differentially affected both ATPase activity and client functionality. Interestingly, the phosphomimetic mutant of residue S485E in Hsp90's middle domain (MD) exhibited the most significantly reduced ATPase activity, despite its distance from the ATP-binding pocket. This observation suggests that the phosphorylation of this particular residue may induce distant structural changes.[42] A study by Yang *et al.* shows that an acetylation mimic mutation on MD-residue 546 results in a substantial decrease in ATP-binding and chaperone function of Hsp90 but increases Hsp90 binding to its inhibitors (e.g. biotinylated geldanamycin).[43] As noted earlier, Hsp90 presents an appealing target for therapeutic intervention, especially in cancer, given its pivotal role in crucial processes such as cell proliferation, survival, and migration. Thus, the inhibition of Hsp90 offers a distinctive approach to concurrently address varied signaling pathways, including those that are mutated or overexpressed in cancers, such as oncogenic kinase pathways.[44] In summary, these various examples illustrate phosphorylation as a dynamic regulatory mechanism in the Hsp90 chaperone cycle.

1.4 Solid phase peptide synthesis (SPPS)

The aspiration to create synthetic proteins dates back to the early 20th century. Emil Fischer managed to synthesize the first ever dipeptide glycyl glycine in 1901, followed by an octadecapeptide 6 years later.[45][46] Following efforts led to the introduction of the first urethane-based protecting group by Bergmann and Zervas in 1932.[47] This in turn, led to accessibility of small peptide hormones like oxytocin in solution.[48] In 1963, Merrifield introduced solid phase peptide synthesis (SPPS). This method is based on attaching an amino acid to an insoluble support via its carboxyl group, so the next amino acid can be coupled to it via its deprotected amino group.[49] Essential to every biological chemist's repertoire, this paved the way for the later introduced automated SPPS, which is common practice in laboratories around the world nowadays.

The majority of synthetic peptides are now prepared by Fmoc solid phase peptide synthesis (SPPS). This method allows site-specific modification of peptides such as introducing non-natural amino acids, fluorescent tags or isotope labeled amino acids. Fmoc-chemistry also provides a solution to the limitations of the Boc method as its deprotection conditions are compatible with modified peptides, such as phosphorylated and glycosylated peptides.[50] Understanding the role of protein post-translational modifications (PTMs) in cell signaling, gene expression, and protein processing and translocation is of enormous interest.[51][52][53][54] For instance, reversible phosphorylation of proteins is found to be involved in signal transduction and on/off control of enzymes.[55] Sulfation is thought to be involved in the modulation of the extracellular protein-protein interactions of secreted and transmembrane proteins.[56-57] Methylation occurs to be involved in protein targeting and signaling and in the epigenetic control of gene expression.[58] Ready access to peptides containing post-translational modified amino acid residues, for use as probes, inhibitors or for raising antibodies against protein PTMs, has been crucial for advances in these areas. The milder chemistry of the Fmoc method allows almost all PTMs such as phosphorylation, sulfation, methylation, citrullation and glycosylation to be introduced during chain elongation using the appropriate preformed protected amino acid building blocks. For SPPS there have been considerable advances in the length of peptides synthesized. This is partly due to the improvements in purity of the Fmoc building blocks, but mostly due to the success of applying pseudo-prolines and backbone protection to the synthesis of long peptides, overcoming the difficult sequence problem. Although the figure of around 50 amino acids is often given as the average target that can be routinely synthesized in publications, in practice, this figure serves as a practical reference point. However, many much shorter sequences are extremely problematic and synthetic success

is not guaranteed in such cases.[50] While requiring optimization, SPPS in batch and flow has been used to produce much longer synthetic peptides (>100 aa). Recent advancements by Hartrampf and Pentelute involve an automated fast-flow instrument, that makes peptide chains of up to 300 residues accessible.[59] However, such long peptides contain closely related by-products that cannot be easily separated. These limitations of peptide synthesis have driven the advancement of ligation chemistry, enabling the assembly of different peptide segments to produce larger proteins and peptides. The origin of the essential native chemical ligation (NCL) will be discussed in the next section.

1.5 Protein synthesis and semi-synthesis

Given the limitations of SPPS, a method capable of synthetically producing longer peptide segments would be highly desirable. In 1994 Stephen B.H. Kent *et al.* described the synthesis of proteins by native chemical ligation (NCL). A schematic overview of the reactions mechanism is shown in figure 3. The first step is the chemoselective reaction of an unprotected peptide-α-thioester with another unprotected peptide segment containing an amino-terminal cysteine residue to give a thioester-linked intermediate as the initial covalent product. Without change in reaction conditions, this intermediate undergoes spontaneous, rapid intramolecular reaction to form a native peptide bond at the ligation site.[60] The capability to conduct this reaction in aqueous solutions at neutral pH, employing entirely unprotected peptides, renders it a potent method for the total synthesis of proteins. It also laid the groundworks for integrating proteins produced through recombinant methods in a process known as protein semi-synthesis or expressed protein ligation.

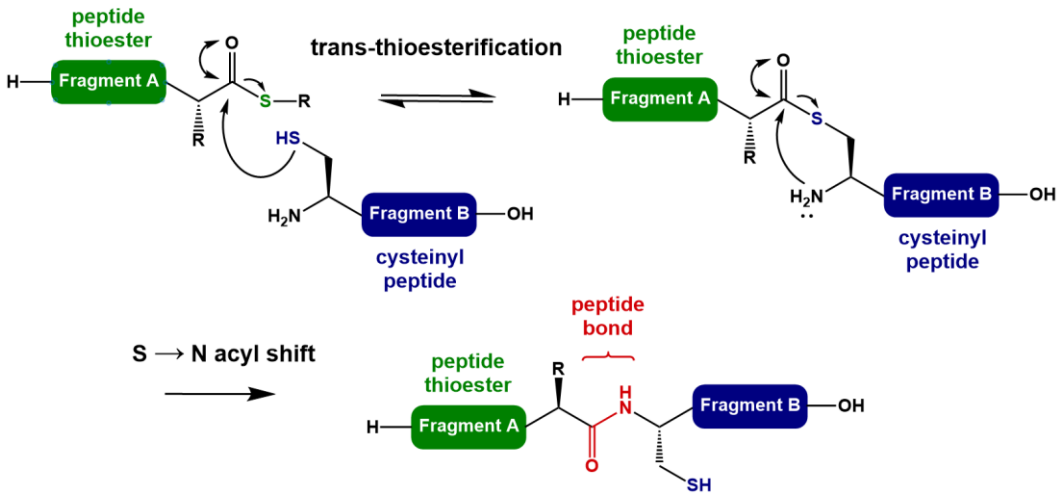


Figure 3 *Mechanism of the Native Chemical Ligation (NCL), involving the reaction between peptide thioester and cysteinyl peptide fragments via an initial trans-thioesterification followed by an S-N acyl shift to afford a native peptide (amide) bond.* [60]

In contrast to the total synthesis of proteins, where multiple synthetic fragments get assembled in a sequential ligation strategy, protein semi-synthesis is generally defined as a strategy that uses at least one or more recombinantly produced segments in the process of building the targeted molecule. Overall, this leads to a number of advantages in terms of types and sizes of biomolecules that can be accessed.[61] There are various strategies to build up a semisynthetic protein, depending on where one wants to introduce modifications.

The first one is to recombinantly express a protein bearing an N-terminal cysteine and ligate it to a synthetic peptide bearing a C-terminal thioester (*figure 4*). This entails the expression of a precursor protein, that consists of the desired polypeptide sequence coupled to a protease recognition motive. After expression and purification with the help of an affinity tag - typically a chitin binding domain (CBD) and/or a hexa- or heptahistidin tag (His_6/His_7), cleavage by a selective protease gives the requisite N-terminal cysteine for subsequent ligation.[61] The synthetic fragment, bearing the C-terminal thioester, is then prepared via specialized linkers or thioester precursors. A protocol by Zheng *et. al.* uses a 2-Cl-(Trt)-Cl resin an attaches hydrazine before continuing with the standard Fmoc-based SPPS protocol. After cleavage, the resulting peptide hydrazides are activated to the corresponding azides *in situ* via $NaNO_2$, which in turn can be converted to the final thioester by thiolysis under the use of a thiol catalyst like 4-mercaptophenylacetic acid (MPAA).[62] This elegant procedure is of special importance for the generation of synthetic thioester peptides, since the very labile thioester bond cannot withstand the basic conditions of Fmoc SPPS. This semi-synthetic strategy is mostly used to create proteins bearing modifications at their N-termini.

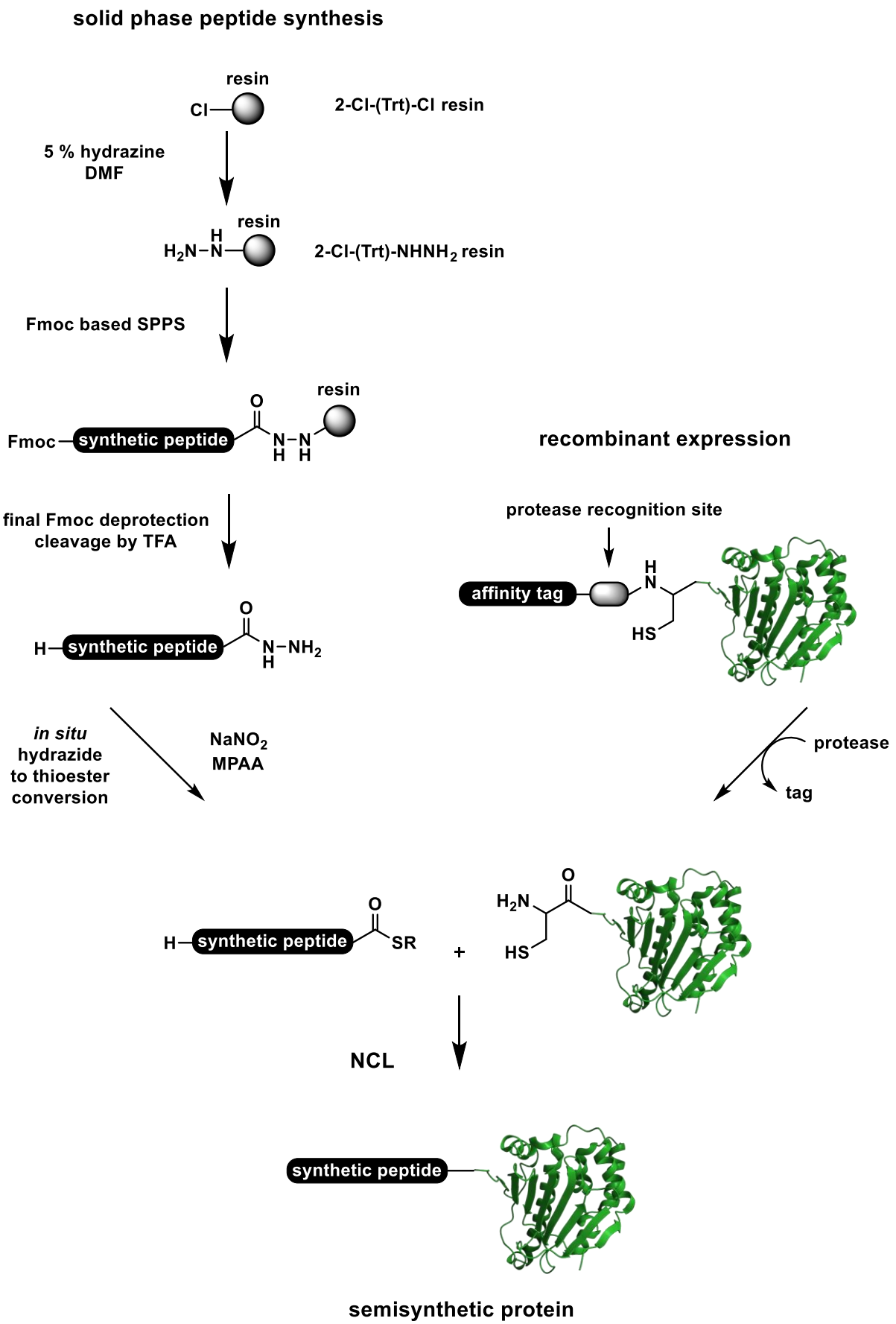


Figure 4 *Strategy 1: generation of modified α-thioester peptides, following the protocol of Zheng et al.* [69]*, using hydrazide to thioester conversion. Specific PTMs can be introduced by the use of modified*

amino acids during standard Fmoc-based SPPS. Expression of the N-terminal cysteine ligation partner is achieved with a protease / protease recognition site-system.

The second strategy of semi-synthesis can be seen as an expansion of the original idea, following the exact same principles but using expression systems that yield a C-terminal thioester moiety on the recombinant protein parts (*figure 5*). This method incorporates the use of protein splicing, a process in which a protein domain called “intein” (derived from internal or intervening protein) gets autocatalytically cleaved off of a given recombinant produced protein and subsequently forms a peptide bond between the two flanking regions called “exteins” (derived from external proteins). The targeted polypeptide chain is expressed as a fusion protein, where its C-terminus is attached to an intein-affinity tag domain. Upon reaction of the purified fusion protein with a suitable thiol (mostly sodium 2-mercaptoethanesulfonate / MesNa) the intein leaves the molecule, leading to the desired free thioester. The synthetic N-terminal cysteine fragment can then be very easily obtained by standard Fmoc SPPS protocol. This second strategy provides opportunities for the assembly of semi-synthetic proteins with modified C-termini.[61]

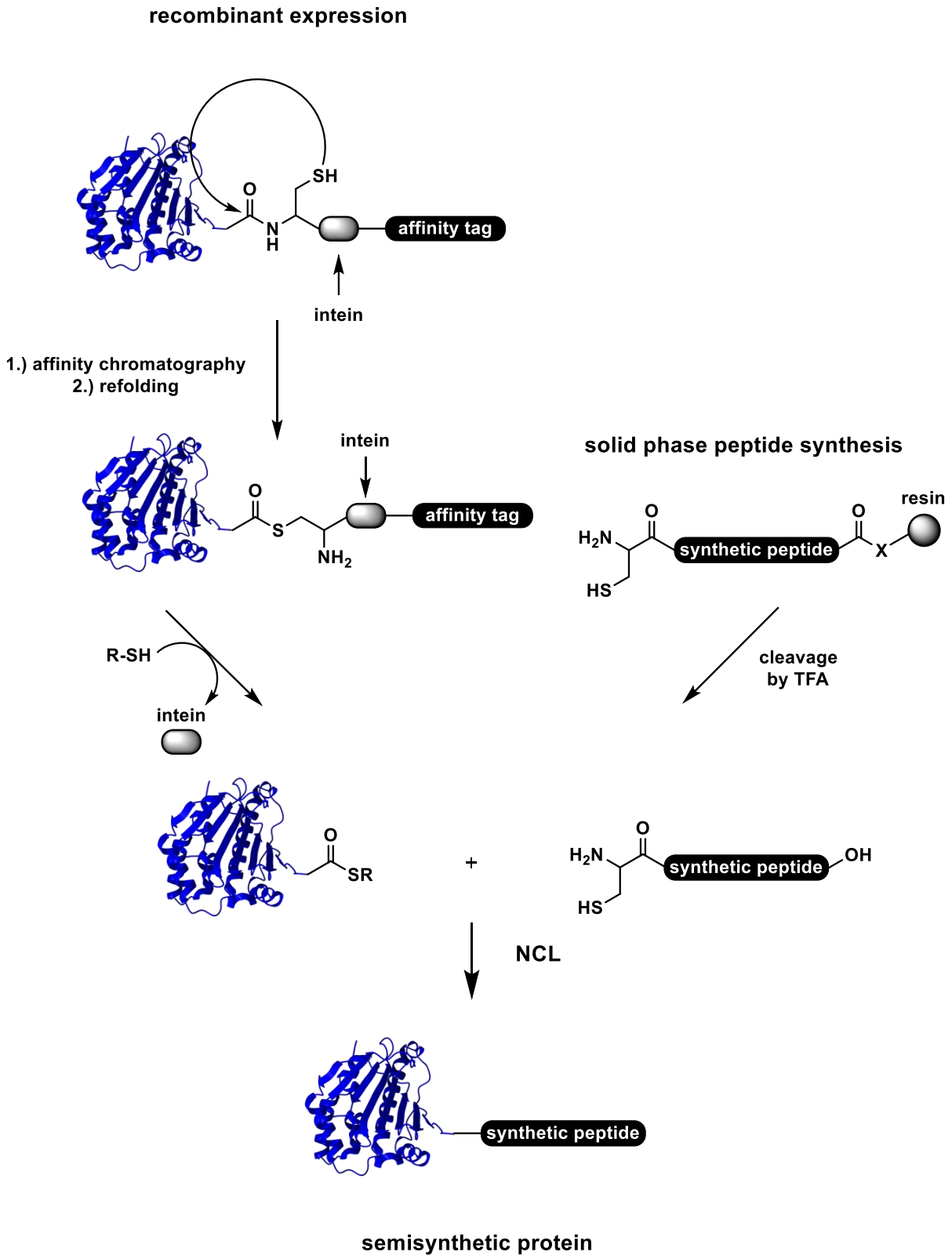


Figure 5 *Strategy 2: The desired α-thioester is generated by expression of the protein segment fused to an intein. Purification and reaction with a suitable thiol catalyst gives the final recombinant product. The N-terminal cysteine ligation partner is prepared by standard Fmoc SPPS strategy.*

1.6 Advancements in Ligation Strategies

The major drawback of the original ligation methodology is the requirement of a cysteine residue for the reaction to proceed. Cysteine is one of the least common proteinogenic amino acids, with 1.8 % abundance. This limitation led to the development of several extensions to the NCL method over the past decade. There has been the significant addition of post-ligation desulfurization chemistry, which allowed to change the requisite cysteine residue to the corresponding alanine after completion of the ligation. The first desulfurization chemistry in combination with NCL was reported in 2001 by Dawson *et. al.*, using Raney®-Ni and H_2 or Pd on Al_2O_3.[63] However, requiring large excesses of the metal catalyst, this method led to undesirable side reactions with side-chain moieties.[61] In 2007 Danishefsky *et. al.* reported a modified radical desulfurization protocol, facilitated by the radical initiator VA-044.[64] Figure 6 shows the proposed mechanism for the radical desulfurization of Cys to Ala that can be performed following an NCL reaction. As alanine residues are of higher native abundance (8.9 %) within proteins, it increases the number of potential ligation junctions within a target protein.

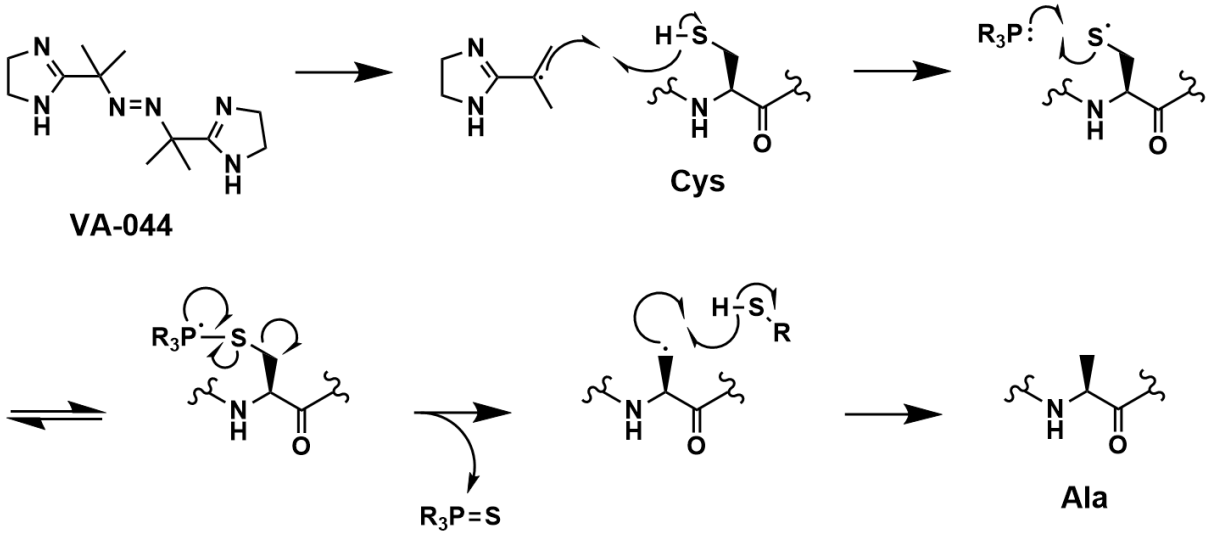


Figure 6 *Radical initiator VA-044 effects homolysis of the S-H Bond. The thiyl radical can then trap a phosphine additive, liberate the phosphoryl sulfide, and abstract a proton from a thiol H-atom donor present in the solution.*[59]

Another limitation of the original method lies in the prolonged reaction times of the thiol and thioester system and the required concentrations of peptide segment can be incompatible with aggregation prone peptides. By replacing the sulfur-based functionalities for selenium moieties, the nucleophilicity of the cysteine-containing fragment can be improved (through the inclusion of selenocysteine *(Sec)*) as well as the electrophilicity of the acyl donor fragment leading to improved reaction kinetics and required concentrations. The state-of-the-art method of the

diselenide-selenoester ligation (DSL), shown in figure 7, can be achieved by reacting an N-terminal selenocysteine peptide with a peptide bearing a C-terminal selenoester. This can be coupled with chemoselective deselenization to afford either the corresponding alanine or a serine residue at the ligation junction via a recently developed oxidative deselenization transformation. N-terminal selenocysteine peptides can be generated simply via SPPS where the suitable protected seleno-derived amino acid residue can be coupled under standard conditions as a 4-methoxybenzyl (Mob)-protected selenol moiety. C-terminal aryl selenoester peptides can be produced either with SPPS by functionalization at the C terminus of the N-terminal fragment as a selenoester, either on resin or in solution.[65] Or alternatively, selenoesters can be generated by hydrazinolysis of expressed intein fusion proteins. The resulting acyl hydrazides then give access to the corresponding selenoester via in situ selenoesterification.[66] The additive-free diselenide-selenoester ligation (DSL) potentially enhances reactivity and ligation rate by an enormous amount, completing ligations within minutes, even at the most sterically hindered junctions as shown in various examples.[59] *Lumbricin-1*, a 62-residue antimicrobial protein, was synthesized with a 52% yield by joining two component fragments through a single, challenging Pro–Pro ligation junction with completion of ligation reached in 5 to 45 minutes.[67] Another example is *Chorismate mutase*, an 83-residue enzyme from Mycobacterium tuberculosis, synthesized from two fragments through a single ligation (completion of ligation in 1 to 10 minutes), achieving a 57% overall yield, including the folding process.[68]

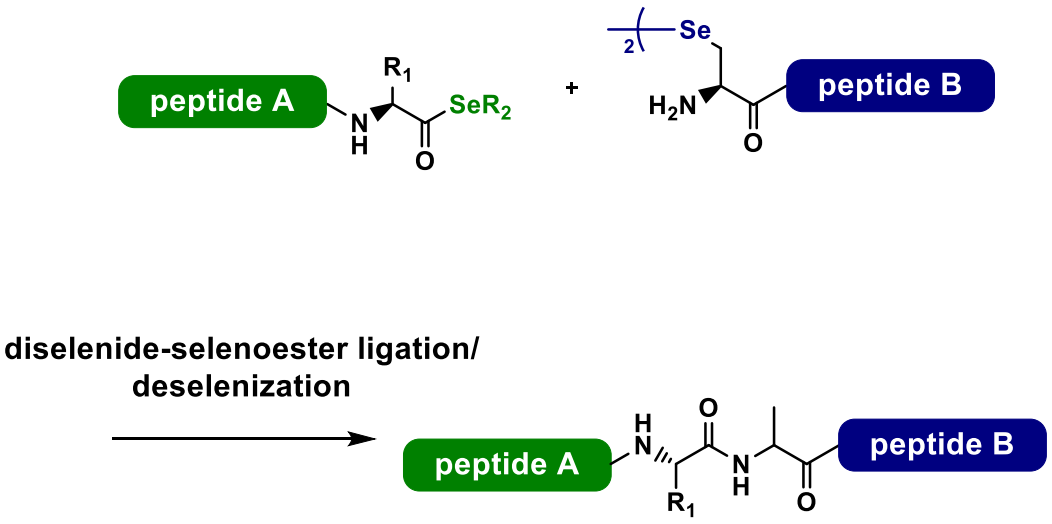


Figure 7 *Schematic overview of the additive-free diselenide-selenoester ligation, utilizing one peptide bearing a N-terminal SeCys residue and a peptide functionalized as a C-terminal selenoester. The N-terminal selenocysteine forms a peptide dimer with a N-terminal diselenide moiety due to its strong oxidation potential.*[60]

2. Aims of the project

This thesis focuses on semi-synthetic strategies to generate modified Hsp90alpha variants, in order to investigate possible influences of post translational modifications on the chaperones structure and function. The content of the thesis can be further split into two separate parts and synthetic strategies involved.

The first goal of the thesis is to build up Hsp90alphas N-terminal domain and implement a novel desulfurization protocol after ligation of synthetic and recombinant fragments (*figure 8. Strategy A*). This part includes the recombinant expression and purification of a Hsp90 N-terminal domain construct (Hsp90 C^8 ; residues 22-235), the ligation of C^8 to a synthetic polypeptide (residues 2-20) previously generated by SPPS, followed by desulfurization of the final product *in situ* with $NaBEt_4$, reported by Xuechen Li et. al.[70] If desulfurization proves to be successful, upscaling and refolding would make proper biochemical characterization of various NTD-variants possible.

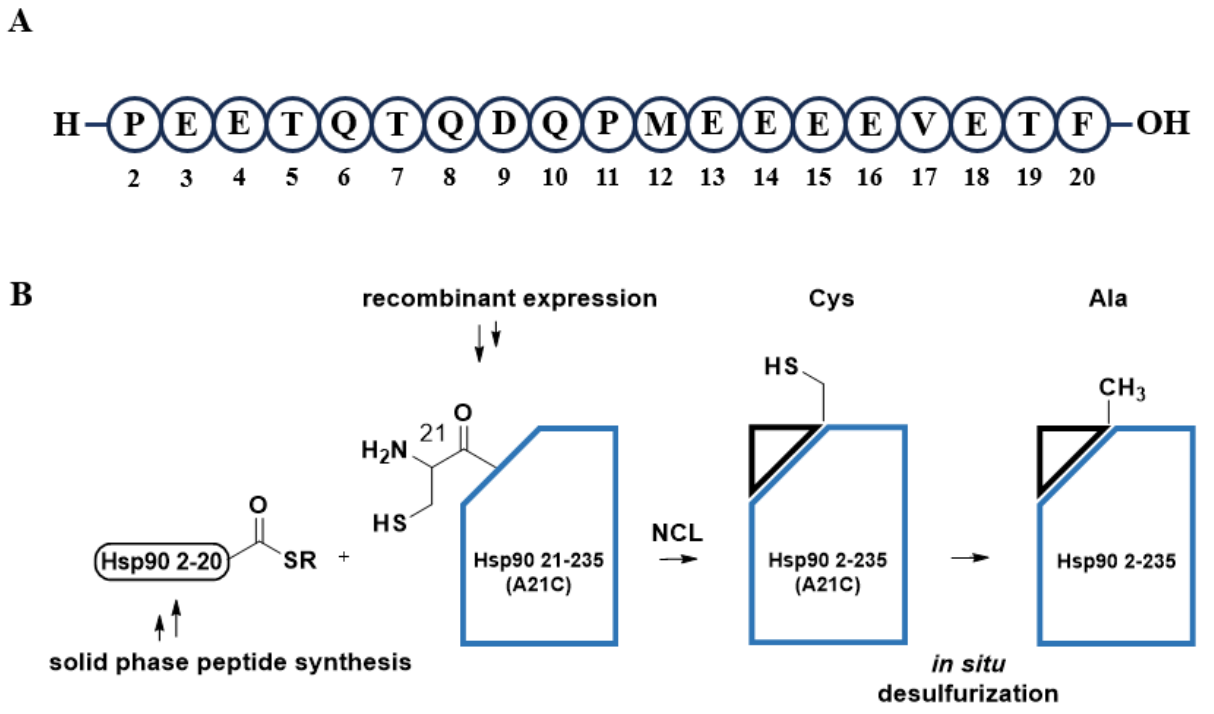
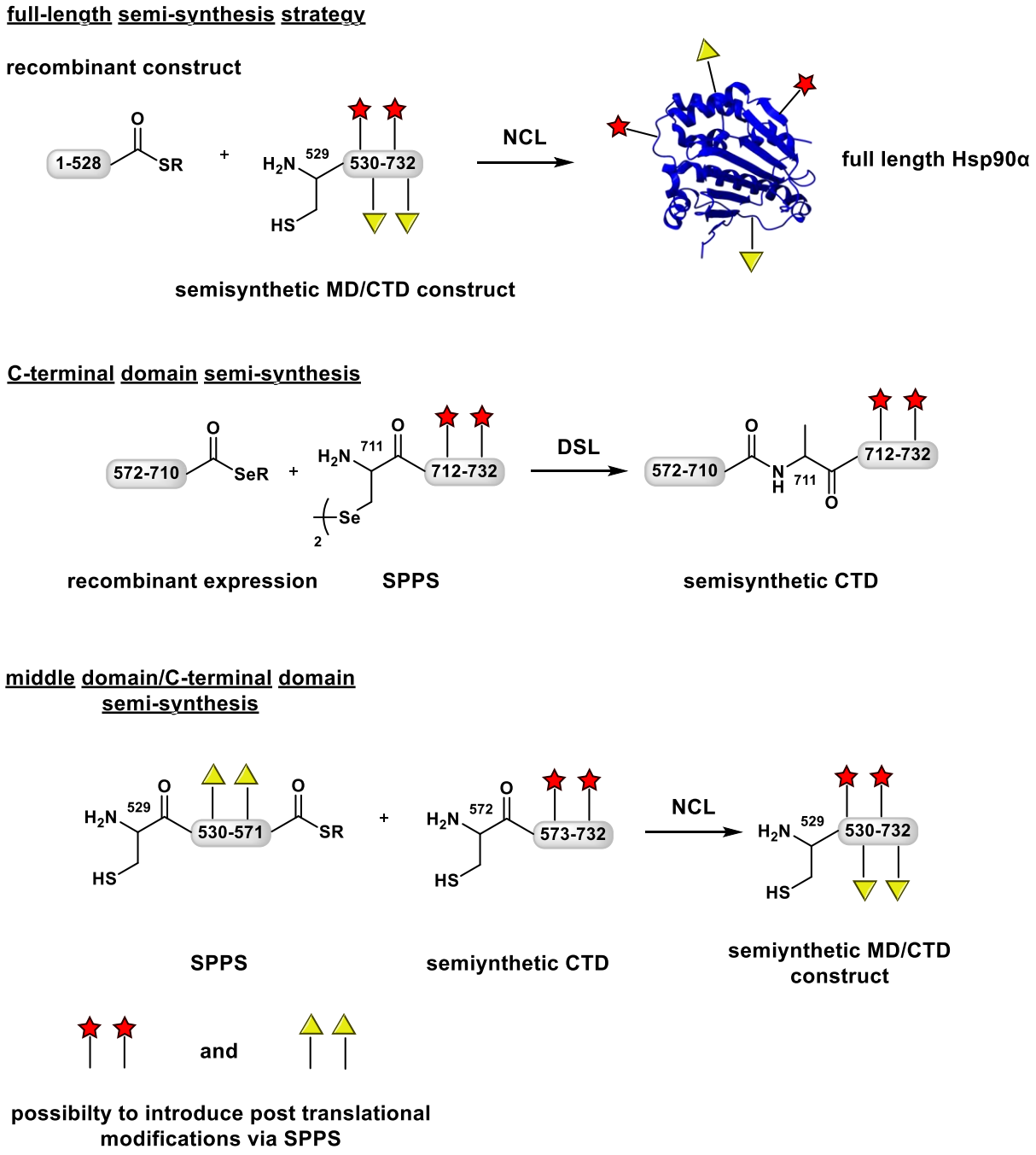


Figure 8. *Strategy A.* ***A)*** *amino acid sequence of unmodified human Hsp90α 2-20 NTD* ***B)*** *schematic overview of the semi-synthetic strategy and in situ desulfurization to create unmodified and modified Hsp90α NTD variants.*

The second goal of the thesis is to achieve the full-length semi-synthesis of an Hsp90alpha monomer using a sequential multiple fragment semi-synthesis strategy from C- to N-terminus (*figure 9. Strategy B*). Starting with plasmid transformation of Hsp90 construct 1-528 (Hsp90 C^{12}) as an intein-fusion protein into various competent cells and consecutive expression screening. Expressing the according construct in the most promising cell-strain and optimize purification and cleavage conditions of the novel recombinant segment, and finally perform a ligation reaction with the remaining residues 529-732 to achieve the first full-length semi-synthesis of an Hsp90alpha monomer. Preparation of fragment 529-732 is achieved via ligation of three different segments total. In a first diselenide selenoester ligation (DSL) the recombinant segment 572-710 is ligated to peptide 711-732 that is prepared by solid phase peptide synthesis (SPPS). In a second ligation, the semisynthetic construct 572-732 is ligated to another peptide 529-571 prepared by SPPS in a native chemical ligation to generate the semisynthetic MD/CTD construct 529-732. This strategy allows for the introduction of modifications in the middle domain of Hsp90 through the synthetic segment 529-571 with the aim of this thesis to produce an acetylated and succinylated lysine variant of full length Hsp90. Consequently, ligation to the semisynthetic CTD fragment 572-732 yields modified 529-732 building blocks, that can be used to synthesize the first ever full-length modified semi-synthetic Hsp90alpha monomer and thus give information about the influence of post-translational modifications that have not been studied yet.

A



B

Figure 9. *Strategy B.* ***A)*** *schematic overview of the semi-synthetic strategy for the generation of full-length Hsp90α unmodified and modified variants* ***B)*** *amino acid sequence of modified middle domain segment Hsp90 529-571 generated by SPPS. Amino acids highlighted in orange indicate post translational modification at this residue.*

3. Materials and Methods

3.1 Chemicals

All chemicals used were purchased from commercial suppliers and used without further purification unless stated otherwise. Solvents for solid phase peptide synthesis DMF and piperidine were purchased from Thermo Scientific Chemicals. The resin [2-Cl-(Trt)-Cl] was ordered from Novabiochem. All amino acid building blocks were purchased from Novabiochem (Darmstadt, Germany) and Iris Biotech (Marktredwitz, Germany) with the exception of building block Fmoc-Lys-(Mtt)-OH, produced by Carbolution (St.Ingbert, Germany). Activating agent HATU was purchased from Novabiochem and DIPEA was produced by Sigma Aldrich. Peptide cleavage chemicals TFA and TIPS as well as additives used for ligation reactions (MPAA, MTG, MesNa, TCEP) were form Sigma Aldrich. Diethylether used for precipitation was produced by VWR chemicals (Darmstadt, Germany). Purified water from a Millipore MilliQ Reference A+ water purification system was used to freshly prepare buffers. For recombinant expression media, tryptone and yeast extract from PanReac AppliChem ITW reagents were used. Sodium chloride was purchased from Carl Roth GmbH + Co. KG.

3.2 Peptide Synthesis

3.2.1 General procedures

3.2.1.1 general manual solid-phase peptide synthesis (SPPS) procedure

Manual solid phase peptide synthesis was carried out in 10 ml polypropylene syringes equipped with polypropylene frits via Fmoc (Fluorenylmethoxycarbonyl) strategy. Fmoc protected L-amino acids and standard sidechain protecting groups [Gln(Trt), Glu(tBu), Lys(Boc), Ser(tBu), Thr(tBu)] were used. The terminal cysteine residue was protected with an acetamidomethyl-group [Cys(Acm)]. Modified residues were introduced in the form of N-α-Fmoc-N-ε-acetyl-L-lysine [Fmoc-Lys(Ac)-OH] for peptide **C1** and N-Allyloxycarbonyl-lysine [Aloc-L-Lys(Fmoc)-OH] for peptide **C2**. The amino acid building blocks (2.5 eq.) were double coupled with hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU, 2.4 eq., 0.5 M in DMF) and diisopropylethylamine (DIPEA, 5 eq.) for 25 to 30 minutes. Fmoc deprotection was done with 20% piperidine in DMF (v/v) three times for 3 min. Two test cleavages of each peptide were performed during the synthesis, by extracting a small amount of resin. After the final coupling the resin was washed thoroughly with DMF and DCM and dried subsequently overnight under vacuum. Cleavage was facilitated by a cleavage cocktail consisting of 92.5 % TFA, 5% TIPS and 2.5% ddH_2O in 2 to 3 hours (5x resin volume). The cleaved peptide was

precipitated in cold diethyl ether (4x cleavage solution volume) and pelleted by centrifugation (4500 rpm, 5 min, 4°C), then dried under ambient conditions before dissolving in 50% water/acetonitrile (v/v, + 0.1% TFA) and lyophilizing.

3.2.1.2 Peptide purification and analysis

Crude peptides were dissolved in either 10 mL of ddH_2O or 10 mL 6 M Gdn·HCl in ddH_2O (pH 4.7) and filtered through a sterile 0.22 µm filter before purification by RP-HPLC on Shimadzu HPLC Purification System coupled with a Shimadzu LC-20AP module, combined with an SPD-10A UV-Vis detector and an FRC-10A fraction collector using a Macherey-Nagl C4 column (300-5-C4, 150 x 4.6 mm, 5 µm particle size). Elution of the peptides was achieved using a gradient of 5-45% solvent B (ACN + 0.05% TFA) in solvent A (ddH_2O + 0.05% TFA) in 45 minutes with a flow rate of 9 mL/min. Fractions were analyzed by electrospray ionization mass spectrometry (ESI-MS), pooled by purity and lyophilized. Purified peptides were characterized using ESI-MS in positive ion operation mode on Waters AutoPurification HPLC/MS System and analytical HPLC analysis on a Thermo Scientific Vanquish instrument using a gradient of 5-65% solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA), a Kromasil C18 column and flow rate of 1 mL/min over 30 minutes, recording chromatograms over 45 minutes and measuring the absorbance at 214 and 280 nm.

3.2.1.3 Synthesis of Hsp90 529-571 peptides C1 and C2

Synthesis of modified Hsp90 529-571 peptides C1 and C2 were performed on a 0.05 mmole scale on preloaded 2-Cl-(Trt)-$NHNH_2$ resin (hydrazine-2-Cl-Trt-resin**),** that is obtained by hydrazination (5% (v/v) NH_2NH_2 in 50 % DMF/DCM (v/v)) of the commercially available 2-Cl-(Trt)-Cl resin. Peptides synthesis was continued manually for **C1** and **C2**, after the scaffold (Hsp90 548-571-SR) of the two modified peptides **C1** and **C2** was prepared by Technician Manuel Felkl up to glycine (G) at position 548 via automated solid phase peptide synthesis on a CEM Liberty Prime microwave peptide synthesizer. Double couplings were performed for all manually coupled residues.

3.3 Protein expression and purification of Hsp90 constructs

3.3.1 Expression and purification of Hsp90 NTD construct

3.3.1.1 General procedures

The N-terminal domain (NTD) construct of Human Hsp90-alpha (UniProt entry: P07900; isoform 2) 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 previously via heat shock transformation. Overnight cultures with a final concentration of 1 µl/mL ampicillin and chloramphenicol were prepared out of fresh LB media. Rosetta II culture was added under semi-sterile conditions with a toothpick. Overnight cultures were 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). Cells were incubated at 37°C at 170 rpm until an OD_{600} ~ 0.7-0.8 and expression was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). Afterwards, the cells were harvested by centrifugation (5500 rpm, 14°C, 20 min), and the cell pellets were resuspended in 30-50 ml TBS (150 mM NaCl, 50 mM Tris, pH 7.5). The resuspended cell solutions underwent homogenization using an UltraTurrax (IKA GmbH, 79219 Staufen, Germany) and the resulting solution decanted into a beaker on ice. Cell lysis was either performed by sonication at 4°C for 10 min. with (pulse: 15/45 seconds; ampl. 60%) or by disruption (Constant Systems Limited, Northants United Kingdom) at 1.93 kbar (three cycles). The cell lysate was then clarified by centrifugation (22000 rpm, 45 min, 7°C), and samples from the supernatant and pellet were taken to ascertain the expression of the protein as inclusion bodies (IB) or in the soluble fraction using SDS-PAGE. The IB pellets were then solubilized in 6M Gdn·HCl, TBS, 0.1M methoxyamine hydrochloride at pH 4.2. Final purification was performed via RP-HPLC on a Macherey-Nagel C4 column at 60°C using a TL105 HPLC column heater from Timberline Instruments (Boulder, Colorado) coupled with a Shimadzu LC-20AP module, combined with an SPD-10A UV-Vis detector and an FRC-10A fraction collector. Collecting fractions at a threshold of 70 mV from 25 to 50 minutes at a fraction size of 2.5 ml. Peak fractions were submitted for LC-MS analysis and pooled accordingly. Pools of interest were combined with 50% ddH_2O/acetonitrile (v/v, + 0.1% TFA) and lyophilized consequently.

3.3.1.2 Expression and purification of Hsp90-NTD 21-235-His_6 ($Hsp90^{C8}$) construct

Expression and purification conditions were optimized by screening different expression temperatures (18 and 24°C), expression duration (5 and 22 hours), pre-purification methods for the IB pellet (varying number of washes with 2M Urea TBS), solubilization and dialysis conditions. The best expression conditions are at 18°C for 22 hours, with induction at an OD_{600}

~ 0.7. Pre-purification of inclusion body pellets trough washes with 2M Urea (10 ml/g pellet) TBS is non-essential. Solubilization is best in 6M Gdn·HCl, TBS at pH 4.2 for 24 to 72 hours. Dialysis to 4M Urea, 100mM Tris, 50mM NaCl at pH 8 and 4°C for 72 hours is optional. After centrifugation of the solubilized IB pellet at 22.000 rpm (7°C) for 35 min. and filtering by syringe filter (Chromafil®Xtra PES 20/25 ; pore size 0.2 µm), the supernatant was purified by loading it onto a C4 column at 60°C. Final purification via RP-HPLC was optimized to the following gradient parameters: 5-35% solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 5 min., and 35-75% solvent B in solvent A in 40 min. (15 min. desalt) at a flow rate of 9 ml/min.

3.3.2 Expression and purification of Hsp90 1-528-*Mxe GyrA*-His_7-CBD construct

3.3.2.1 General procedures

Human Hsp90-alpha 1-528 was ordered from BioCat GmbH (69120 Heidelberg, Germany) (UniProt entry: P07900; isoform 2) in a pET-21a+ bacterial expression vector including a C-terminal fused *Mxe GyrA*-His_7-CBD (CBD, chitin binding domain) intein and purification tag. Transformation into E. coli BL21 (DE3) Gold, XL1 (Blue) and Rosetta 2 (DE3) competent cells was performed via heat shock method, described in "*Transformation and Expression of Hsp90 1-528-Mxe GyrA-His7-CBD construct*". Small scale test expressions (200 mL LB medium) were performed at 25°C and 37°C in BL21 (DE3) Gold and Rosetta 2 (DE3) to identify optimal expression conditions and strains. Only E. coli BL21 strains were further used for large scale expressions. RP-HPLC was performed on a Macherey-Nagel C4 column at 60°C using a TL105 HPLC column heater from Timberline Instruments (Boulder, Colorado) coupled with a Shimadzu LC-20AP module, combined with an SPD-10A UV-Vis detector and an FRC-10A fraction collector. Peak fractions were submitted for LC-MS analysis and pooled accordingly. Pools of interest were combined with 50% ddH2O/acetonitrile (v/v, + 0.1% TFA) and lyophilized consequently.

3.3.2.2 Transformation and Expression of Hsp90 1-528-Mxe GyrA-His_7-CBD (Hsp90^{C12}) construct

The plasmid was dissolved in 25 µl ddH_2O and 100 µl precooled aliquots of the competent cells were incubated with 5 µl (Rosetta 2) and 3 µl (BL21(DE3)Gold, XL1) of plasmid solution. The competent cells were incubated for 10 min on ice, and heat-shocked for 90s at 42°C afterwards. After placing the cells on ice for 5 min, 1 ml of sterile LB-media was added to the cells before incubating at 37°C for 45 min with gentle shaking. The cells pellets were harvested by centrifugation and most of the supernatant removed, leaving 50-100 µl for resuspension. The resuspended cell solutions were spread over prepared agar plates including the corresponding

antibiotics under semi sterile conditions. The plates were incubated at 37°C overnight under humid conditions without shaking. Single cell colonies were picked and placed in a test tube with 4 ml sterile LB medium containing 100 µg/ml ampicillin (BL21, XL1) or 100 µg/ml ampicillin + 30 µg/ml chloramphenicol (Rosetta 2). The cell cultures were prepared by shaking overnight at 37°C at 200 rpm and used for test expressions and to prepare 20% glycerol stocks that were stored at -80°C. To prepare the glycerol stocks 900 µl of the overnight culture were mixed with 600 µl of 50% glycerol. The mixture was then vortexed and freezed at -80°C.

Large scale expressions were performed by inoculating 2 L of LB medium containing 1 µl/ml ampicillin with 50 ml overnight culture, followed by growth at 37°C to an OD_{600} of ~ 0.7. Protein expression was induced with 1 mM IPTG at 37°C for 4 to 5 hours at 170 rpm. The cells were collected by centrifugation (8000 rpm, 10°C, 15 min). Cell pellets were resuspended in TBS buffer and the solution homogenized using an UltraTurrax homogenizer. Cells were lysed by sonification at 4°C for 10 min. with (pulse: 15/45 seconds; ampl. 60%) and the lysate cleared by centrifugation (22000 rpm, 1 h). The obtained pellets contained the protein as inclusion bodies as well as the supernatant in lower concentration. The fusion protein was solubilized using 6 M Gdn·HCl, TBS pH 7.5 overnight.

3.3.2.3 Intein cleavage and purification of Hsp90 1-528-MxeGyrA-His_7-CBD (Hsp90 ^{C12}) construct

After centrifugation of the cell lysate, pre-purification of the resulting supernatant was optimized by testing different NiNTA column systems. Firstly, pre-purification of the supernatant was tested on an ÄKTAPurifier at 4°C and a flow rate of 1 ml/min, using a 20 to 500 mM imidazole 6 M Gdn·HCl, TBS pH 7.5 buffer system. Fractions of interest were pooled, and small aliquots taken for SDS samples. In a second approach, the supernatant was loaded onto a NiNTA-batch system consisting of two 20 ml syringes, capped with a polypropylene filter, and filled with Ni-Sepharose resin (40 mg/mL binding capacity). The batch system was prepared by removing the EtOH solution out of the syringes, filled with the resin. Then, the resin was washed with ddH_2O for 10 minutes, followed by two washes with 6 M Gdn·HCl, TBS buffer pH 7.2 – 7.5 for 10 minutes each. The sample was split into equal parts and loaded onto the resin, then incubated for at least two hours. Therefore, the sample was loaded in 6 M Gdn·HCl, TBS pH 7.5. Washes were performed with 10 mL 40 mM imidazole in 6 M Gdn·HCl, TBS pH 7.5 for each syringe three times total for 10 minutes. Elution was achieved via 500 mM imidazole buffer in 6 M Gdn·HCl, TBS pH 7.4 for 35 to 40 minutes. SDS samples were taken after every step of the procedure (load, flow through, wash(es), elution).

Subsequently, MesNa cleavage was optimized by dialyzing the fractions of interest against different Urea concentrations (5, 4, 3, 0 M) for a MesNa cleavage condition screen. MesNa cleavage was first performed by dropwise 1:1 dilution in a 20 ml Falcon tube, using 1M MesNa, TBS pH 7.5 solution at RT. SDS samples were taken for over 48 hours. In a second approach, dialysis to a concentration of 2 M Urea in TBS with a volume of 3 L at pH 7.4 was performed overnight. The MesNa Cleavage was carried out by adding 500 mM MesNa as solid to the stirring solution over a time frame of approximately 10 minutes at pH 7.5 before finalizing the cleavage after over 24 hours.

The final purification step was optimized by trying a whole array of different purification methods on the resulting solution after MesNa cleavage. In a first approach, centrifugation of the cleavage solution led to pellet and supernatant which were purified separately. The supernatant was loaded for a "reverse" Ni-NTA on an ÄKTAPurifier at a flow rate of 1 ml/min, using a 20 to 500 mM imidazole 6 M Gdn·HCl, TBS pH 7.5 buffer system at RT. SDS samples were taken of the load, flow through and the eluted peak. The pellet from the MesNa cleavage was dissolved in 6M Gdn and analyzed on LC-MS and analytical HPLC on a C4 waters column with a gradient of 5-65 % solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 45 min. at 60°C and a flow rate of 1 ml/min. SDS samples of the eluted peaks of interest were taken. In a second strategy, purification of solubilized pellet and supernatant was performed via RP-HPLC. The following parameters were used: 5-35% solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 5 min., and 35-65% solvent B in solvent A in 30 min. (15 min. desalt) at a flow rate of 9 ml/min. Collecting fractions at a threshold of 35 mV from 20 to 45 minutes at a fraction size of 3 ml. The third purification procedure finalized the previous MesNa cleavage with a “reverse” NiNTA batch system, using the same two syringe system that was used for the first NiNTA-batch system. SDS samples were taken analogue to the first NiNTA after every step of the procedure. The flow through was collected and stored at 4°C until further purification. Final purification of the flow through of the "reverse" NiNTA was then performed via RP-HPLC and size exclusion chromatography (SEC). For SEC, the load of the “reverse” NiNTA was concentrated under the use of Amicon®Ultra centrifugal filters. Concentrated to 5 mL total, the solution was loaded onto a SEC/SuperdexTM 75 (HiLoadTM 16/60 ; prep grade) column on an GE ÄKTA Purifier 10 FPLC System w/ UV-900 Detector at 4°C in a first purification run. Afterwards, the resulting fractions were collected with a fraction collector assembled to the purification system, analyzed, and pooled accordingly. In a second purification round the further concentrated solution with a total volume

of 500 µL was loaded onto a smaller SEC/ SuperdexTM 200 (10/300 GL) column on the same system as the first purification at 4°C.

The best procedure for the intein cleavage and purification of $Hsp90^{C12}$ is the batch NiNTA system with two syringes. MesNa cleavage was the most efficient with 2M Urea at pH 7.5, adding the MesNa as solid with a final concentration of 500mM over approximately 10 minutes. "Reverse" batch NiNTA and subsequent, final purification of the resulting load by RP-HPLC with the following gradient led to starting material for the subsequent ligation reactions: 5-20% solvent B (ACN + 0.08% TFA) in solvent A (ddH_2O + 0.1% TFA) in 5 min., and 20-75% solvent B in solvent A in 30 min. (5 min. desalt) at a flow rate of 9 ml/min. Collecting fractions at a threshold of 35 mV from 22.5 to 41 minutes at a fraction size of 3 ml.

3.4 Ligation reactions

3.4.1 General procedures

Semipreparative purification of the ligation reactions was performed via RP-HPLC on a C4 Waters analytical column using a Shimadzu system with the following modules: LC-20AP, FCV-200AL, CBM-20CL, SBD-10 UV Vis detector, FRC-10A fraction collector).

3.4.2 NTD semi-synthesis

EPL of $Hsp90^{C8}$ to peptide thioester OG_{13} ($Hsp90^{C8}$ = Hsp90 21-235) ; (OG_{13} = 2-20-SR)

Small scale expressed protein ligations of $Hsp90^{C8}$ to peptide thioester Hsp90 2-20 **OG_{13}** (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 under 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. Continuing with the desulfurization, the pH was adjusted to 3.0 and ethyl acetate extraction was performed three times when 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 $NaBEt_4$ was added to a final concentration of 100mM, according to the procedure of Xuechen Li by preparing a 1 g/ml $NaBEt_4$ stock solution in ddH_2O. 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 (ddH2O + 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 gradient of 5-75% solvent B (ACN + 0.08% TFA) in solvent A (ddH2O + 0.1% TFA) in 30 min., detecting wavelengths 214 and 280 nm (Dionex).

3.4.3 Hsp90 MD-CTD (529-732) and Hsp90 full-length semi-synthesis

3.4.3.1 DSL of Hsp90^{C9} to selenocysteine peptide Sec$_{001}$ (Hsp90^{C9} = Hsp90 572-711) ; (Sec$_{001}$ = 712 - 732)

To convert the recombinant Hsp90^{C9} peptide hydrazide to a selenoester, 1.5 ml of the selenoester conversion buffer (SerCon) was prepared, containing 6 M Gdn, 200 mM HEPES, 100 mM TCEP and 50 mM DPDS. The buffer was sonicated for 30 minutes and degassed under argon for another 10 minutes. The protein was dissolved in 1 ml of the freshly prepared SerCon-Buffer (final concentration 1.3 mM) and the pH adjusted to 1.5. Selenoesterification was initiated by the addition of 3.7 µl acetylacetone (30 equiv.) and the reactions progress was monitored for the following 30 minutes by LC-MS. For the diselenide-selenoester ligation, a peptide buffer containing 6 M of Gdn, 200 mM of HEPES and 5 mM DPDS was freshly prepared by sonicating and degassing, analogue to the SerCon-Buffer. Afterwards the selenocysteine containing peptide fragment 711-732 (1.2 equiv.) was dissolved in 2 ml of the peptide buffer, combined with the selenoester – conversion solution and adjusted to a pH of ~ 3-4. The reaction was monitored after 1 hour by LC-MS. Afterwards, any excess of DPDS was extracted with Hexane. This was done by the addition of equal volumes of Hexane six times total before the final deselenization step. For the deselenization of the ligated product, 3 ml of deselenization-buffer consisting of 6 M of Gdn, 200 mM of HEPES, 250 mM TCEP and 25 mM DTT was added to the ligation mixture via 1:1 dilution. The solution was adjusted to a pH of ~ 7.2 and the reaction was incubated for three hours at RT before shock freezing at – 80 °C. After thawing and dilution of the reaction mixture using 6M Gdn in ddH_2O (pH 4.2), semipreparative reversed phase chromatography was used for purification (5ml/min., 5-30% B in 5 min., 30-60% B in 40 min.) (Shimadzu LC-20AP, FCV-200AL, CBM-20CL, SBD-10 UV Vis detector, FRC-10A fraction collector), fractions of interest were analyzed by LCMS and HPLC (Vanquish Horizon, Thermo Scientific).

3.4.3.2 EPL of Hsp90^{C12} to peptide thioester CTD$_{Lig01}$ (Hsp90^{C12} = Hsp90 1-528) ; (CTD$_{Lig01}$ = 529-732-SR)

Small scale expressed protein ligation of Hsp90^{C12} to peptide thioester Hsp90 529-732 CTD$_{Lig01}$ (1.5 equiv.) was performed by dissolving the recombinant protein in the ligation buffer consisting of 6M Gdn·HCl, 200 mM NaPi, 50 mM TCEP and 100 mM MPAA to a concentration of 0.7 mM. Subsequently the pH was adjusted to 7.2, and the solution was added to the peptide thioester under 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 (2 µl of sample diluted in 20 µl of 6M Gdn·HCl) and SDS-PAGE (2 µl of sample incubated in 18 µl of EtOH 98%). After completion of the reaction, ethyl acetate extraction was performed three times, the reaction mixture diluted with equal volumes of 6M Gdn·HCl in ddH_2O and β-mercaptoethanol (20%, v/v) and purified on a C4 Waters column using a gradient of 5-65% solvent B (ACN + 0.08% TFA) in solvent A (ddH2O + 0.1% TFA) in 30 min. at a flow rate of 1 mL/min at 60°C (Dionex). After LCMS measurements, the fractions of interest were pooled, diluted with ddH_2O:ACN (50/50), frozen at -80°C and lyophilized consequently.

3.4.3.3 Ligation of semisynthetic Hsp90 CTD 572-732 to acetyl lysine peptide 529-571 (AcLys546) CTD$_{N03}$

This ligation reaction of the semisynthetic CTD and the synthetic acetyl lysine peptide CTD$_{N03}$ was performed in a ligation buffer consisting of 6M Gdn·HCl, 150 mM NaPi, 100 mM TCEP and 100 mM MPAA. The semisynthetic building block was dissolved in the degassed ligation buffer to a concentration of 1 mM. Afterwards, the synthetic peptide CTD$_{N03}$ (1.5 equiv.) was added, and the pH adjusted to 7.2. The reaction was monitored by LC-MS and SDS-PAGE for 2 hours. Subsequently, the reaction mixture was quenched by the addition of an equivalent volume of 6M Gdn·HCl in ddH_2O and β-mercaptoethanol (20%, v/v). The solution was further diluted, filtered, and loaded on C4 semi-prep column.

4. Results and Discussion

4.1 Expression and purification of Hsp90 constructs

4.1.1 Expression and purification of Hsp90 NTD construct C^8 (21-235 (A21C))

Expression level monitoring by SDS-PAGE revealed strong expression levels of the construct ($MW_{calc.}$ 24889 Da) as a single band at a molecular weight of ~ 30 kDa. A schematic representation of construct C^8 is depicted in figure 10. After cell lysis, analysis of supernatant and pellet showed that the construct is exclusively expressed in inclusion bodies. Overall expression purity was acceptable throughout all expressions.

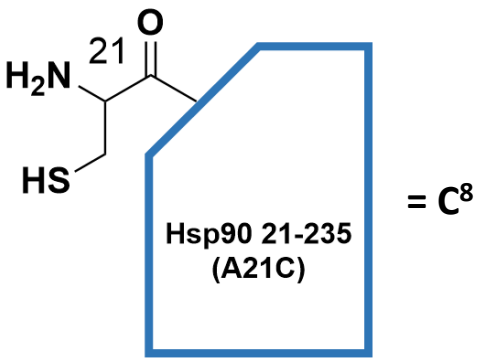


Figure 10. *Schematic representation of construct C^8. The construct comprises the N-terminal domain with its residues 21-235 and a cysteine at position 21, replacing the native alanine at this position.*

The screening of different expression temperatures and expression durations showed that late induction with 1mM IPTG at an OD of approximately 1 and expression at 25°C for 5 hours yields big inclusion body pellets with gel-like consistency. Subsequent solubilization with TBS buffer containing 6M Gdn·HCl (pH 7.5) over night (20 hours) and over the weekend (72 hours) as well as final purification turned out to be challenging under these conditions. Dialysis parameters can be seen in table 1.

Table 1. *Overview of the dialysis parameters for expressions 1-3. For Nr. 3 no dialysis was carried out. The other expressions were dialyzed in a buffered TBS system at pH 8. (RT = room temperature; approx. 20-22°C)*

Nr.	From →	To	pH	Method	Temperature	Additive
1	6M Gdn·HCl	4M Urea	8	Dialysis	4°C	0.1M methoxyamine hydrochloride
2	6M Gdn·HCl	6M Urea	8	Dialysis	4°C	-
3	6M Gdn·HCl	-	4.2	-	RT	-

The addition of methoxyamine hydrochloride during dialysis can be a precautionary measure to maintain the integrity and functionality of recombinantly expressed protein. Specifically, methoxyamine reacts with aldehyde or ketone groups to form stable oximes, which can prevent unwanted side reactions such as protein crosslinking or aggregation. After earlier induction with 1mM IPTG at an OD ~ 0.7 and expression at 18°C for 5 hours, the protein expressed in inclusion bodies was solubilized with TBS buffer containing 6M Gdn·HCl (pH 7.5). After dialysis to 4M Urea and consequent filtering through a syringe filter, the protein solution was purified by preparative RP-HPLC, and fractions of interest were collected and pooled accordingly before submitting for ESI-MS and HPLC analysis (*figure 11*). Low yield of final product was observed due to significant loss of the desired protein trough bad recovery during purification via preparative RP-HPLC. However, solubility issues were not as troubling, as with previous expression parameters at 25°C. This suggests, that the expression of construct C^8 might be at its optimum at 18°C.

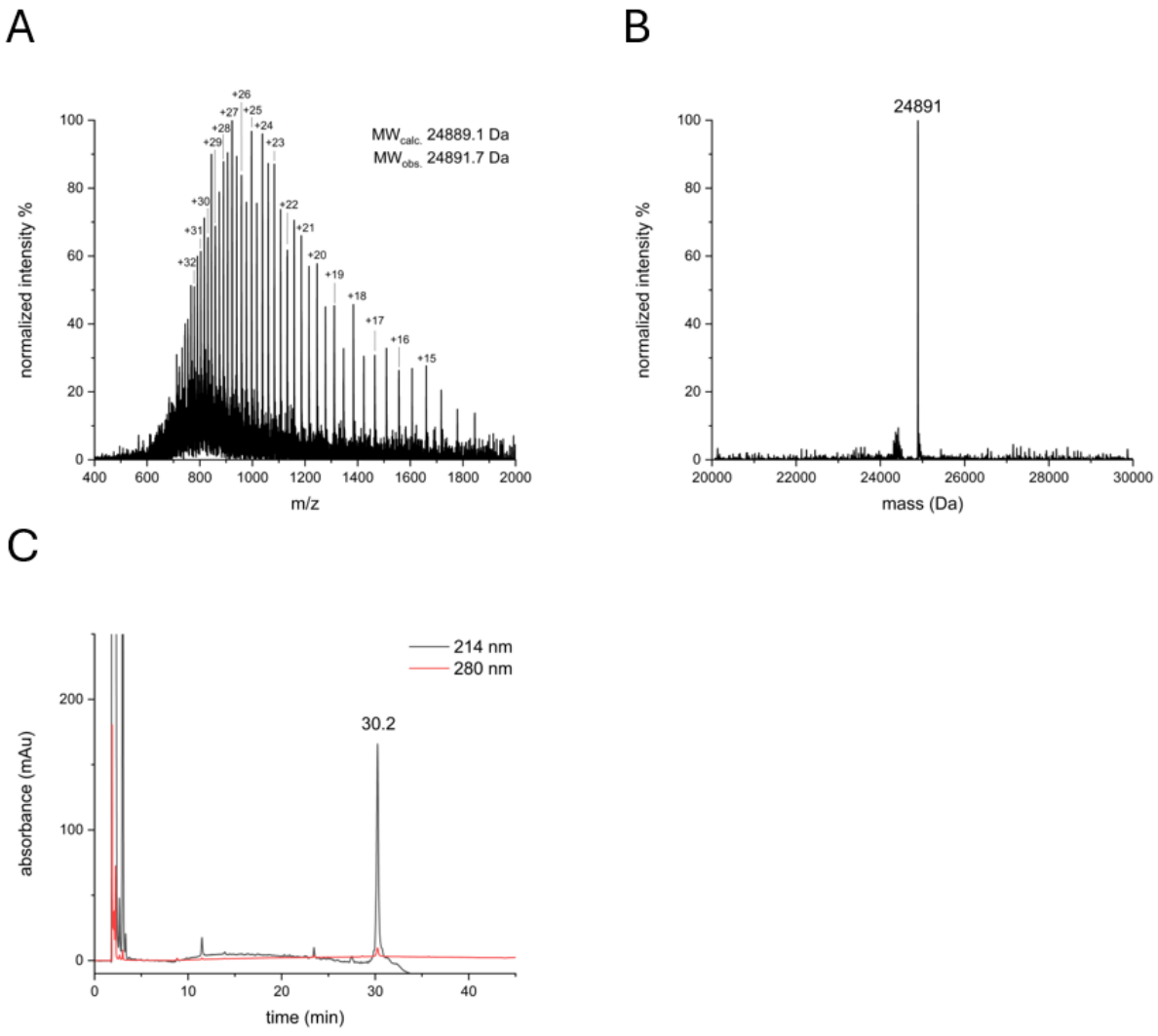


Figure 11. final analysis construct C^8: *Characterization of NTD construct Hsp90 21-235 (A21C).* ***A)*** *ESI-MS spectrum of the purified main pool shows that the observed mass ($MW_{obs.}$ 24891.7 Da) is in good agreement with the calculated mass ($MW_{calc.}$ 24889.1 Da).* ***B)*** *Deconvoluted mass spectrum of purified Hsp90 21-235 (A21C).* ***C)*** *HPLC chromatogram of purified Hsp90 21-235 (A21C).*

Optimized parameters were found for the induction with 1mM IPTG at an OD ~ 0.7 and expression at 18°C for 22 hours. After dissolving the inclusion body pellets in 6M Gdn·HCl, TBS at pH 4.2 for 24 to 72 hours, no dialysis was performed. The resulting solution was filtered through a syringe filter and directly loaded onto a preparative RP-HPLC system for purification. Average expression yield throughout purification of the inclusion body pellets with this optimized procedure was around 2 mg/L LB media.

Discussion of Expression and purification of Hsp90 NTD construct C^8 (21-235 (A21C))

Designing a semi-synthetic strategy for a desired protein under the use of native chemical ligation is delicate, and a lot of requirements and criteria need to be considered. First and foremost, a suitable ligation site needs to be selected, so that the synthetic fragment is still within reach of the standard Fmoc-SPPS method lengthwise. Secondly, the manually synthesized thioester peptide (acyl-donor fragment) should not include residues that are prone to transthioesterification towards the sidechain (Asp, Glu), electronically deactivating because the amino group is part of a five-membered ring, which reduces its availability for resonance or hydrogen bonding interactions compared to other amino acids (Pro) or sterically hindered beta-branched amino acids (Ile, Val) at the C-terminus to prevent isopeptide bond formation or low ligation yields.[66] Furthermore, a cysteine residue is needed to perform the classic reaction of the native chemical ligation. Since there are no cysteine residues to be found in the native sequence of Hsp90α N-terminal domain, an alanine residue at position 21 of the NTD was mutated to cysteine and chosen as ligation site. Consequent desulfurization after ligation would yield the native sequence with alanine. The synthetic fragment consists of Hsp90α residues 2-20, excluding the methionine at position 1 to resemble the natural occurring process of methionine cleavage at the N-terminus of proteins by methionine aminopeptidases present in eukaryotic cells. Finally, one would strive to keep the number of ligation sites as low as possible, not only to reduce effort and time consumption but also to eliminate sources of error to an absolute minimum.

Literature about the expression and purification of Hsp90 N-terminal domain (NTD) in its α and β isoforms, typically details the purification of the protein from the soluble fraction of the cell lysate, avoiding challenges related to refolding or precipitation. The expression and

purification of truncated construct C^8 proved to be challenging mainly because the designed construct is exclusively expressed as inclusion bodies and therefore solubility issues not only in 6M Gdn·HCl but also after dialysis at different urea concentrations have to be dealt with. This highlights the importance of the naturally highly conserved N-terminal domain of Hsp90α for proper solubility. Thus, removing a relatively short motif of 19 amino acids can lead to radical changes on the physical-chemical properties of the protein in terms of solubility. The inclusion body pellets also contain other expressed proteins (e.g. chloramphenicol esterase) which might add to the fact that the IB pellets could not be fully dissolved. Even though screening different expression and purification parameters helped in addressing these major issues, the expression, purification, and subsequent refolding of construct C^8 could not be fully optimized in the terms of this work. The main loss of final product is observed during the final purification step via RP-HPLC due to strong interactions with the stationary phase and in turn poor recovery. Nevertheless, the yielded recombinant protein after expression was enough to carry on with small scale test ligations (*"4.3.1 EPL of Hsp90α C8 to peptide thioester Hsp90 2-20 (pThr5/pThr7) (OG13) and in situ desulfurization"*) and focus on implementing the subsequent, novel *in situ* desulfurization protocol.

4.1.2 Transformation and expression of Hsp90 construct C^{12} (1-528-SR)

The construct C^{12} of Human Hsp90-alpha was cloned into a pET-21a+ bacterial expression vector including a C-terminal Mxe GyrA-His_7-CBD fusion intein. A schematic representation of the construct C^{12} is depicted in figure 12.

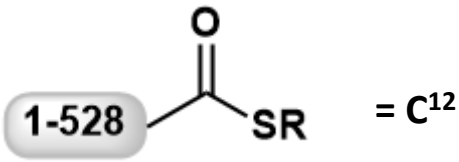


Figure 12. *Schematic representation of construct C^{12}. The construct comprises the whole N-terminal domain and parts of the middle domain with its residues 1-528.*

Transformation into E. coli BL21 (DE3) Gold, XL1 (Blue) and Rosetta 2 (DE3) competent cells was performed via heat shock method, as described above (see *transformation of Hsp90 1-528-MxeGyrA-His_7-CBD*). First, we started with small scale test expressions (200 mL LB medium) at 25°C and 37°C in BL21 (DE3) Gold and Rosetta 2 (DE3) to identify optimal expression conditions and strains, as well as verifying successful expression of the construct itself. Expressions were monitored by SDS-PAGE, depicted in figure 13. Samples were taken directly before induction with IPTG at OD_{600} ~ 0.7 (t=0), and during the expression (t=1-4h for

37°C)(t=1-4h and overnight *(o.n.)* for 25°C). Only E. coli BL21 strains were further used for large scale expressions, due to strong expression of the desired construct and satisfactory expression purity. Optimal conditions for large scale expressions (2L LB medium) were found to be at 37°C for 4 hours upon induction with 1mM IPTG at an OD of approximately 0.7. After harvesting, the cells were subsequently disrupted by sonification, the lysate cleared by centrifugation, and the resulting pellets were solubilized in 6 M Gdn·HCl, TBS pH 7.5 overnight, according to the description in the section *"Expression and purification of Hsp90 full-length constructs" "General procedures"*.

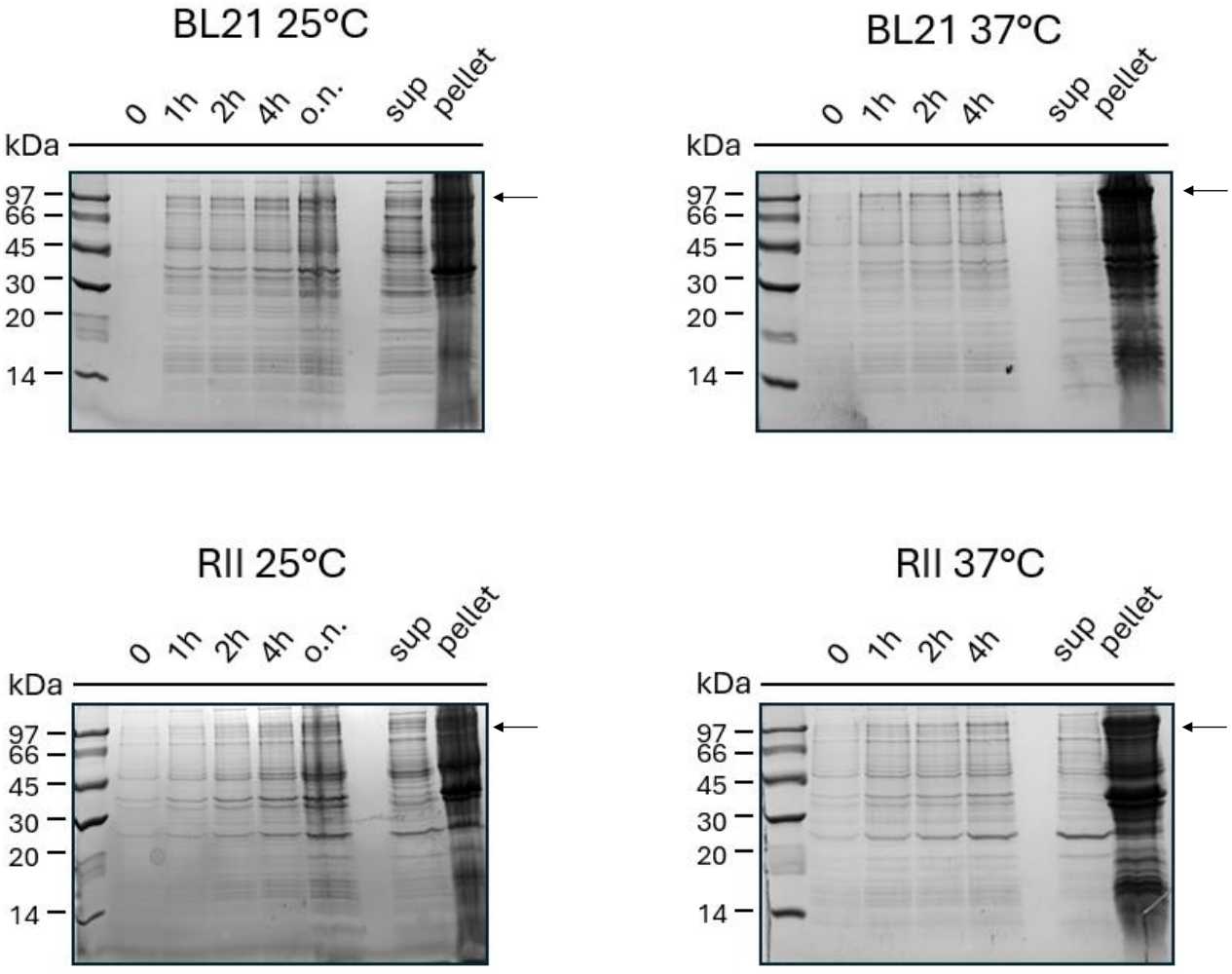


Figure 13. *Expression screening of Hsp90C^{12} in E. coli BL21 (DE3) gold and E. coli Rosetta 2 (DE3) at 25 and 37°C. The gel shows the expression of the construct as an intense band at $MW_{app.}$ of ~97 kDa. SDS-PAGE analysis of BL21 and Rosetta 2 cell lysate (supernatant and pellet) shows that the construct Hsp90C^{12} is almost exclusively expressed in inclusion bodies. The black arrows indicate the expected position of the overexpressed protein.*

4.1.3 MesNa cleavage and protein purification of Hsp90 construct C^{12} (1-528-SR)

MesNa induced cleavage was optimized by screening different cleavage times, temperatures, chaotropic salt conditions. First, the resulting fractions of interest of IB-pellets and supernatant were united and dialyzed against different urea concentrations at 4°C overnight after an initial purification step via ÄKTA-purifier. The cleavage was performed by 1:1 dilution with 1M MesNa, TBS pH 7.5 at room temperature. The process was monitored by SDS-PAGE for over 20 hours. The results of the cleavage condition screen, depicted in figure 14, show that cleavage was only performed around 50% regardless of the applied conditions.

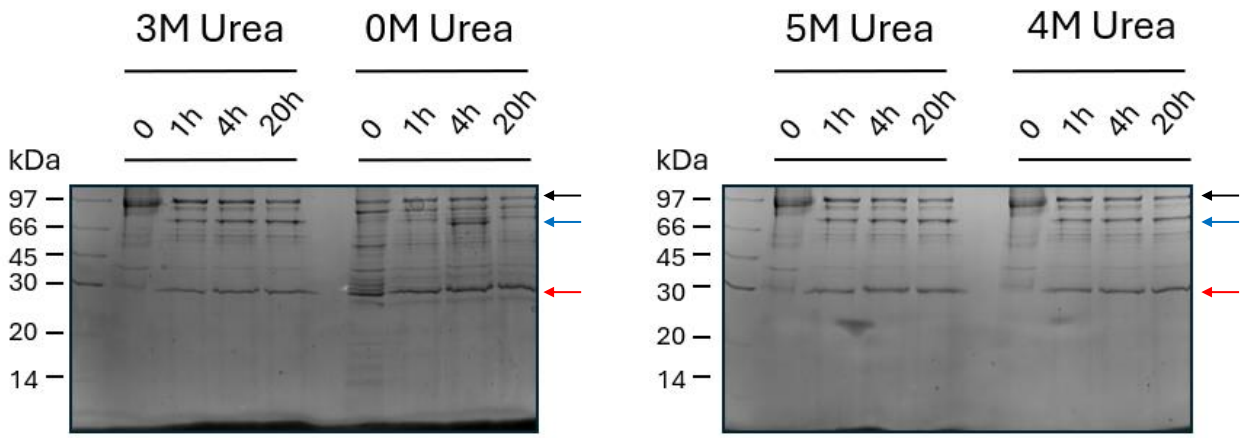


Figure 14. *MesNa cleavage condition screen, monitored by SDS-PAGE directly before 1:1 dilution with 1M MesNa (t=0) and after 1, 4 and 20 hours. The band appearing at MW_{app} 97kDa resembles the expressed construct with intein bound. The two bands at MW_{app} ~ 70kDa and 30kDa indicate that cleaved construct and intein are produced over time. The black arrows show uncleaved starting material, the blue ones cleaved product and the red ones intein.*

In another approach, pre cleavage purification was done via a NiNTA-batch system in syringes according to the section "*Intein cleavage and purification of Hsp90 1-528-MxeGyrA-His7-CBD (Hsp90 C12) construct*". Therefore, the inclusion body pellets of expressed construct C^{12} were dissolved in approximately 20ml of 6M Gdn·HCl, TBS pH 7.5 overnight. On the following day, the solution was centrifuged, the supernatant split into equal parts and loaded onto the syringes filled with Ni-Sepharose resin. After incubation for 2 hours, followed by three washes with 40mM imidazole total, the construct was eluted with 500mM imidazole in 30 to 45 minutes. The process was monitored by SDS-PAGE and samples were taken after each step (load, flow through, wash(es), elution). The pre-purified eluate was then further submitted to dialysis. Optimal dialysis conditions were found to be 2M urea TBS pH 7.5 at 4°C overnight. Optimal conditions for the subsequent MesNa cleavage are slow addition of 500mM MesNa as

a solid over 10 to 15 minutes. Afterwards pH was adjusted to 7.5. The cleavage was performed overnight at room temperature. The process was monitored by SDS-PAGE once more, taking samples right before MesNa was added as a solid and throughout the cleavage up to 24 hours. When precipitation during cleavage was observed, the solution was cleared again by adding 5 to 10 ml of 6 Gdn·HCl, TBS pH 7.5. To the clear solution, 40mM imidazole were added as a solid before incubation of the solution with two times 5 mL pre-equilibrated Ni-Sepharose resin over 60 to 90 minutes for an affinity separation of uncleaved His-tagged containing starting material and His-tag containing intein from the desired product ("reverse" Ni-NTA). After one wash with 40mM imidazole, elution was achieved with 500mM imidazole. SDS-PAGE samples were taken after each step. Figure 15 shows the process of purification from the first "batch" Ni-NTA to the "reverse" Ni-NTA as a final pre purification step. The cleaved construct C^{12} was obtained in satisfactory purity in the flow through and submitted for a final purification step via RP-HPLC.

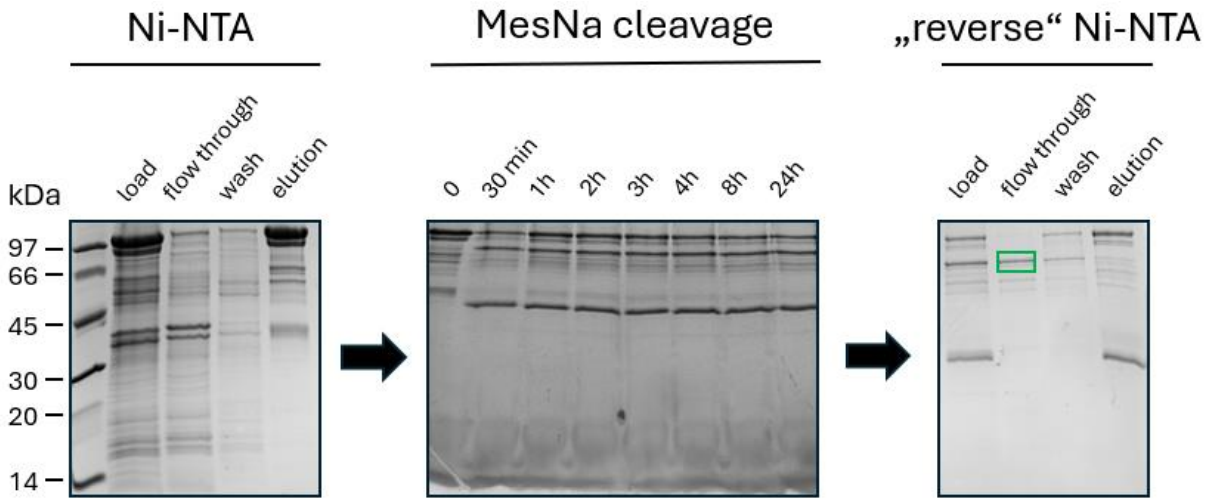


Figure 15. *SDS-PAGE monitoring of Hsp90α C^{12} purification process. Ni-NTA in a syringe batch system leads to pre purification of the construct. Dialysis of the elution and following MesNa cleavage afford the product as cleaved protein. In a third step the cleaved protein is purified via "reverse" Ni-NTA in the same syringe batch system, which leads to nearly pure product of cleaved protein as a MesNa ester. The resulting flow through, boxed in green, is submitted to a final purification process via RP-HPLC.*

Purification of the flow through via RP-HPLC at 60°C did not lead to any product (see *supplementary data* **S3**). This further confirms the assumption that the expressed construct might be too hydrophobic and therefore shows too strong interactions with the C4 stationary phase of the RP-HPLC which leads to the observed detrimental recovery rate.

To combat these poor recovery rates and since the results of the very first purification screen indicate that it is challenging to keep the protein in solution during MesNa cleavage, cleavage

was performed on a larger scale under optimized conditions (3M urea, TBS, pH 7.5, 500 mM MesNa) and the precipitating protein was simply dissolved in 6M Gdn before final purification on RP-HPLC. Purification of the dissolved protein pellet at 60°C did not lead to any product (see *supplementary data* **S4**), however, RP-HPLC purification of the supernatant of the MesNa cleavage at 60°C led to 2mg of relatively clean final product, judged on a rather challenging purification, as depicted in figure 16.

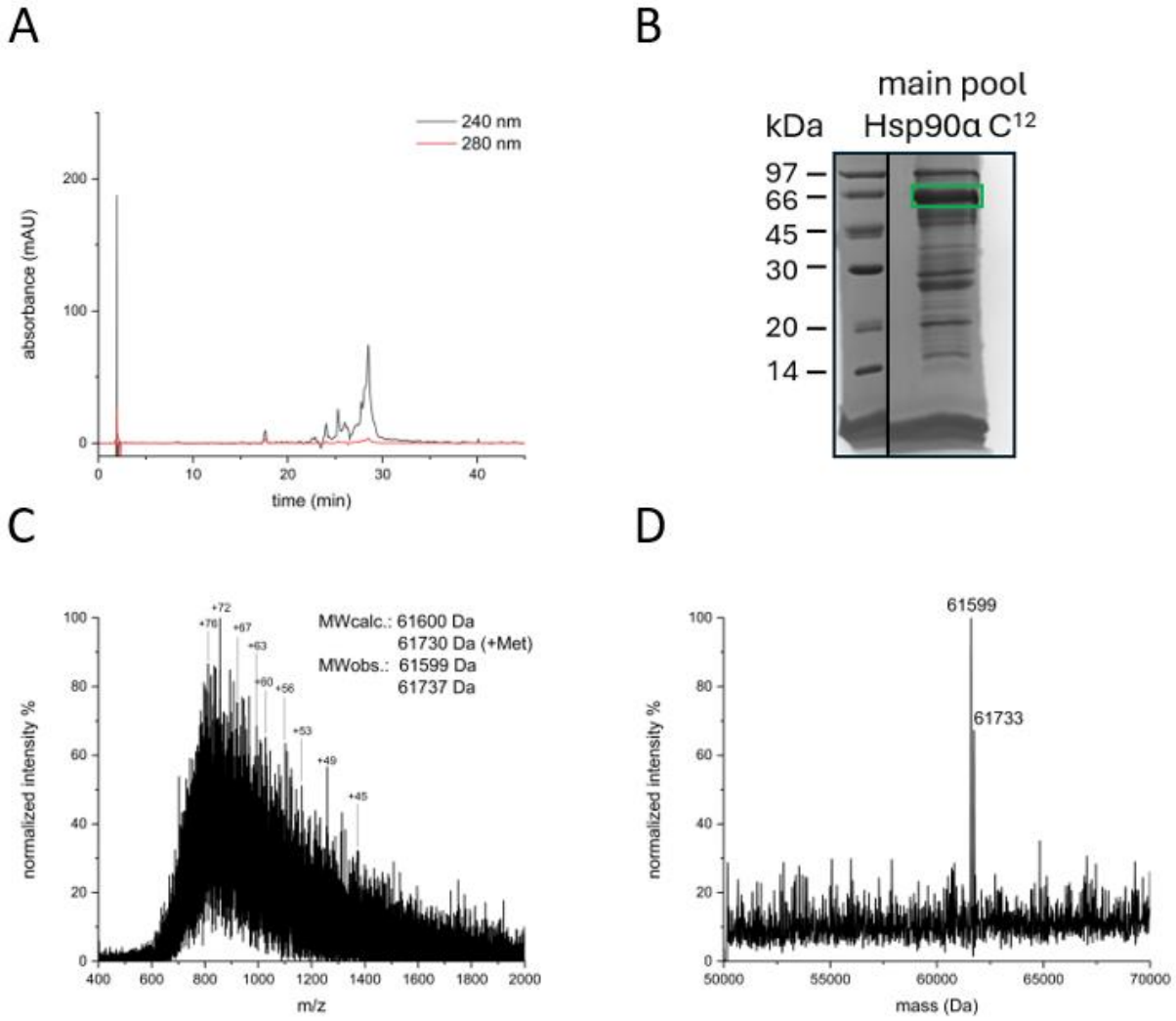


Figure 16. A) *Analysis of Hsp90α C^{12} via RP-HPLC* ***B)*** *SDS-PAGE characterization of the main pool after purification. Cleaved product is shown on the right side with a strong band at MW_{app} 66kDa boxed in green* ***C)*** *Characterization of Hsp90α C^{12} via LC-MS as a raw spectrum. $MW_{calc.}$ and $MW_{2calc.}$ are the calculated masses of C^{12} as a MesNa ester with expression initiating methionine and without methionine. The observed masses are in good accordance with the calculations* ***D)*** *deconvoluted LC-MS spectrum of C^{12}.*

Overall, transformation and expression of full-length construct C^{12} was successful. Pre-purification was optimized by experimenting with different set ups ("batch" Ni-NTA, MesNa

cleavage and subsequent “reverse” batch Ni-NTA). Finally, 1-528-SR product was obtained by purification of the supernatant after MesNa cleavage and led to a yield of 2 mg/L LB media. The obtained product was then further used for the first semi-synthesis of full-length Hsp90α.

Discussion of expression and purification of Hsp90 construct C^{12} (1-528-SR)

The main objective of the thesis is to establish a semi-synthesis protocol to generate the first semi-synthetic full-length Hsp90α protein. If successful, this strategy can be used to synthesize modified full-length constructs by introducing different post translational modifications in the synthetic parts of the final protein. Subsequently, the researcher could perform biochemical assays on these modified Hsp90α variants, to deduce structure-activity relationships, helping to encrypt the chaperone code further.

There are a few major problems for the purification of construct C^{12} that need to be discussed. The first one being that the cleavage does not work exceptionally well. This inefficiency of cleavage in intein fusion constructs often arises due to improper folding of the intein. Secondly, Hsp90 can interact with other proteins or reactants present in the mixture due to its natural function as a chaperone. Lastly, the final purification step could not be optimized fully. Purification of the construct via RP-HPLC on a C4 column proofed to be inefficient, even at 60°C. Therefore, one has to consider the use of "special" RP-HPLC stationary material like C3 columns or think about other ways of final purification and desalting such as ion exchange chromatography or hydrophobic interaction chromatography (HIC), that have not been pursued further in this work.

4.2 Solid phase peptide synthesis (SPPS)

4.2.1 Solid phase peptide synthesis of MD peptides Acm(Cys)-529-571-SR

Fmoc-based-solid phase peptide synthesis of middle-domain peptides 529-571-SR with modifications of a lysine residue at position 546 with an acetyl- or alloc group was initiated automatically on a Liberty Prime until residue 548 as depicted in figure 17. The alloc-protected lysine can be used strategically to introduce succinylation in a subsequent process. After the last automatic coupling of glycine residue 548, the resin was split into equal parts, to prepare the acetyl-lysine peptide for one and the alloc-lysine for another. Both peptides were synthesized on a 0.05 mM scale. Manual Fmoc SPPS was achieved according to *"3.2 Peptide Synthesis"*.

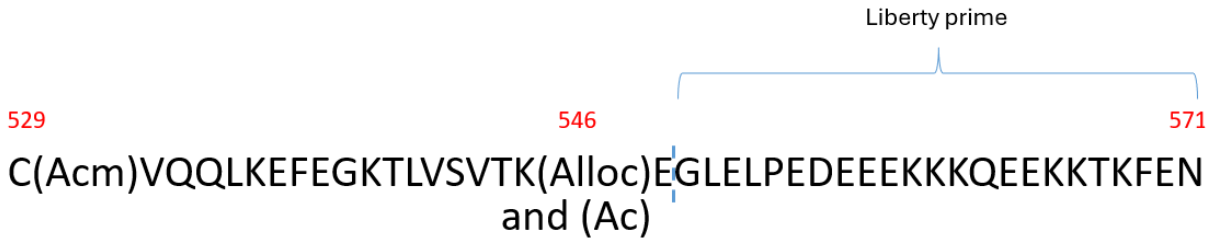


Figure 17. *Schematic overview of MD peptides Acm(Cys)-529-571-SR sequence and semi-automated Fmoc SPPS approach. Residues 548-571 were synthesized automatically with a Liberty Prime peptide synthesizer. Residues 529-547 were synthesized manually.*

4.2.1.1 Synthesis of MD peptide Acm(Cys)-529-571-SR (AcLys) (= C1)

The synthesis of MD peptide **C1** followed a standard Fmoc SPPS protocol. Double couplings were performed for each amino acid. Test cleavages were performed once right after the automated synthesis on the Liberty Prime synthesizer, twice during the manual synthesis and once more on the finished peptide **C1**. The peptide was synthesized on a 0.05 mM scale and product was obtained in a yield of 5.5 % (14.5 mg). Analytical HPLC chromatogram and ESI-MS spectrum of purified peptide **C1** is shown in figure 18.

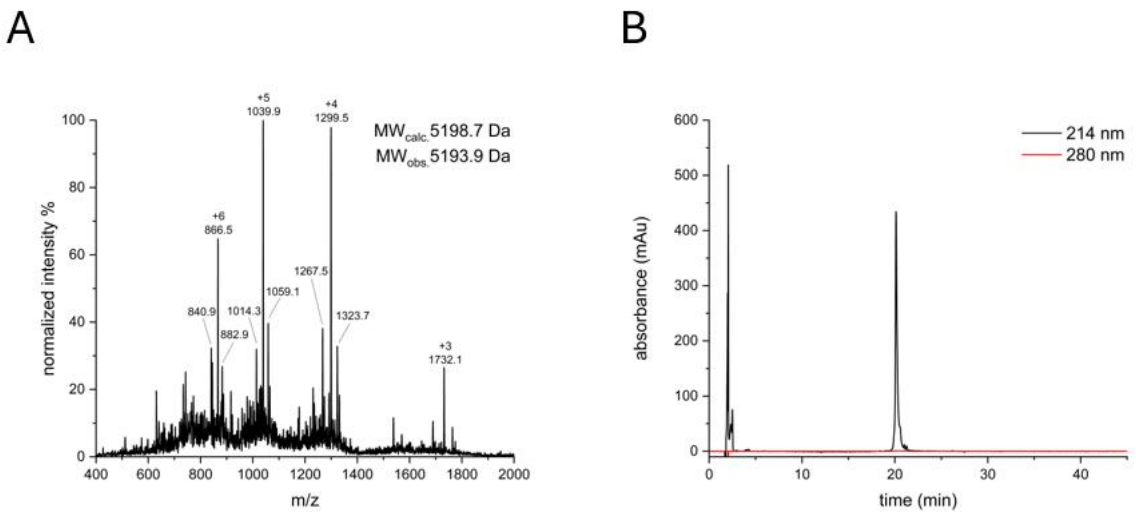


Figure 18. *Characterization of purified Hsp90 528-571-SR peptide* ***C1****.* ***A)*** *ESI-MS of Acm(Cys)-529-571-SR (AcLys). The observed mass* $MW_{obs.}$ *5193.9 Da is in good agreement with the calculated mass of* $MW_{calc.}$ *5198.7 Da. The observed masses of 840.9, 1014.3 and 1267.5 Da can be assigned to* $[M+6H]^{6+}$*,* $[M+5H]^{5+}$ *and* $[M+4H]^{4+}$ *leucine (L) deletion products respectively. The observed masses of 882.9, 1059,1 and 1323.7 Da can be assigned to* $[M+6H]^{6+}$*,* $[M+5H]^{5+}$ *and* $[M+4H]^{4+}$ *valine (V) addition products respectively* ***B)*** *HPLC analysis of Acm(Cys)-529-571-SR (AcLys) shows a single peak around 20 minutes.*

Overall, manual Fmoc SPPS for peptide **C1** was considered successful and the final mass of the obtained product is in good agreement with the calculated mass for this peptide. However, the final analysis of peptide **C1** shows slight impurities of +97 Da, which could correspond to an additional Valine (V), of which three are present in the peptide or a TFA ester. An addition product with proline (P) is highly unlikely since proline was only coupled during automated peptide synthesis and test cleavage after automated peptide synthesis indicated no impurities or adducts. The second impurity found in the final analysis of **C1** corresponds to a mass of -132 Da. This indicates leucine (L) deletion, which appears in the sequence. One of the leucine residues sits right next to the beta-branched amino acid valine, which could be the explanation for an unsuccessful coupling. This could be improved by coupling leucine a third time. These impurities could not be separated after final cleavage, leaving room for optimization of synthesis and purification of the MD peptides. The relatively low yield of the final peptide can be explained by accidental loss of dried resin before cleavage. In summary, the 41 amino acids long peptide **C1** could be obtained in satisfactory purity and yield for following ligation reactions to the semi-synthetic C-terminal construct 572-732.

4.2.1.2 Synthesis of MD peptides Acm(Cys)-529-571-SR (AllocLys) (= C2)

The synthesis of MD peptide **C2** followed the exact same procedure as for the synthesis of peptide **C1**. LC-MS analysis of the test cleavages is depicted in figure 19. Although the two test cleavages during manual peptide synthesis display almost exact accordance of observed to calculated mass, the analysis of the peptide shows that the final observed mass if off by a big margin when compared to the final calculated mass. Notably, the final mass of the analyzed peptide is only 132 Da heavier than the mass of the peptide after the second test cleavage over 7 residues before. These results cannot be explained logically and will therefore not be discussed any further. Additionally, the solid phase peptide synthesis was performed towards the end of the practical work for this thesis and therefore was not repeated. However, there is no indication that a middle domain peptide with an alloc-modified lysine (K) is not feasible, especially because the synthesis of **C1** was so straightforward and uncomplicated.

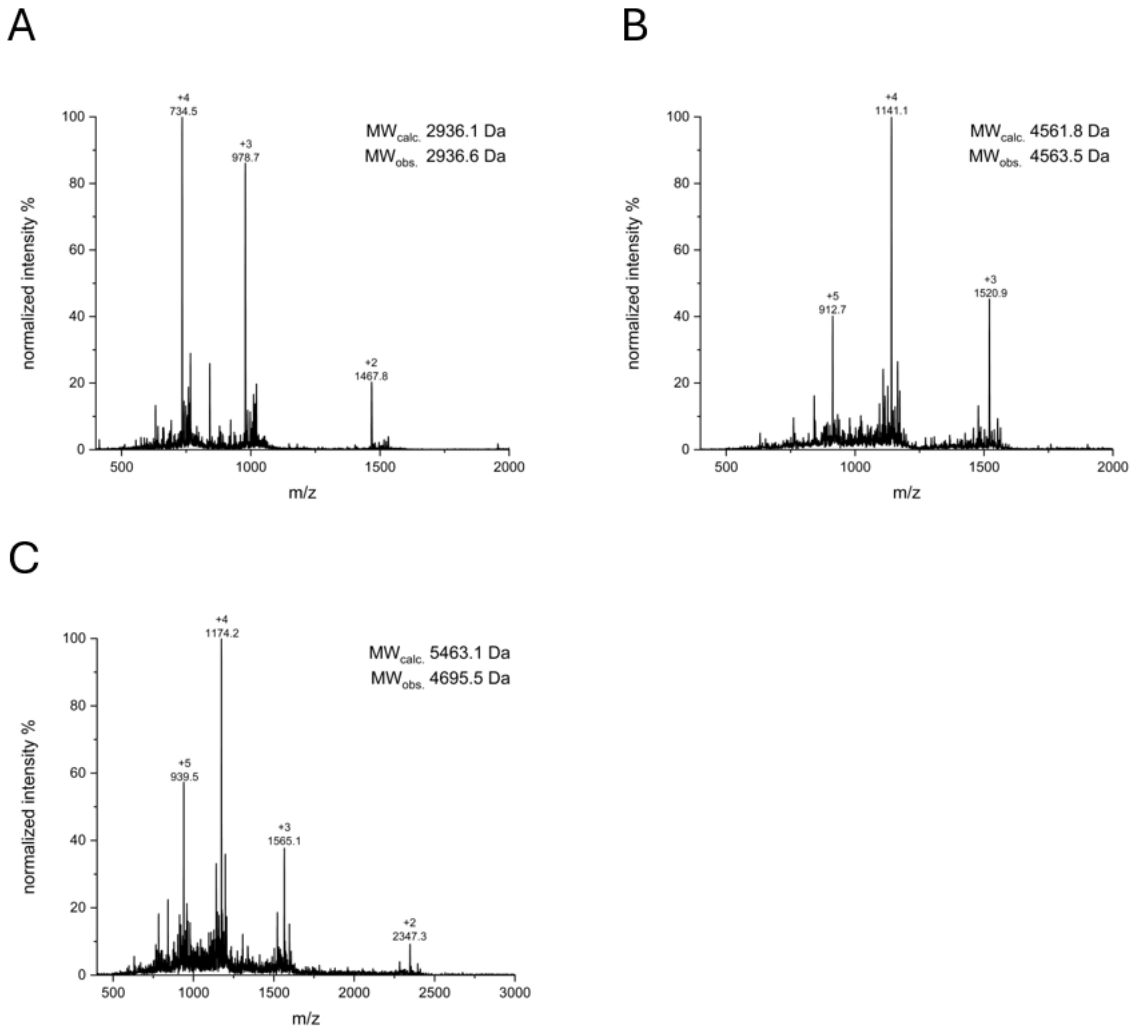


Figure 19. *Characterization of Hsp90 528-571-SR peptide* ***C2****.* ***A)*** *LC-MS of first test cleavage for residues 544-571-SR (AllocLys). The observed mass MW$_{obs.}$ 2936.6 Da is in good agreement with the calculated mass of MW$_{calc.}$ 2936.1 Da.* ***B)*** *LC-MS of second test cleavage for residues 536-571-SR*

(AllocLys). The observed mass $MW_{obs.}$ 4563.5 Da is in good agreement with the calculated mass of $MW_{calc.}$ 4561.8 Da. ***C)*** *LC-MS of final test cleavage for residues 529-571-SR (AllocLys). The observed mass $MW_{obs.}$ 4695.5 Da displays a molecular weight difference of 767,6 Da to the calculated mass of $MW_{calc.}$ 4561.8 Da.*

4.3 Expressed protein ligations (EPL)

4.3.1 EPL of Hsp90α C^8 to peptide thioester Hsp90 2-20 (pThr5/pThr7) (OG^{13}) and *in situ* desulfurization

The small-scale test ligations of recombinant protein segment Hsp90α C^8 to the synthetic peptide thioester OG^{13} was monitored by LC-MS as well as SDS-PAGE, as depicted in figures 20 and 21. Both confirm emerging ligation product over time, indicating successful ligation with estimated conversion rates of ~ 80% based on SDS-PAGE band intensities depicted in figure 20.

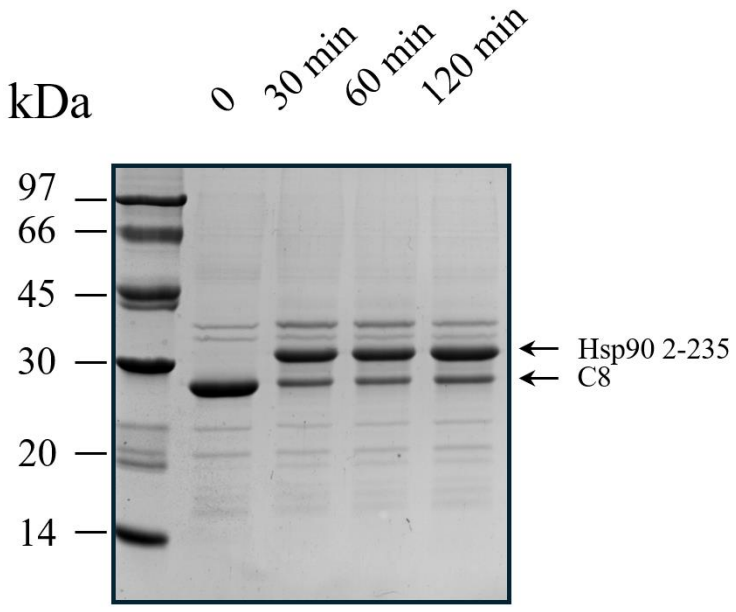


Figure 20. Monitoring of the reaction between recombinant Hsp90α C^8 and synthetic Hsp90 2-20 peptide thioester OG^{13} via SDS-PAGE. *Samples for the time point zero (t=0), were taken after reactants were dissolved and the pH of the reaction mixture adjusted.* ***A)*** *LC-MS monitoring taken at the time points t=0 (after combination of the reactants and pH adjustment), t=30 min., t=1 h. and t=2 h.*

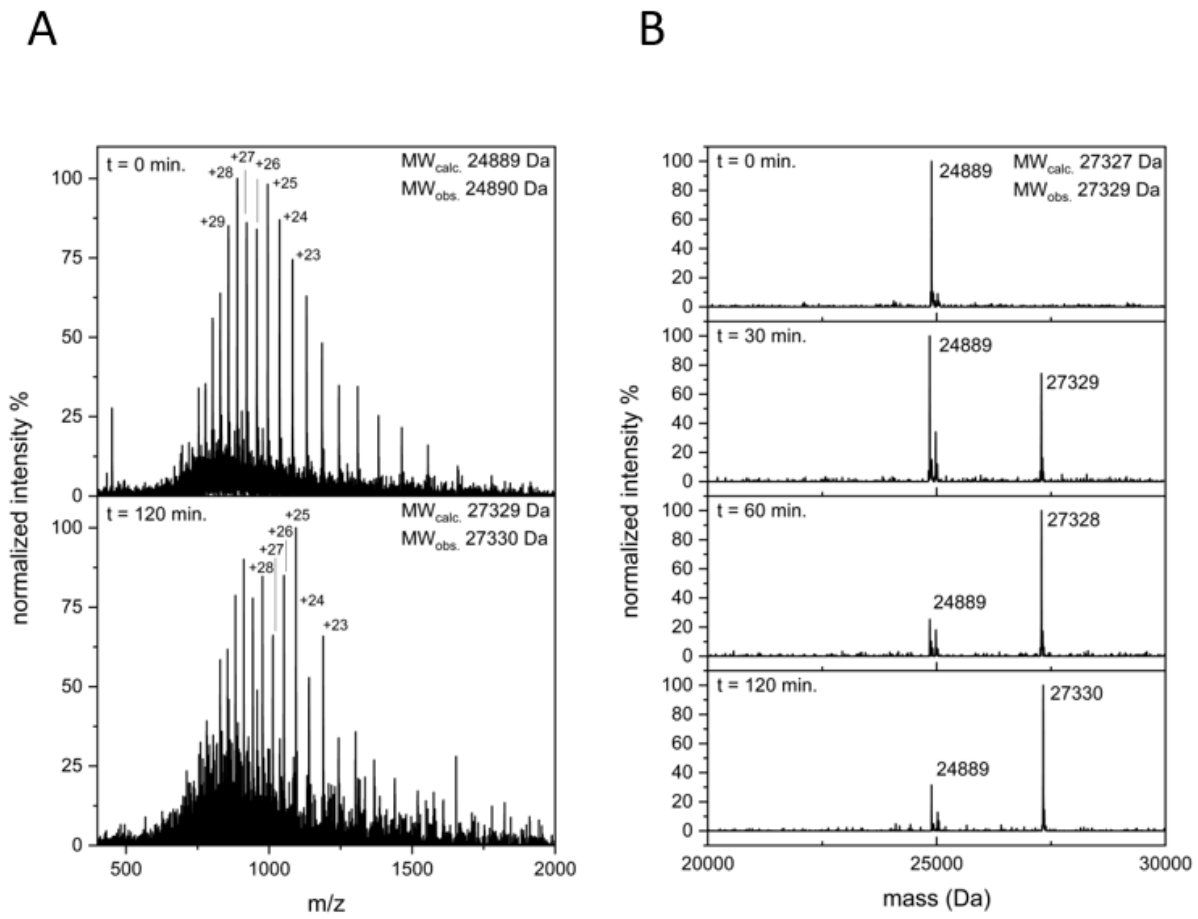


Figure 21. Monitoring of the reaction between recombinant Hsp90α C[8] and synthetic Hsp90 2-20 peptide thioester OG[13] via LC-MS. *Samples for the time point zero (t=0), were taken after reactants were dissolved and the pH of the reaction mixture adjusted.* ***A)*** *LC-MS monitoring taken at the time points t=0 (after combination of the reactants and pH adjustment), t=30 min., t=1 h. and t=2 h.* ***B)*** *Deconvoluted LC-MS spectra at the time points t=0, t=30 min., t=1 h and t=2 h. The observed mass $MW_{obs.}$ 27329 Da is in good accordance with the calculated mass $MW_{calc.}$ 27327 Da.*

The first reaction was performed at room temperature in standard conditions of 6M Gdn·HCl, 150 mM NaP_i at pH 7.2, 50 mM TCEP and 100 mM thiol additive MPAA were used for the ligation buffer. MPAA showed to be the most efficient additive for this ligation reaction, gives rise to fast reactions rates and is easy to handle. After two hours ligation time, the solution got quenched by acidification with HCl (pH lowered to ~ 3) and ethyl acetate extraction (three times with 1.5 equiv. of volume) to get rid of any excess of thiol catalyst used (since MPAA is a known radical quencher and needs to be removed), before adding the desulfurization buffer with 500 mM TCEP and 100mM of $NaBEt_4$.[69] Desulfurization buffer with a concentration of 1g ml $NaBEt_4$ per ml was prepared freshly on the same day for every ligation. The reaction was monitored by LC-MS, depicted in figure 22, and showed no conversion to the final desulfurized product after three hours reaction time. The outcome did not change when the thiol catalyst was

changed in a second ligation reaction for sodium methane thiolate (MesNa) while all the other parameters were kept equal to the first ligation. In a third attempt, the thiol catalyst was exchanged once again for methyl thioglycolate (MTG). Since MTG does not quench the radical reaction of the desulfurization, no extraction of the ligation reaction is needed before desulfurizing. However, this approach did also not lead to conversion of the successful emerging ligation product to the final desulfurized product. Prolonged reaction times of more than 3 hours showed no effect, as well as adding more desulfurization buffer or $NaBEt_4$ exclusively after 1 hour reaction time. The results of an easy-to-handle, superfast, add-and-done reaction of Xuechen Li et. al. [70] could not be reproduced, when working with the exact same protocol and parameters the researchers used.

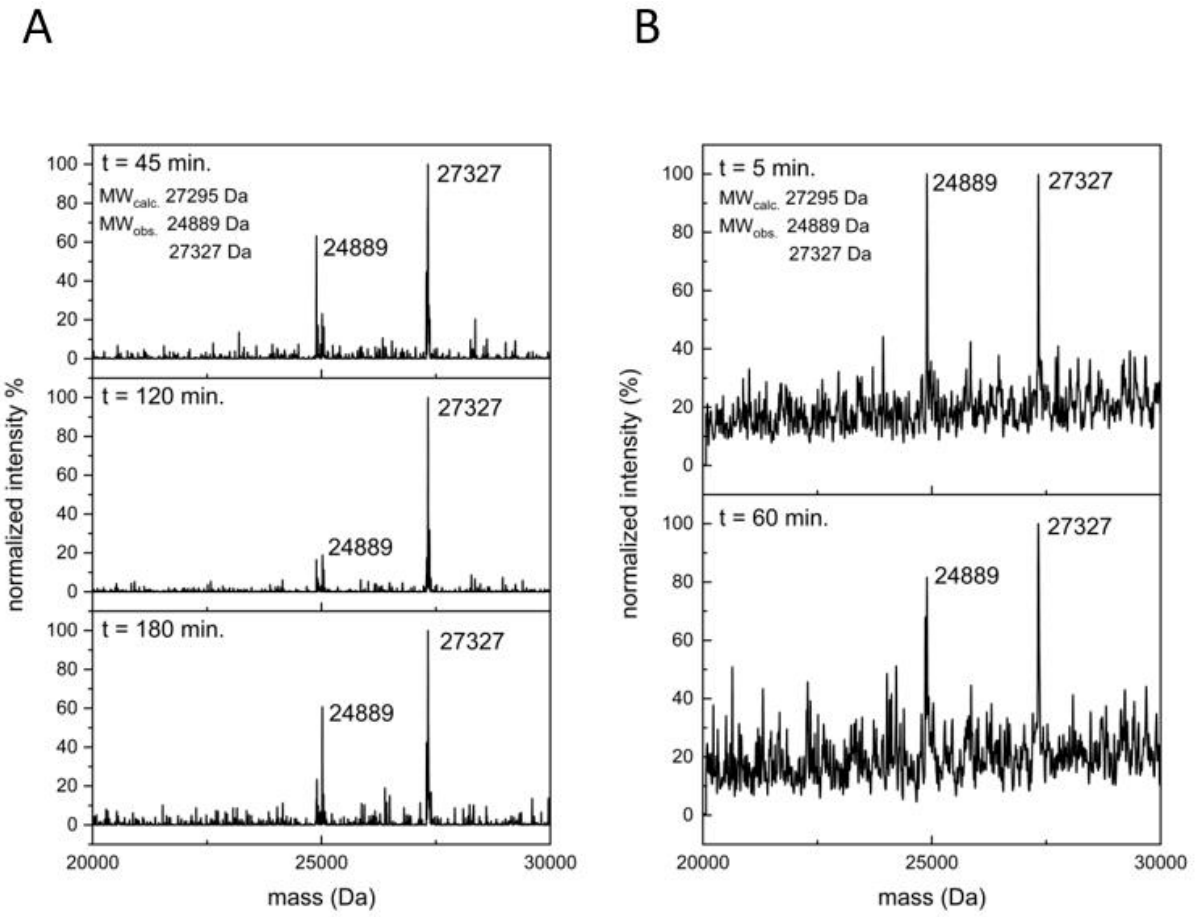


Figure 22. Monitoring of the in situ desulfurization reaction of semisynthetic Hsp90α 2-235 via LC-MS for Lig01 and Lig02. A) *Deconvoluted LC-MS spectra of monitoring Lig01 with thiol additive MPAA taken at the time points t=45 min., t=2 h. and t=3 h.* ***B)*** *Deconvoluted LC-MS spectra of monitoring Lig02 with thiol additive MesNa at the time points t=5 min., t=1 h. The observed masses of $MW_{obs.}$ 27327 Da and 24889 are in exact agreement to the semisynthetic ligation product and the recombinant educt C^{8}. The calculated mass $MW_{calc.}$ 27295 Da for the final desulfurized product could not be observed for either ligation.*

Discussion of EPL of Hsp90α C^8 to peptide thioester Hsp90 2-20 (pThr5/pThr7) (OG^{13}) and in situ desulfurization

There are a lot of factors that can be discussed for the unsuccessful results of this particular *in situ* desulfurization. First and foremost MS signal intensity decreased overall with time, indicating protein aggregation and/or precipitation. Aggregation due to hydrophobic interactions with other proteins (or protein segments) present in the reaction mixture would be plausible, since the NTD construct displayed highly hydrophobic behavior in general. Solubility can be a crucial parameter as well, especially when working with small test scales and in the presence of salts that can reduce protein solubility because of the competition for water molecules, resulting in precipitation. Interestingly, desulfurization of semisynthetic targets by Xuechen Li et. al. is only described for recombinant segments found in the soluble fraction of the cell lysate. More interestingly, when discussing the mechanism of this novel desulfurization protocol the researchers state, that $NaBEt_4$ is oxidized in solution, leading to the formation of Et · (ethyl radical) and BEt_3. Subsequently the ethyl radical initiates the desulfurization chain reaction by abstracting the hydrogen atom from the thiol of the target molecule to generate the corresponding thiyl radical and ethane gas with vigorous bubbling. This characteristic vigorous bubbling was observed for all of the three ligations, indicating deprotonation via the ethyl radical must be present. Reasonable explanations for these unsuccessful results would be non-sufficient amounts of TCEP, since the reagent is crucial to create β-scission and eliminate the sulfur atom to create the alkyl radical. However, this work used the upper threshold of TCEP. A more theoretical explanation would be the reaction of a different non-cysteine residue, triggering immediate deprotonation with the generation of bubble after mixing. Afterwards re-protonation of this non-cysteine residue back to the native corresponding side chain would be mediated by abstraction of one proton from the triethylboron-water complex (Et_3B-H_2O). But then one has to assume that the cysteine residue gets deprotonated as well, questioning the subsequent reaction with TCEP yet again. Another angle would be low accessibility of cysteine residues, which was found to be corresponding to lower overall yields and the necessity to add more equivalents of $NaBEt_4$ over time.[69]

4.3.2 EPL of C^{12} (1-528-SR) to unmodified semisynthetic construct 529-732

A central part of this thesis was the attempt at a full-length Hsp90α semi-synthesis. Therefore, the recombinant construct C^{12} (1-528-SR; $MW_{calc.}$ 61600 Da) obtained from previous expression was used in a native chemical ligation to the unmodified semisynthetic construct 529-732-SR ($MW_{calc.}$ 23070 Da), that was previously prepared by MSc. Oliver Gajsek via NCL of residues 529-571 and 572-732. The reaction was monitored over 21 hours by SDS-PAGE and LC-MS, the results are depicted in figures 23 and 24.

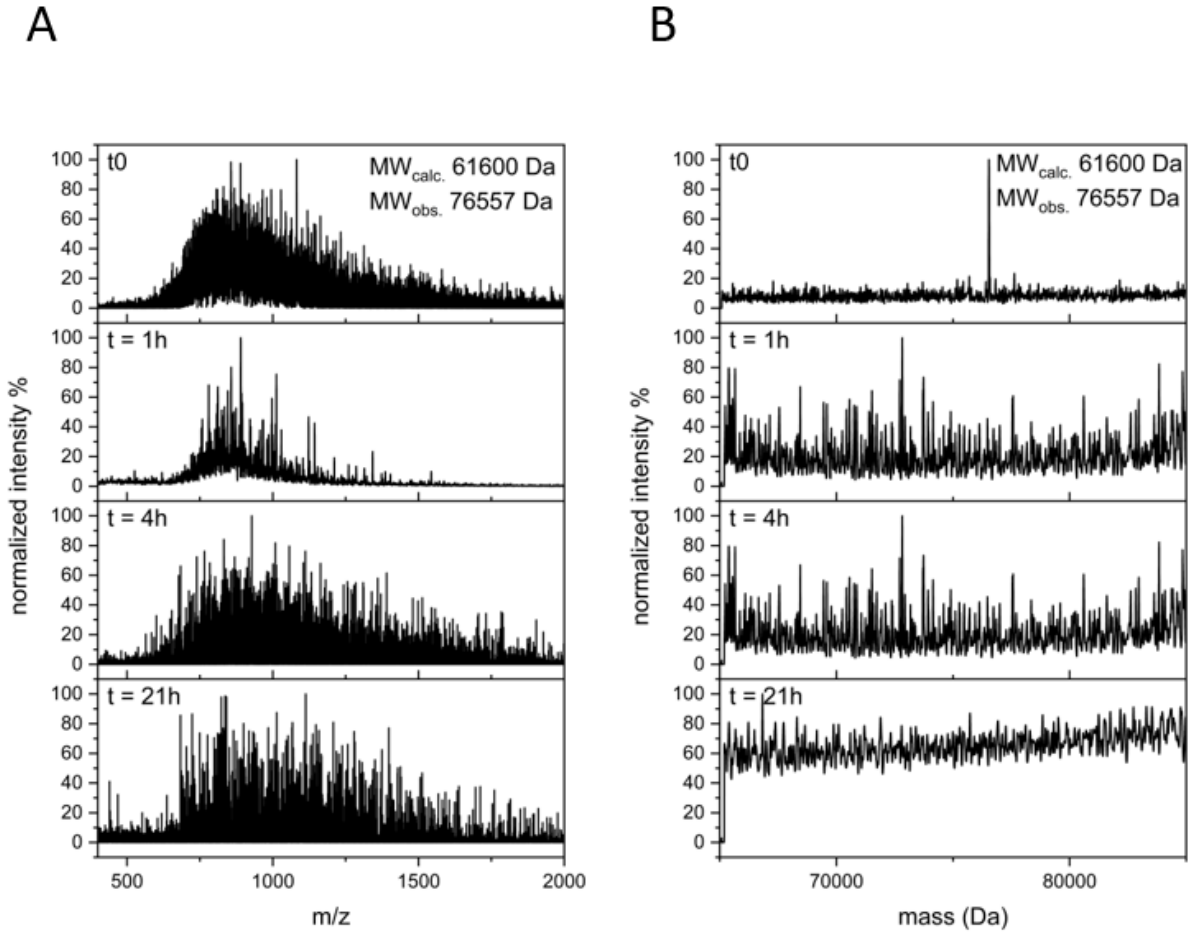


Figure 23. *Hsp90α full-length ligation of 1-528-SR to 529-732-SR monitored by LC-MS. Measurements were taken right after adding semisynthetic construct 529-732-SR and adjusting the pH at time point zero (t0), t=1h, t=4h and t=21h.* ***A)*** *raw mass spectra of Hsp90α full-length ligation* ***B)*** *Deconvoluted mass spectrum of Hsp90α full-length ligation*

Monitoring this reaction via LC-MS is challenging because of the large size of educts and product. To deconvolute such high masses out of a rather noisy background rarely leads to a hit. That is why the results of the SDS-PAGE monitoring are more meaningful for this ligation.

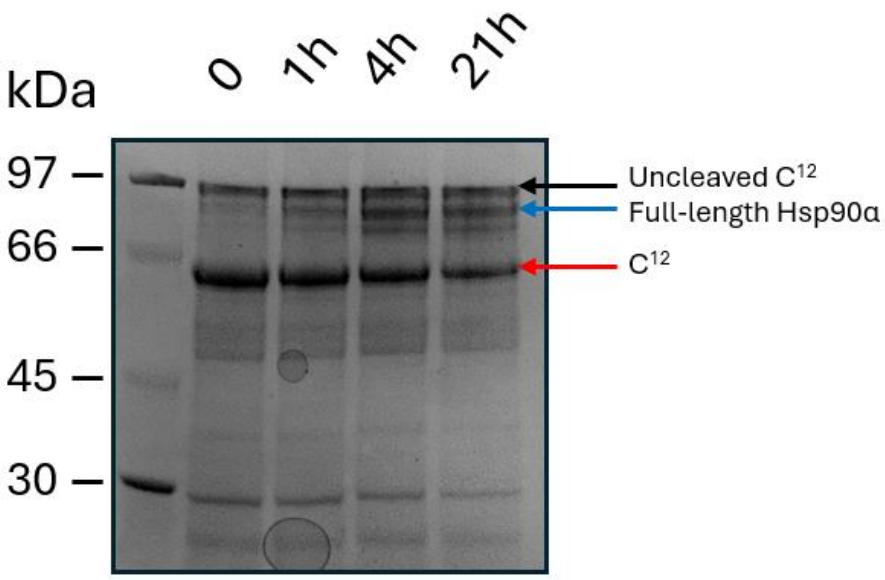


Figure 24. *Hsp90α full-length ligation monitored by SDS-PAGE at the time points t=0, t=1 h, t=4 h and t=21 h. The ligation product is expected to run at an apparent molecular weight of ~ 84 kDa while Hsp90C^{12} is running at an apparent molecular weight of ~ 60 kDa.*

For this small scale test ligation, standard ligation conditions for a native chemical ligation were used (*"3.4.2.2 EPL of Hsp90C^{12} to peptide thioester CTD$_{Lig01}$"*). The gel shows the emergence of a band at MW$_{app.}$ ~ 85 kDa, resembling the full-length Hsp90α product. This indicates successful full-length ligation between expressed construct C^{12}, that apparently gets consumed over time with a decreasing intensity of the band at MW$_{app.}$ ~ 66 kDa and the semisynthetic construct 529-732, also decreasing in intensity over time at an MW$_{app.}$ of ~ 30 kDa. An exact mass of the product could not be confirmed by LC-MS analysis which was expected due to the already difficult ionization and signal generation of the smaller C^{12} construct which displays the limitation of the used single single-quadrupole system. SDS monitoring showed only minor product formation between 4h and 21h and therefore the reaction was considered complete after 21 hours, although there was still ~ 30% educt present.

Discussion of EPL of C^{12} (1-528-SR) to unmodified semisynthetic construct 529-732

The final construct is assembled from 4 peptide segments (2 recombinantly produced and 2 made by SPPS). The first is a rather large recombinant construct C^{12} that covers residues 1-528(-SR), the second is a recombinant C-terminal domain (CTD) construct of moderate size, including residues 572-710(-SeR). In a first chemoselective ligation, the C-terminus of the CTD construct is completed with a synthetic diselenide peptide (residues 711-732) in a diselenide-selenoester ligation (DSL), with the possibility to introduce modifications of choice via the synthetic fragment. This semisynthetic CTD construct is then ligated to another synthetic (and modifiable) thioester fragment via NCL (native chemical ligation), resembling the middle domain of the protein with residues 528-571-SR. In the last step of the semi-synthetic route, the resulting semi-synthetic MD/CTD construct, consisting of a synthetic, a recombinant and a synthetic fragment, is ligated to the large C^{12} construct in a NCL to yield the final full-length Hsp90α protein. The choice of strategy is based on requirements for semi-synthesis in general, by keeping synthetic fragments within reach of the standard Fmoc-SPPS procedure lengthwise. But it is also based on prior experience of working with Hsp90 constructs. Since truncated NTD constructs have shown to be difficult to handle because of poor solubility, the first recombinant segment was chosen to be larger with residues 1-527, making the expression and purification easier to handle. However, this leaves the researcher no room for modifications in this part of the protein other than mutation. Furthermore, earlier works on the semi-synthesis of the C-terminal domain have shown that diselenide-selenoester ligation between 572-710 and 711-732 is successful, which is why these two sequences have been selected. The remaining residues 528-571 of Hsp90α´s middle domain are manageable in size for standard Fmoc-SPPS, and the required thioester moiety can be produced rather easy via *in situ* hydrazide to thioester conversion, when synthesizing the peptides as hydrazide precursors.

The incomplete ligation of C^{12} (1-528-SR) to unmodified semisynthetic construct 529-732 can be a consequence of the thioester stability of the recombinant C^{12} construct. When a thioester is too stable, the initial nucleophilic attack by the cysteine thiol becomes less favorable. This in turn would reduce the formation of thioester intermediate and thus slow down the overall reaction rate, leading to incomplete ligation. Aryl thioesters that are generated when using MPAA are usually rather stable because the aromatic ring can delocalize electron density and provide resonance stabilization. The reaction mixture was purified right after completion of the reaction by RP-HPLC at 60°C. Fractions of interest were pooled, lyophilized, and analyzed via SDS-PAGE consequently, because detection via LC-MS turned out to be problematic due to high molecular weights of the reactants and product. Analysis of the fractions revealed product

in multiple fractions but separation via HPLC was not satisfactory, leading to rather impure fractions that contain unconverted C^{12} although most of the semisynthetic CTD/MD construct could be separated. In summary, ligation conditions and the subsequent purification can be optimized. But the main goal of this first test ligation was to validate the proof of concept, which was successful. The SDS-PAGE analysis reveals that a full-length semi-synthesis of Hsp90α with the herein described strategy works. This can be interpreted as a solid foundation to create numerous semi-synthetic full-length variants of the targeted protein. Ultimately, this gives researchers the opportunity to introduce modifications that have not been studied yet or do not occur naturally which can help to decipher the chaperone code even more. Further optimization of ligation and purification would enable larger scale production and future structure-function relationship studies of post translationally modified full-length Hsp90α. The remaining part of this work focuses on the synthesis of modified middle-domain peptides (529-571) at a lysin residue (position 545) through the introduction of acetyl-lysine for one and alloc-lysine as a succinylation mimic for another variant. Currently, there is no research data for these modifications at this exact residue. The peptides can then be ligated to the semisynthetic CTD (572-732) and finally to construct C^{12} (1-528) in pursuit of the first modified semi-synthetic full-length Hsp90α variant.

4.3.3 Ligation of 572-710-SeR to unmodified peptide 711-732 (CTD)

In order to build up a new modified semi-synthetic middle domain/C-terminal domain construct, the unmodified peptide Sec001 was ligated to recombinant segment Hsp90 C^9 (572-710-SeR) in a diselenide-selenoester ligation. The Ligation was performed on a 1.2 µM scale and 1.5 equiv. of unmodified peptide were used for the reaction. Reaction conditions followed the standard protocol for a DSL, details are mentioned in *“3.4 Ligation reactions””3.4.2 Full-length semi-synthesis””3.4.2.2 DSL of Hsp90^{C9} to selenocysteine peptide Sec$_{001}$ (Hsp90^{C9} = Hsp90 572-711) ; (Sec$_{001}$ = 712 -732)”*. A DSL can be split into three reactions total, with the first one being the selenoester conversion of the recombinant construct bearing a hydrazide moiety at its C-terminus. The reaction was monitored by LC-MS, since SDS-PAGE would not reveal such minor differences in mass (15589 to 15714 Da). The results are shown in figure 25. Selenoester conversion was performed over one hour.

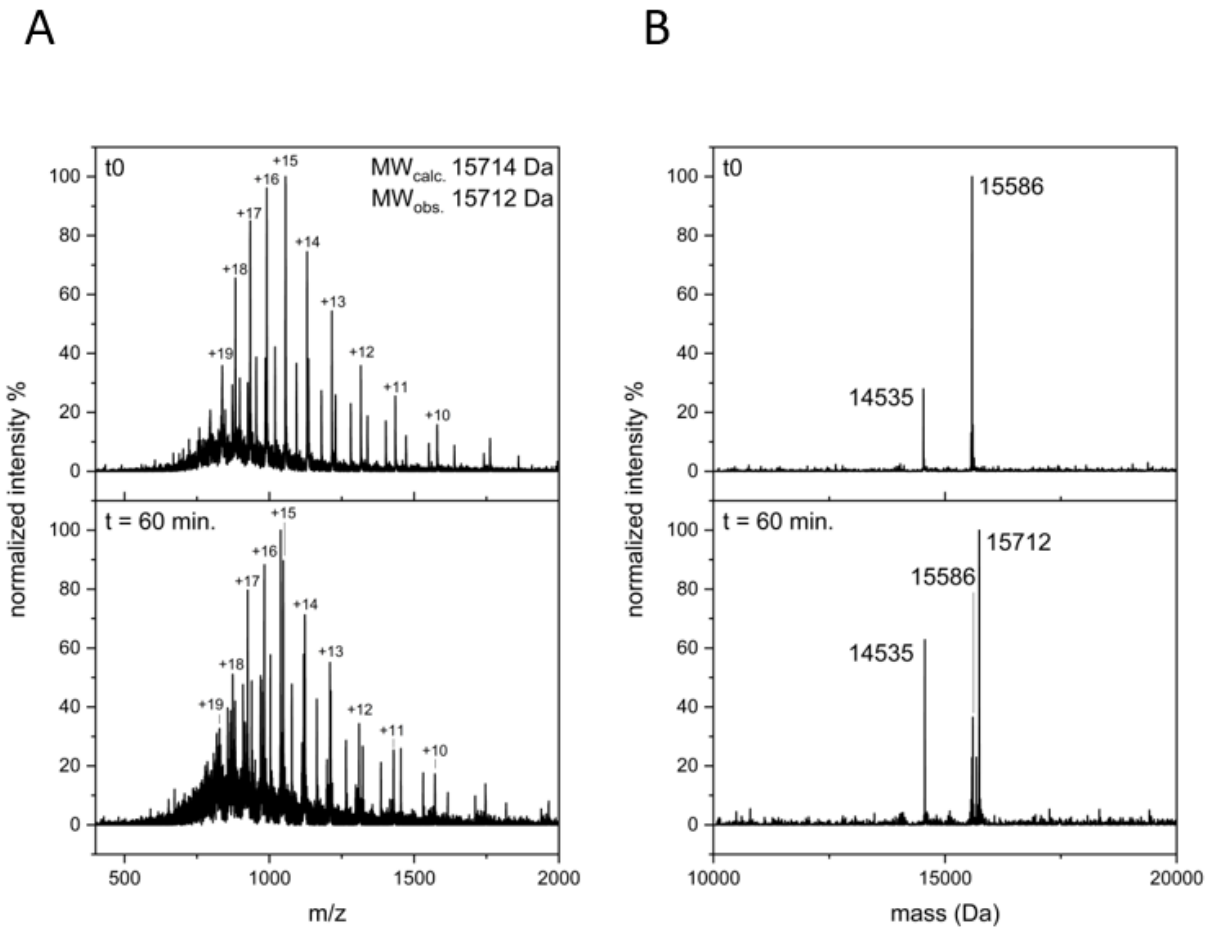


Figure 25. DSL – Selenoester Conversion: *Characterization of selenoester conversion of Hsp90 572-710 before diselenide selenoester ligation (DSL).* ***A)*** *ESI-MS spectrum shows that product is built after 60 minutes. The observed mass (MW$_{obs.}$ 15712 Da) is in good agreement with the calculated mass (MW$_{calc.}$ 15714 Da).* ***B)*** *Deconvoluted mass spectrum of Hsp90 572-710-SeR. The additional observed mass of 14535 Da cannot be identified further and is likely an impurity carried over from the recombinant construct. The other observed mass of 15586 Da can be assigned to the educt of the selenoester conversion (MW$_{calc.}$ 15589 Da).*

The selenoester conversion of recombinant hydrazide construct C^9 is successful, as indicated by the results of LC-MS analysis. The observed mass of 15712 Da is in almost exact agreement with the calculated mass of 15714 Da. LC-MS analysis also reveals, that after one hour there is still educt present, indicated by the smaller peak of mass 15589 Da. Longer selenoester conversion reaction time could help in obtaining more starting material for the subsequent ligation, but in this case, judged on the relative intensity of the peaks after 60 minutes, around 60-70% conversion were achieved which is a satisfactory result. The third observed mass of 14535 Da could not be identified further. It is most likely carried over as an impurity of the recombinant construct C^9, that could not be separated fully in the previous purification. The selenoester conversion was followed by the next step of the DSL, which is the actual ligation

with unmodified synthetic peptide Sec001. The reaction was performed for approximately 60 minutes and analyzed by LC-MS. The results are depicted in figure 26.

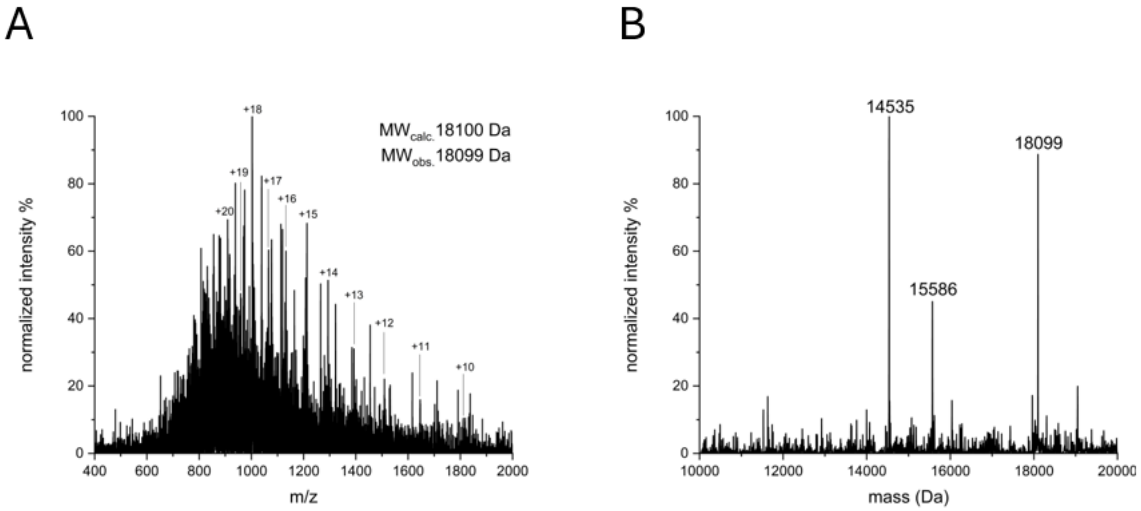


Figure 26. Diselenide selenoester ligation (DSL): *Characterization of diselenide selenoester ligation (DSL) between Hsp90 572-710-SeR and Hsp90 711-732.* ***A)*** *ESI-MS spectrum shows that the observed mass ($MW_{obs.}$ 18099 Da) is in good agreement with the calculated mass ($MW_{calc.}$ 18100 Da).* ***B)*** *Deconvoluted mass spectrum of Hsp90 572-732. The additional observed masses of 14535 Da and 15586 Da can be assigned to impurities and the educt ($MW_{calc.}$ 15589 Da).*

The diselenide selenoester ligation of recombinantly expressed Hsp90α C^9 and the synthetic unmodified peptide Sec001 is successful as well, as the results of the LC-MS analysis show. Unconverted educt of the selenoester conversion can still be seen with the mass of 155867 Da, as well as the mass of 14535 Da, characterized as an impurity. Nevertheless, after one hour a band resembling the mass of $MW_{obs.}$ 18099 Da emerges, indicating that ligation product is successfully formed. The mass is in good agreement with the calculated mass of $MW_{calc.}$ 18100 Da. After the presence of the ligation product was confirmed, the reaction solution was extracted with hexane 6 times total to quench any DPDS leftovers before going into the final step of the DSL, which is the deselenization of the ligation product to obtain the native alanine residue at the ligation junction. Deselenization was monitored via SDS-PAGE for over 4 hours, the results are shown in figure 27. The deconvoluted LC-MS spectra reveal 4 different masses found after 4 hours reaction time total. The lowest one being the impurity of 14535 Da, followed by the mass of the initial educt before the first reaction step (15586 Da, recombinant hydrazide C^9). Then the mass $MW_{obs.}$ 18021 Da of the deselenized product indicates successful deselenization over time. The last mass that is found is the ligation product with a selenocystine residue at the ligation junction (the educt before deselenization, 18099 Da). The relative

intensity of the last two masses 18021 Da and 18099 Da shows that more product is generated between one and two hours of reaction time.

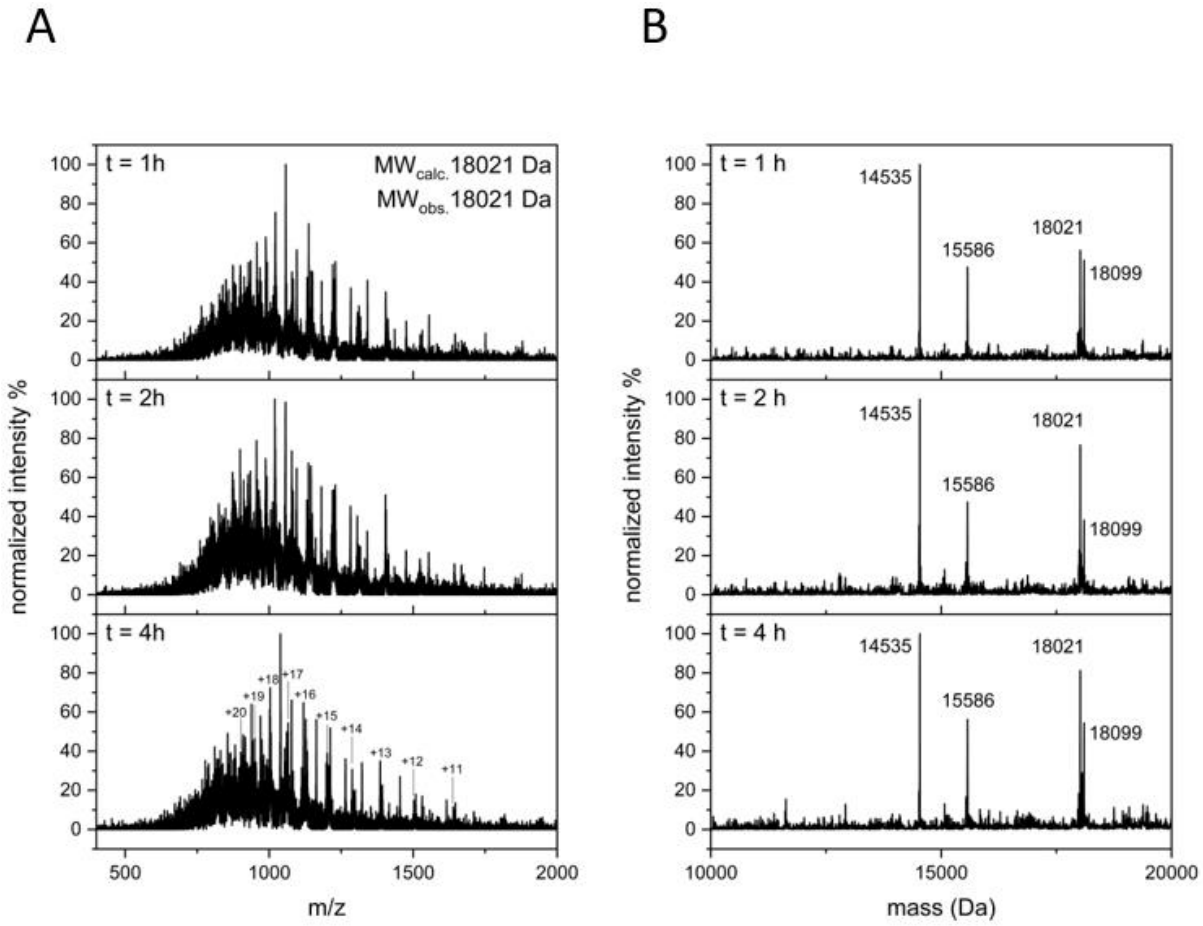


Figure 27. Deselenization DSL: *Characterization of deselenized Hsp90 572-732 after diselenide selenoester ligation (DSL).* ***A)*** *ESI-MS spectrum shows that the observed mass (MW$_{obs.}$ 18021 Da) is in exact agreement with the calculated mass (MW$_{calc.}$ 18021 Da).* ***B)*** *Deconvoluted mass spectrum of Hsp90 572-732 shows that product is built after 1 hour. The additional observed masses of 14535 Da, 15586 Da and 18099 Da can be assigned to an impurity and the corresponding educts (MW$_{calc.}$ 15586 Da Hsp90 $_{572\text{-}710\text{-}NH\text{-}NH2}$)(MWcalc. 18100 Da Hsp90 $_{572\text{-}732(A711U)}$).*

Subsequently, the reaction mixture was frozen and stored at -80°C until further purification on the next day via semipreparative HPLC. Fractions of interest were analyzed and pooled accordingly. The main pool was submitted for a final analysis by HPLC and ESI-MS, the results are depicted in figure 28.

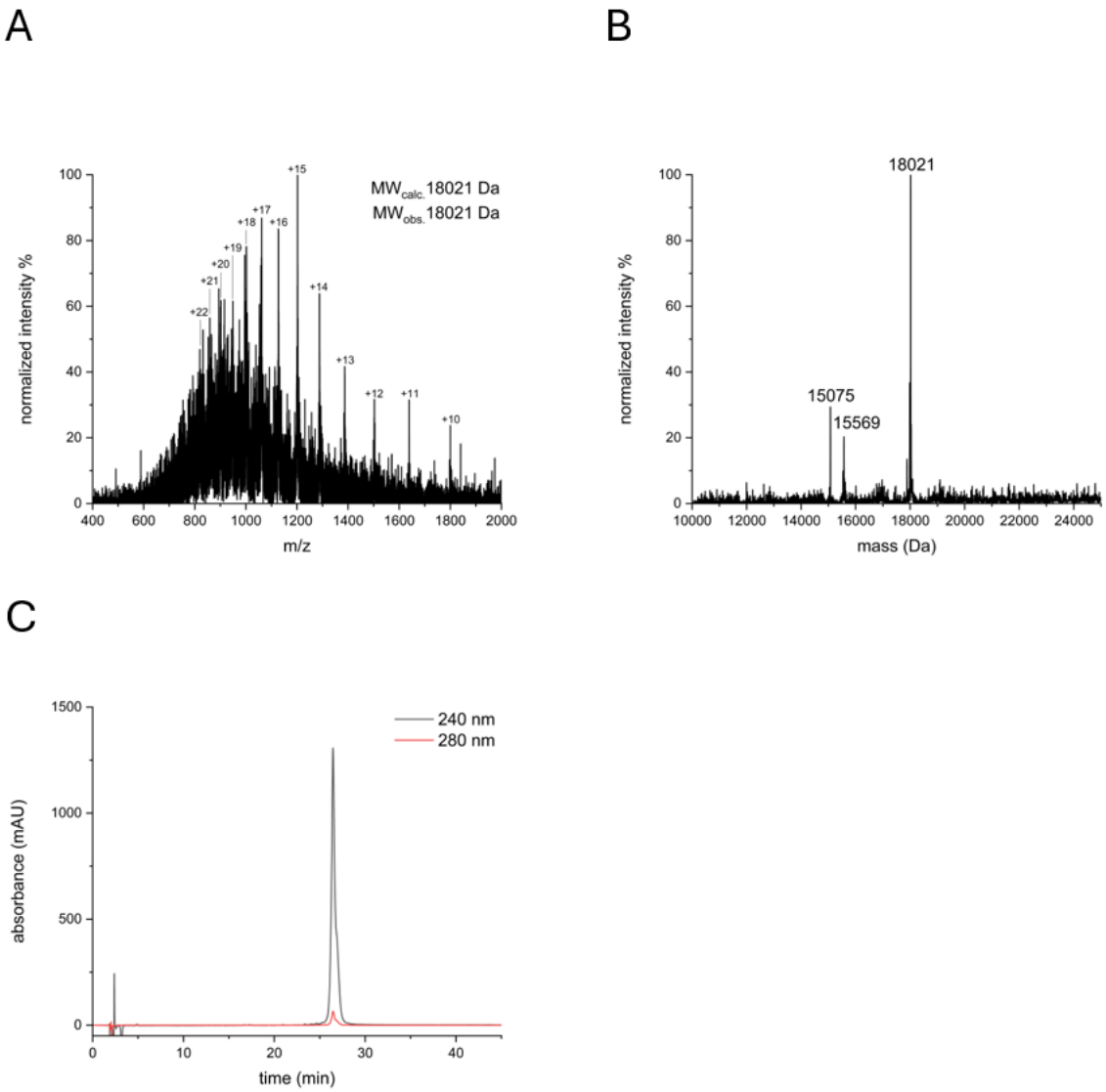


Figure 28. Final analysis DSL: *Characterization of deselenized Hsp90 572-732 after diselenide selenoester ligation (DSL).* ***A)*** *ESI-MS spectrum of the purified main pool shows that the observed mass ($MW_{obs.}$ 18021 Da) is in exact agreement with the calculated mass ($MW_{calc.}$ 18021 Da).* ***B)*** *Deconvoluted mass spectrum of purified Hsp90 572-732. The additional observed mass of 15075 Da and could not be identified further. The mass of 15569 Da corresponds to the hydrolyzed selenoester* ***C)*** *HPLC chromatogram of purified Hsp90 572-732 main pool.*

The final analysis confirms a successful diselenide-selenoester conversion between recombinant hydrazide segment 572-710 and synthetic unmodified diselenide peptide 711-732, as the calculated mass of 18021 Da is in good agreement with the observed mass of 18021 Da. Overall, the reaction yield is satisfactory with about 57% (12.4 mg; main pool). The minor impurity of 15075 Da could not be identified. The mass of 15569 Da corresponds to the hydrolyzed selenoester, which might be prevented by shorter reaction times for the selenoester conversion in the future. The obtained product was lyophilized and stored for the following ligation reaction to build up new, now modified, semi-synthetic MD/CTD construct.

4.3.4 Ligation of 529-571(AcLys) to 572-732 (CTD)

The ligation between the semisynthetic construct obtained by DSL of Hsp90 572-710-SeR to Hsp90 711-732 and the modified acetyl-lysine peptide 529-571(AcLys) previously generated by semi-automatic Fmoc SPPS was performed on a 0.05 µM scale, with 1.5 equiv. of synthetic peptide used. Reaction conditions followed the standard protocol for native chemical ligations reactions, details are described in *"3.4 Ligation reactions" "3.4.2 Full-length semi-synthesis" "3.4.2.3 Ligation of semisynthetic DSL^{008} to acetyl lysine peptide CTD$_{N03}$ (DSL^{008} = Hsp90 572-732) ; (CTD$_{N03}$ = 529-571)"*. The reaction was monitored by LC-MS and SDS-PAGE for over 90 minutes, the results are shown in figure 29 and figure 30.

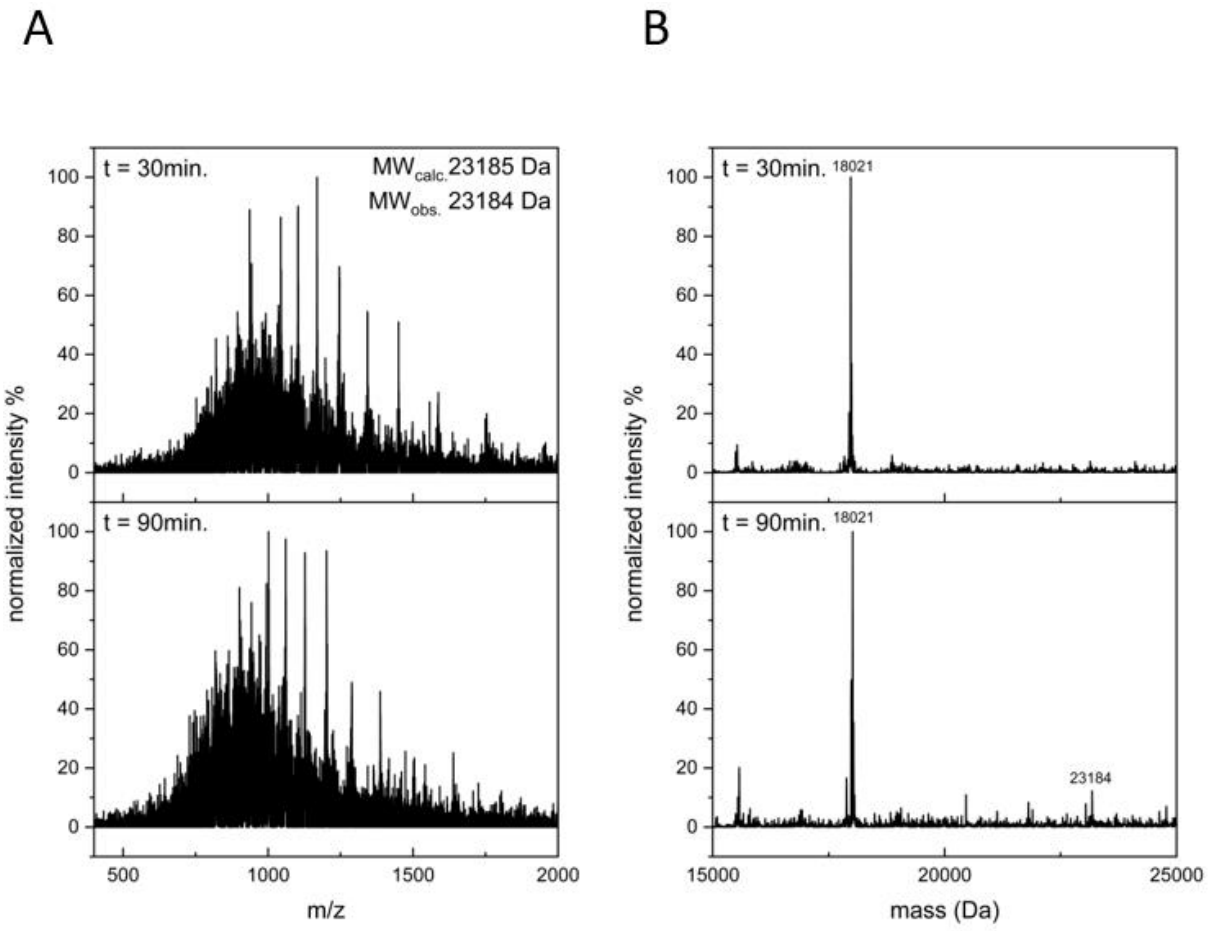


Figure 29. *Reaction monitoring of ligation between Hsp90 572-732 after diselenide selenoester ligation (DSL) and synthetic thioester peptide Hsp90 529-571.* ***A)*** *raw ESI-MS spectrum shows that the observed mass (MW$_{obs.}$ 23184 Da) is in good agreement with the calculated mass (MW$_{calc.}$ 23185 Da).* ***B)*** *Deconvoluted mass spectrum shows that product is built after 90 minutes. The additional observed mass of 18021 Da can be assigned to educt (MW$_{calc.}$ 18021 Da Hsp90 $_{572-732}$).*

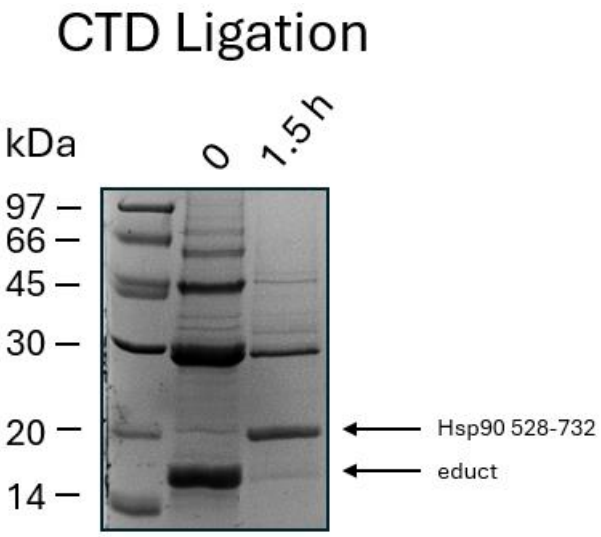


Figure 30. *Reaction monitoring of ligation between Hsp90 572-732 after diselenide selenoester ligation (DSL) and synthetic thioester peptide Hsp90 529-571 via SDS-PAGE. The observed mass at the 20kDa band is in good agreement with the calculated mass ($MW_{calc.}$ 23185 Da). Product is built after 90 minutes. The additional observed mass between the 14 and 20kDa band can be assigned to the educt ($MW_{calc.}$ 18021 Da Hsp90 $_{572-732}$).*

Product formation can be observed over time, as indicated by the results of SDS-PAGE. However, this was a small scale test ligation that indicates that either much longer reaction times or optimized reaction conditions (thiol catalyst, pH, higher temperature) are needed. When optimized, this reaction can be followed up with a full-length ligation according to the newly established full-length semi-synthesis strategy. This would, in turn, lead to the first modified semi-synthetic full-length variant of Hsp90α.

5. Conclusion

5.1 Hsp90α NTD semi-synthesis

Concluding the results of the N-terminal domain semi-synthesis of Hsp90α, the established semi-synthetic strategy turned out to be successful, however, expression and purification of truncated recombinant constructs is a challenging task that highlights the importance of sequences as small as 19 amino acids for the solubility of the protein. Subsequent native chemical ligation proofed to be successful as well. However, the essential *in situ* post ligation desulfurization could not be achieved when using $NaBEt_4$ radical chemistry despite varying various parameters and reaction conditions such as different thiol additives. Therefore, an additional purification step after the ligation and before attempting desulfurization still needs to be explored in order to generate the semi-synthetic Hsp90alpha NTD variant with the native Alanine residue at position 21. This remains a major challenge due to the significant yield loses during RP-HPLC purification due to the high hydrophobicity of the NTD and could require redesigning the ligation strategy involving cleavable solubility tags.

5.2 Hsp90α full-length semi-synthesis

For the assembly of the full-length semi-synthetic Hsp90α, an entirely new semi-synthetic route was successfully employed. The strategy focused on building the target molecule out of four smaller, more manageable segments of the protein, which enables the researcher to introduce modifications with a vast array of combinations at the very C-terminal end and in the middle domain. The transformation of construct C^{12} which comprises the first 528 residues of the protein, expression and purification was successful although further optimization of the final purification step of the recombinant construct is needed in order to upscale the production of this crucial segment. The first synthesis of full-length semi-synthetic Hsp90α by native chemical ligation of recombinant segment 1-528 to semisynthetic construct 529-732 can be considered a major achievement towards our ambitions to decipher the complex influences of PTMs on the function and structure of the targeted protein.

In total, both parts of the project successfully utilized various different semi-synthetic strategies to create unmodified and modified semi-synthetic variants of Hsp90α N-terminal domain and Hsp90α in full-length. Ligations were based on the principle of native chemical ligation as well as recent advancements of the original methodology and successful in all cases. Both parts of the project included the expression and purification of truncated recombinant segments, which proved to be an extremely complex, challenging, and time-consuming practice. The

implementation of a full-length synthetic strategy opportunities for the reliable generation of modified Hsp90α in a vast combinatorial array.

6. Outlook

The implementation of a new *in situ* desulfurization protocol for the semi-synthesis of Hsp90α N-terminal domain could not be achieved successfully. Although different thiol catalysts and minor differences in the reaction conditions as well as different buffers for the purification were tried, there are still unchecked possibilities left. The desulfurization protocol has yet to be tried on purified product. However, even if successfully applied on purified semi -synthetic Hsp90 NTD this would take away the proclaimed benefit of *in situ* compatibility which is crucial due to the detrimental yield loss when submitting the NTD to RP-HPLC. The semi-synthetic strategy was already implemented and proofed to work, even though challenging due to the poor solubility of the truncated N-terminal recombinant construct. If successful desulfurization is achieved, optimization of upscaling and refolding would yield modified Hsp90α NTD in sufficient amounts for subsequent characterization and biochemical assays to determine structure-activity relationships of said modified variants.

The implementation of a new full-length semi-synthetic strategy for the generation of unmodified and modified Hsp90α variants was successful. The final purification step of the recombinant product still needs to be optimized, and there are lots of options that have not been considered yet, like hydrophobic interactions chromatography (HIC). When purification conditions are optimized, upscaling and refolding protocol needs to be established in order to characterize and investigate the effects of different modifications in the middle and C-terminal domain on Hsp90α function and activity. Additionally, one could think about altering the established strategy to a small amount by truncating the recombinant construct of the NTD (1-527). This would enable the researcher to implement modifications throughout the whole Hsp90α molecule, since modifications can only be introduced in the middle and C-terminal domain with the herein described protocol.

7. References

(1) Schopf, F. H., Biebl, M. M., and Buchner, J. (2017) The HSP90 chaperone machinery, *Nature reviews. Molecular cell biology 18,* 345–360 published online Apr 21, 2017. DOI: 10.1038/nrm.2017.20.

(2) Hu, C., Yang, J., Qi, Z., Wu, H., Wang, B., Zou, F., Mei, H., Liu, J., Wang, W., and Liu, Q. (2022) Heat shock proteins: Biological functions, pathological roles, and therapeutic opportunities, *MedComm 3,* e161 published online Aug 2, 2022. DOI: 10.1002/mco2.161.

(3) Hartl, F. U. (1996) Molecular chaperones in cellular protein folding, *Nature 381,* 571–579. DOI: 10.1038/381571a0.

(4) Rehn, A. B., and Buchner, J. (2015) p23 and Aha1, *Sub-cellular biochemistry 78,* 113–131. DOI: 10.1007/978-3-319-11731-7_6.

(5) Biebl, M. M., and Buchner, J. (2019) Structure, Function, and Regulation of the Hsp90 Machinery, *Cold Spring Harbor perspectives in biology 11 published online* Sep 3, 2019. DOI: 10.1101/cshperspect.a034017.

(6) Miyata, Y., Nakamoto, H., and Neckers, L. (2013) The therapeutic target Hsp90 and cancer hallmarks, *Current pharmaceutical design 19,* 347–365. DOI: 10.2174/138161213804143725.

(7) Jaeger, A. M., and Whitesell, L. (2019) HSP90: Enabler of Cancer Adaptation, *Annu. Rev. Cancer Biol. 3,* 275–297. DOI: 10.1146/annurev-cancerbio-030518-055533.

(8) Scheibel, T., Neuhofen, S., Weikl, T., Mayr, C., Reinstein, J., Vogel, P. D., and Buchner, J. (1997) ATP-binding properties of human Hsp90, *Journal of Biological Chemistry 272,* 18608–18613. DOI: 10.1074/jbc.272.30.18608.

(9) Kumalo, H. M., Bhakat, S., and Soliman, M. E. (2015) Heat-shock protein 90 (Hsp90) as anticancer target for drug discovery: an ample computational perspective, *Chemical biology & drug design 86,* 1131–1160 published online Jun 11, 2015. DOI: 10.1111/cbdd.12582.

(10) Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1997) Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone, *Cell 90,* 65–75. DOI: 10.1016/s0092-8674(00)80314-1.

(11) Jackson, S. E. (2013) Hsp90: structure and function, *Topics in current chemistry 328,* 155–240. DOI: 10.1007/128_2012_356.

(12) Prodromou, C., and Pearl, L. H. (2003) Structure and functional relationships of Hsp90, *Current cancer drug targets 3,* 301–323. DOI: 10.2174/1568009033481877.

(13) Pearl, L. H., and Prodromou, C. (2000) Structure and in vivo function of Hsp90, *Current opinion in structural biology 10,* 46–51. DOI: 10.1016/s0959-440x(99)00047-0.

(14) Hainzl, O., Lapina, M. C., Buchner, J., and Richter, K. (2009) The charged linker region is an important regulator of Hsp90 function, *The Journal of biological chemistry 284,* 22559–22567 published online Jun 24, 2009. DOI: 10.1074/jbc.M109.031658.

(15) Jahn, M., Rehn, A., Pelz, B., Hellenkamp, B., Richter, K., Rief, M., Buchner, J., and Hugel, T. (2014) The charged linker of the molecular chaperone Hsp90 modulates domain contacts and biological function, *Proceedings of the National Academy of Sciences of the United States of America 111,* 17881–17886 published online Dec 2, 2014. DOI: 10.1073/pnas.1414073111.

(16) Zuehlke, A. D., and Johnson, J. L. (2012) Chaperoning the chaperone: a role for the co-chaperone Cpr7 in modulating Hsp90 function in Saccharomyces cerevisiae, *Genetics 191,* 805–814 published online Apr 13, 2012. DOI: 10.1534/genetics.112.140319.

(17) Tsutsumi, S., Mollapour, M., Prodromou, C., Lee, C.-T., Panaretou, B., Yoshida, S., Mayer, M. P., and Neckers, L. M. (2012) Charged linker sequence modulates eukaryotic heat shock protein 90 (Hsp90) chaperone activity, *Proceedings of the National Academy of Sciences of the United States of America 109,* 2937–2942 published online Feb 6, 2012. DOI: 10.1073/pnas.1114414109.

(18) Shiau, A. K., Harris, S. F., Southworth, D. R., and Agard, D. A. (2006) Structural Analysis of E. coli hsp90 reveals dramatic nucleotide-dependent conformational rearrangements, *Cell 127,* 329–340. DOI: 10.1016/j.cell.2006.09.027.

(19) Dollins, D. E., Warren, J. J., Immormino, R. M., and Gewirth, D. T. (2007) Structures of GRP94-nucleotide complexes reveal mechanistic differences between the hsp90 chaperones, *Molecular cell 28,* 41–56. DOI: 10.1016/j.molcel.2007.08.024.

(20) Hellenkamp, B., Wortmann, P., Kandzia, F., Zacharias, M., and Hugel, T. (2017) Multidomain structure and correlated dynamics determined by self-consistent FRET networks, *Nature methods 14,* 174–180 published online Dec 5, 2016. DOI: 10.1038/nmeth.4081.

(21) Cunningham, C. N., Krukenberg, K. A., and Agard, D. A. (2008) Intra- and intermonomer interactions are required to synergistically facilitate ATP hydrolysis in Hsp90, *Journal of Biological Chemistry 283,* 21170–21178 published online May 20, 2008. DOI: 10.1074/jbc.M800046200.

(22) Horwich, A. L., Buchner, J., Smock, R. G., Gierasch, L. M., and Saibil, H. R. (2012) 3.10 Chaperones and Protein Folding. in *Comprehensive biophysics.* (Egelman, E. H., and Wirtz, D., Eds.), Elsevier, Amsterdam.

(23) Panaretou, B., Siligardi, G., Meyer, P., Maloney, A., Sullivan, J. K., Singh, S., Millson, S. H., Clarke, P. A., Naaby-Hansen, S., Stein, R., Cramer, R., Mollapour, M., Workman, P., Piper, P. W., Pearl, L. H., and Prodromou, C. (2002) Activation of the ATPase activity of hsp90 by the stress-regulated cochaperone aha1, *Molecular cell 10,* 1307–1318. DOI: 10.1016/s1097-2765(02)00785-2.

(24) Li, J., Richter, K., and Buchner, J. (2011) Mixed Hsp90-cochaperone complexes are important for the progression of the reaction cycle, *Nature structural & molecular biology 18,* 61–66 published online Dec 19, 2010. DOI: 10.1038/nsmb.1965.

(25) Pratt, W. B., and Toft, D. O. (1997) Steroid receptor interactions with heat shock protein and immunophilin chaperones, *Endocrine reviews 18,* 306–360. DOI: 10.1210/edrv.18.3.0303.

(26) Richter, K., Muschler, P., Hainzl, O., Reinstein, J., and Buchner, J. (2003) Sti1 is a non-competitive inhibitor of the Hsp90 ATPase. Binding prevents the N-terminal dimerization reaction during the atpase cycle, *Journal of Biological Chemistry 278,* 10328–10333 published online Jan 13, 2003. DOI: 10.1074/jbc.M213094200.

(27) Csermely, P., Miyata, Y., Söti, C., and Yahara, I. (1997) Binding affinity of proteins to hsp90 correlates with both hydrophobicity and positive charges. A surface plasmon resonance study, *Life sciences 61,* 411–418. DOI: 10.1016/S0024-3205(97)00398-6.

(28) Prodromou, C., Siligardi, G., O'Brien, R., Woolfson, D. N., Regan, L., Panaretou, B., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1999) Regulation of Hsp90 ATPase activity by tetratricopeptide repeat (TPR)-domain co-chaperones, *The EMBO journal 18,* 754–762. DOI: 10.1093/emboj/18.3.754.

(29) Meyer, P., Prodromou, C., Liao, C., Hu, B., Roe, S. M., Vaughan, C. K., Vlasic, I., Panaretou, B., Piper, P. W., and Pearl, L. H. (2004) Structural basis for recruitment of the ATPase activator Aha1 to the Hsp90 chaperone machinery, *The EMBO journal 23,* 1402–1410. DOI: 10.1038/sj.emboj.7600141.

(30) Richter, K., Walter, S., and Buchner, J. (2004) The Co-chaperone Sba1 connects the ATPase reaction of Hsp90 to the progression of the chaperone cycle, *Journal of molecular biology 342,* 1403–1413. DOI: 10.1016/j.jmb.2004.07.064.

(31) Vaughan, C. K., Gohlke, U., Sobott, F., Good, V. M., Ali, M. M. U., Prodromou, C., Robinson, C. V., Saibil, H. R., and Pearl, L. H. (2006) Structure of an Hsp90-Cdc37-Cdk4 complex, *Molecular cell 23,* 697–707. DOI: 10.1016/j.molcel.2006.07.016.

(32) Backe, S. J., Sager, R. A., Woodford, M. R., Makedon, A. M., and Mollapour, M. (2020) Post-translational modifications of Hsp90 and translating the chaperone code, *The Journal of biological chemistry 295,* 11099–11117 published online Jun 11, 2020. DOI: 10.1074/jbc.REV120.011833.

(33) Cloutier, P., and Coulombe, B. (2013) Regulation of molecular chaperones through post-translational modifications: decrypting the chaperone code, *Biochimica et biophysica acta 1829,* 443–454 published online Feb 28, 2013. DOI: 10.1016/j.bbagrm.2013.02.010.

(34) Cloutier, P., Lavallée-Adam, M., Faubert, D., Blanchette, M., and Coulombe, B. (2013) A newly uncovered group of distantly related lysine methyltransferases preferentially interact with molecular chaperones to regulate their activity, *PLoS genetics 9,* e1003210 published online Jan 17, 2013. DOI: 10.1371/journal.pgen.1003210.

(35) Nitika, and Truman, A. W. (2017) Cracking the Chaperone Code: Cellular Roles for Hsp70 Phosphorylation, *Trends in biochemical sciences 42,* 932–935 published online Nov 5, 2017. DOI: 10.1016/j.tibs.2017.10.002.

(36) Somale, D., Di Nardo, G., Di Blasio, L., Puliafito, A., Vara-Messler, M., Chiaverina, G., Palmiero, M., Monica, V., Gilardi, G., Primo, L., and Gagliardi, P. A. (2020) Activation of RSK by phosphomimetic substitution in the activation loop is prevented by structural constraints, *Scientific reports 10,* 591 published online Jan 17, 2020. DOI: 10.1038/s41598-019-56937-3.

(37) Mollapour, M., Tsutsumi, S., Truman, A. W., Xu, W., Vaughan, C. K., Beebe, K., Konstantinova, A., Vourganti, S., Panaretou, B., Piper, P. W., Trepel, J. B., Prodromou, C., Pearl, L. H., and Neckers, L. (2011) Threonine 22 phosphorylation attenuates Hsp90 interaction with cochaperones and affects its chaperone activity, *Molecular cell 41,* 672–681. DOI: 10.1016/j.molcel.2011.02.011.

(38) Nguyen, M. T. N., Knieß, R. A., Daturpalli, S., Le Breton, L., Ke, X., Chen, X., and Mayer, M. P. (2017) Isoform-Specific Phosphorylation in Human Hsp90β Affects Interaction with Clients and the Cochaperone Cdc37, *Journal of molecular biology 429,* 732–752 published online Jan 19, 2017. DOI: 10.1016/j.jmb.2017.01.011.

(39) Wang, X., Lu, X., Song, X., Zhuo, W., Jia, L., Jiang, Y., and Luo, Y. (2012) Thr90 phosphorylation of Hsp90α by protein kinase A regulates its chaperone machinery, *The Biochemical journal 441,* 387–397. DOI: 10.1042/BJ20110855.

(40) Muller, P., Ruckova, E., Halada, P., Coates, P. J., Hrstka, R., Lane, D. P., and Vojtesek, B. (2013) C-terminal phosphorylation of Hsp70 and Hsp90 regulates alternate binding to co-chaperones CHIP and HOP to determine cellular protein folding/degradation balances, *Oncogene 32,* 3101–3110 published online Jul 23, 2012. DOI: 10.1038/onc.2012.314.

(41) Woo, S. H., An, S., Lee, H.-C., Jin, H.-O., Seo, S.-K., Yoo, D.-H., Lee, K.-H., Rhee, C. H., Choi, E.-J., Hong, S.-I., and Park, I.-C. (2009) A truncated form of p23 down-regulates telomerase activity via disruption of Hsp90 function, *The Journal of biological chemistry 284,* 30871–30880 published online Sep 9, 2009. DOI: 10.1074/jbc.M109.052720.

(42) Soroka, J., Wandinger, S. K., Mäusbacher, N., Schreiber, T., Richter, K., Daub, H., and Buchner, J. (2012) Conformational switching of the molecular chaperone Hsp90 via regulated phosphorylation, *Molecular cell 45,* 517–528. DOI: 10.1016/j.molcel.2011.12.031.

(43) Yang, Y., Rao, R., Shen, J., Tang, Y., Fiskus, W., Nechtman, J., Atadja, P., and Bhalla, K. (2008) Role of acetylation and extracellular location of heat shock protein 90alpha in tumor cell invasion, *Cancer research 68,* 4833–4842. DOI: 10.1158/0008-5472.CAN-08-0644.

(44) Echeverria, P. C., Bhattacharya, K., Joshi, A., Wang, T., and Picard, D. (2019) The sensitivity to Hsp90 inhibitors of both normal and oncogenically transformed cells is determined by the equilibrium between cellular quiescence and activity, *PloS one 14,* e0208287 published online Feb 6, 2019. DOI: 10.1371/journal.pone.0208287.

(45) Fischer, E., and Fourneau, E. (1901) Ueber einige Derivate des Glykocolls, *Ber. Dtsch. Chem. Ges. 34,* 2868–2877. DOI: 10.1002/cber.190103402249.

(46) Fischer, E. (1907) Faraday lecture. Synthetical chemistry in its relation to biology, *J. Chem. Soc., Trans. 91,* 1749. DOI: 10.1039/ct9079101749.

(47) Bergmann, M., and Zervas, L. (1932) Über ein allgemeines Verfahren der Peptid-Synthese, *Ber. dtsch. Chem. Ges. A/B 65,* 1192–1201. DOI: 10.1002/cber.19320650722.

(48) Du Vigneaud, V., Ressler, C., Swan, C. J. M., Roberts, C. W., Katsoyannis, P. G., and Gordon, S. (1953) THE SYNTHESIS OF AN OCTAPEPTIDE AMIDE WITH THE HORMONAL ACTIVITY OF OXYTOCIN, *J. Am. Chem. Soc. 75,* 4879–4880. DOI: 10.1021/JA01115A553.

(49) Merrifield, R. B. (1963) Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide, *J. Am. Chem. Soc. 85,* 2149–2154. DOI: 10.1021/ja00897a025.

(50) Behrendt, R., White, P., and Offer, J. (2016) Advances in Fmoc solid-phase peptide synthesis, *Journal of peptide science : an official publication of the European Peptide Society 22,* 4–27. DOI: 10.1002/psc.2836.

(51) Prabakaran, S., Lippens, G., Steen, H., and Gunawardena, J. (2012) Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding, *Wiley interdisciplinary reviews. Systems biology and medicine 4,* 565–583 published online Aug 15, 2012. DOI: 10.1002/wsbm.1185.

(52) Walsh, C. T., Garneau-Tsodikova, S., and Gatto, G. J. (2005) Protein posttranslational modifications: the chemistry of proteome diversifications, *Angewandte Chemie (International ed. in English) 44,* 7342–7372. DOI: 10.1002/anie.200501023.

(53) Lothrop, A. P., Torres, M. P., and Fuchs, S. M. (2013) Deciphering post-translational modification codes, *FEBS letters 587,* 1247–1257 published online Feb 10, 2013. DOI: 10.1016/j.febslet.2013.01.047.

(54) Kouzarides, T. (2007) Chromatin modifications and their function, *Cell 128,* 693–705. DOI: 10.1016/j.cell.2007.02.005.

(55) Cohen, P. (2000) The regulation of protein function by multisite phosphorylation--a 25 year update, *Trends in biochemical sciences 25,* 596–601. DOI: 10.1016/s0968-0004(00)01712-6.

(56) Kehoe, J. W., and Bertozzi, C. R. (2000) Tyrosine sulfation: a modulator of extracellular protein-protein interactions, *Chemistry & biology 7,* R57-61. DOI: 10.1016/s1074-5521(00)00093-4.

(57) Ludeman, J. P., and Stone, M. J. (2014) The structural role of receptor tyrosine sulfation in chemokine recognition, *British journal of pharmacology 171,* 1167–1179. DOI: 10.1111/bph.12455.

(58) Paik, W. K., Paik, D. C., and Kim, S. (2007) Historical review: the field of protein methylation, *Trends in biochemical sciences 32,* 146–152 published online Feb 8, 2007. DOI: 10.1016/j.tibs.2007.01.006.

(59) Hartrampf, N., Saebi, A., Poskus, M., Gates, Z. P., Callahan, A. J., Cowfer, A. E., Hanna, S., Antilla, S., Schissel, C. K., Quartararo, A. J., Ye, X., Mijalis, A. J., Simon, M. D., Loas, A., Liu, S., Jessen, C., Nielsen, T. E., and Pentelute, B. L. (2020) Synthesis of proteins by

automated flow chemistry, *Science (New York, N.Y.) 368,* 980–987. DOI: 10.1126/science.abb2491.

(60) Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) Synthesis of proteins by native chemical ligation, *Science (New York, N.Y.) 266,* 776–779. DOI: 10.1126/science.7973629.

(61) Conibear, A. C., Watson, E. E., Payne, R. J., and Becker, C. F. W. (2018) Native chemical ligation in protein synthesis and semi-synthesis, *Chemical Society reviews 47,* 9046–9068 published online Nov 12, 2018. DOI: 10.1039/c8cs00573g.

(62) Zheng, J.-S., Tang, S., Qi, Y.-K., Wang, Z.-P., and Liu, L. (2013) Chemical synthesis of proteins using peptide hydrazides as thioester surrogates, *Nature protocols 8,* 2483–2495 published online Nov 14, 2013. DOI: 10.1038/nprot.2013.152.

(63) Yan, L. Z., and Dawson, P. E. (2001) Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization, *J. Am. Chem. Soc. 123,* 526–533. DOI: 10.1021/ja003265m.

(64) Wan, Q., and Danishefsky, S. J. (2007) Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides, *Angewandte Chemie (International ed. in English) 46,* 9248–9252. DOI: 10.1002/anie.200704195.

(65) Kulkarni, S. S., Watson, E. E., Premdjee, B., Conde-Frieboes, K. W., and Payne, R. J. (2019) Diselenide-selenoester ligation for chemical protein synthesis, *Nature protocols 14,* 2229–2257 published online Jun 21, 2019. DOI: 10.1038/s41596-019-0180-4.

(66) Kulkarni, S. S., Watson, E. E., Maxwell, J. W. C., Niederacher, G., Johansen-Leete, J., Huhmann, S., Mukherjee, S., Norman, A. R., Kriegesmann, J., Becker, C. F. W., and Payne, R. J. (2022) Expressed Protein Selenoester Ligation, *Angewandte Chemie (International ed. in English) 61,* e202200163 published online Mar 9, 2022. DOI: 10.1002/anie.202200163.

(67) Sayers, J., Karpati, P. M. T., Mitchell, N. J., Goldys, A. M., Kwong, S. M., Firth, N., Chan, B., and Payne, R. J. (2018) Construction of Challenging Proline-Proline Junctions via Diselenide-Selenoester Ligation Chemistry, *J. Am. Chem. Soc. 140,* 13327–13334 published online Oct 8, 2018. DOI: 10.1021/jacs.8b07877.

(68) Mitchell, N. J., Malins, L. R., Liu, X., Thompson, R. E., Chan, B., Radom, L., and Payne, R. J. (2015) Rapid additive-free selenocystine-selenoester peptide ligation, *J. Am. Chem. Soc. 137,* 14011–14014 published online Nov 3, 2015. DOI: 10.1021/jacs.5b07237.

(69) Sun, Z., Ma, W., Cao, Y., Wei, T., Mo, X., Chow, H. Y., Tan, Y., Cheung, C. H., Liu, J., Lee, H. K., Tse, E. C., Liu, H., and Li, X. (2022) Superfast desulfurization for protein chemical synthesis and modification, *Chem 8,* 2542–2557. DOI: 10.1016/j.chempr.2022.07.017.

Supplementary Data

A

```
TGT TTT CAG GCA GAA ATA GCT CAA CTA ATG TCG CTG ATC ATC AAC ACC TTT
TAC AGC AAC AAG GAG ATC TTT CTG CGT GAA TTG ATA AGC AAT AGC TCC GAC
GCT CTG GAT AAG ATT CGT TAT GAG TCC TTA ACC GAT CCA AGC AAA TTG GAC
AGC GGT AAA GAG CTC CAT ATT AAT CTG ATT CCG AAT AAG CAA GAT CGT ACC
TTG ACG ATT GTC GAT ACT GGT ATT GGC ATG ACC AAG GCG GAT CTG ATT AAC
AAC CTG GGT ACG ATC GCC AAA TCC GGC ACG AAG GCG TTC ATG GAA GCC CTG
CAA GCT GGC GCA GAC ATC TCT ATG ATC GGT CAG TTC GGC GTT GGT TTT TAC
AGC GCG TAT CTG GTT GCG GAA AAG GTG ACC GTG ATC ACC AAG CAC AAC GAC
GAC GAA CAA TAC GCG TGG GAA AGC AGT GCA GGT GGT AGC TTC ACC GTG CGC
ACC GAC ACC GGC GAA CCG ATG GGC CGT GGT ACG AAG GTC ATC CTT CAC CTG
AAG GAG GAT CAG ACC GAA TAT CTG GAA GAG CGC CGT ATT AAA GAG ATC GTG
AAA AAA CAT TCT CAG TTC ATC GGC TAC CCG ATT ACC TTG TTT GTT GAA AAA
GAG CGC GAT AAA GAG GTT AGC GAC GAC GAG GCG CAT CAT CAC CAC CAC CAC
```

B

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

S1: **A)** Codon sequence of Hsp90^{C8}. **B)** Amino acid sequence of Hsp90^{C8}.

A

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

CTA TTT GTG CCG CGT CGT GCG CCG TTT GAT CTG TTC GAA AAC CGT AAA AAA
AAG AAT AAC ATC AAG TTA TAT GTG CGT CGC GTG TTT ATT ATG GAT AAC TGC
GAA GAA CTG ATC CCG GAG TAC CTG AAC TTT ATC CGT GGT GTC GTG GAT AGC
GAA GAT TTG CCG CTG AAC ATT AGC CGT GAA ATG CTG CAA CAG AGC AAG ATT
CTG AAA GTA ATC CGC AAA AAT CTG GTT AAA AAG TGC CTG GAG TTG TTC ACG
GAA TTA GCG GAG GAT AAA GAG AAC TAT AAA AAA TTC TAC GAA CAA TTC AGC
AAA AAT ATC AAA CTG GGC ATT CAT GAA GAT AGC CAA AAC CGC AAA AAG CTC
TCC GAG TTG CTG CGT TAT TAC ACC AGC GCC AGC GGT GAT GAG ATG GTT AGC
CTG AAG GAC TAC TGT ACC CGC ATG AAA GAG AAC CAG AAA CAC ATC TAC TAC
ATT ACT GGT GAG ACG AAG GAC CAG GTT GCT AAT AGT GCT TTC GTG GAG CGC
CTG AGA AAG CAC GGC CTG GAA GTG ATA TAT ATG ATC GAA CCG ATT GAC GAG
TAC

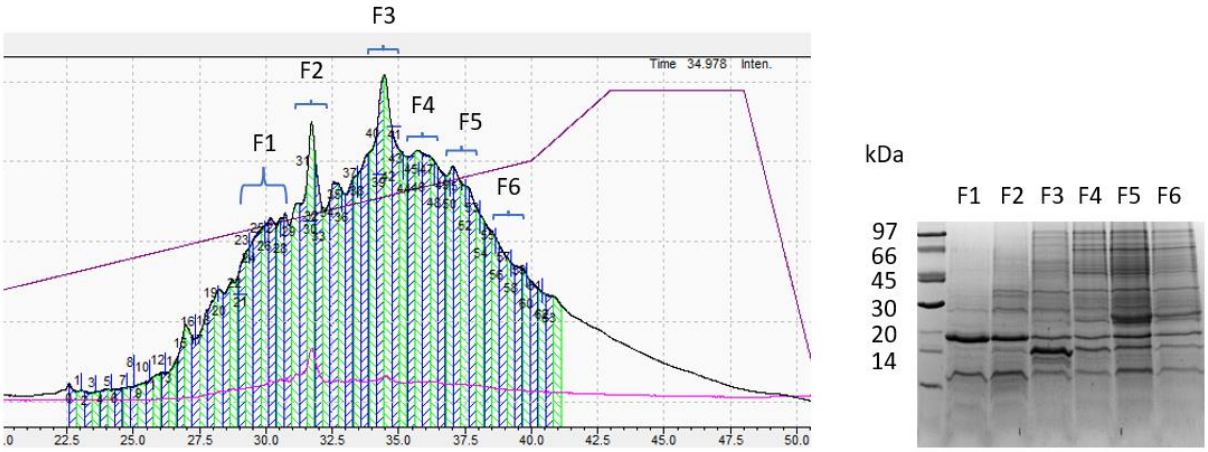
TGT ATA ACA GGA GAT GCA TTA GTA GCT CTA CCG GAA GGT GAG AGC GTT CGT
ATT GCT GAC ATC GTT CCG GGC GCA CGC CCT AAT TCG GAT AAC GCA ATC GAC
CTG AAG GTG CTG GAT CGT CAT GGC AAT CCG GTC CTG GCC GAC CGC CTG TTT
CAT TCC GGC GAG CAC CCG GTG TAC ACC GTA CGC ACC GTT GAA GGC CTG CGT
GTT ACG GGC ACC GCC AAT CAT CCG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

B

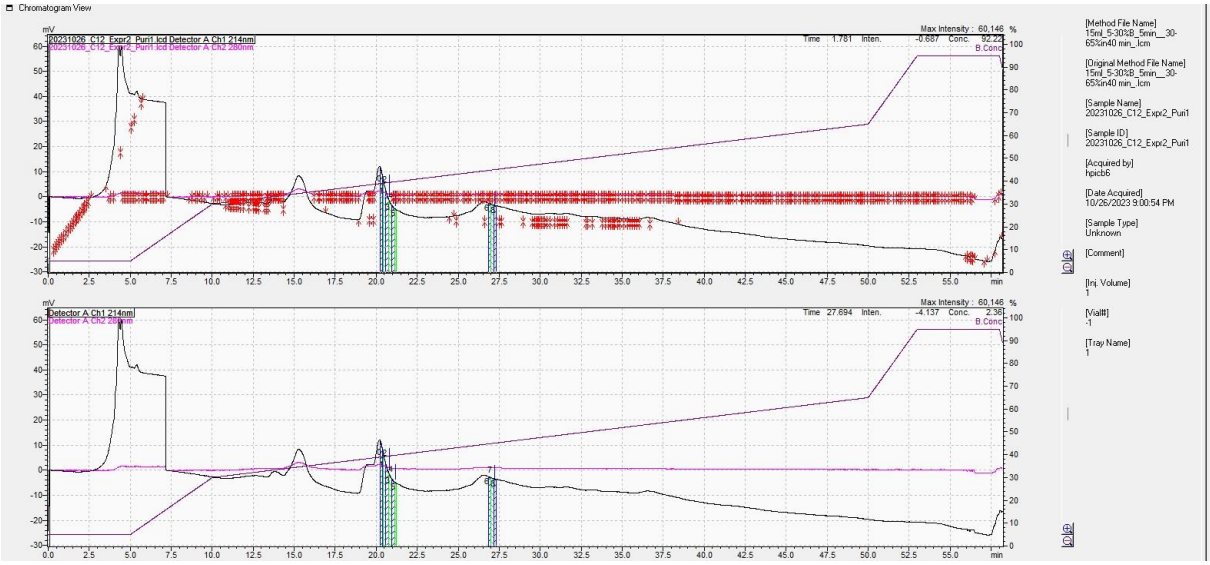
MPEETQTQDQ PMEEEEVETF AFQAEIAQLM SLIINTFYSN KEIFLRELIS NSSDALDKIR
YESLTDPSKL DSGKELHINL IPNKQDRTLT IVDTGIGMTK ADLINNLGTI AKSGTKAFME
ALQAGADISM IGQFGVGFYS AYLVAEKVTV ITKHNDDEQY AWESSAGGSF TVRTDTGEPM
GRGTKVILHL KEDQTEYLEE RRIKEIVKKH SQFIGYPITL FVEKERDKEV SDDEAEEKED
KEEEKEKEEK ESEDKPEIED VGSDEEEEKK DGDKKKKKKI 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

CITGDALVAL PEGESVRIAD IVPGARPNSD NAIDLKVLDR HGNPVLADRL FHSGEHPVYT
VRTVEGLRVT GTANHPLLCL VDVAGVPTLL WKLIDEIKPG DYAVIQRSAF SVDCAGFARG
KPEFAPTTYT VGVPGLVRFL EAHHRDPDAQ AIADELTDGR FYYAKVASVT DAGVQPVYSL
RVDTADHAFI TNGFVSHATG LTGIHHHHHH HSGLNSGLTT NPGVSAWQVN TAYTAGQLVT
YNGKTYKCLQ PHTSLAGWEP SNVPALWQLQ

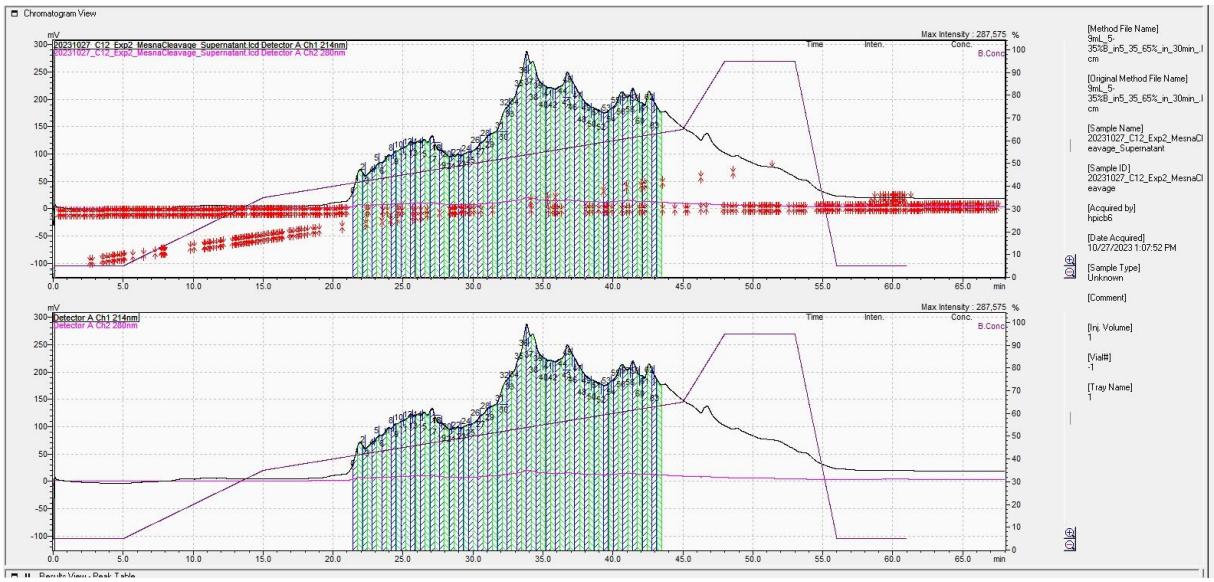
S2: **A)** Codon sequence of Hsp90^{C12}. **B)** Amino acid sequence of Hsp90^{C12}.



S3: **Left:** Chromatogram of semipreparative purification of the flow through after "reverse" Ni-NTA column for full-length construct C^{12}. **Right:** SDS-PAGE of the collected fractions of interest after semipreparative purification.



S4: Chromatogram of semipreparative purification of the solubilized pellet after MesNa cleavage of full-length construct C^{12}.



S5: Chromatogram of semipreparative purification of the supernatant after MesNa cleavage of full-length construct C^{12}.

Acknowledgements

First and foremost I would like to thank Univ.-Prof. Dr. Christian F.W. Becker for the opportunity to finish my master's degree in his group and for his expert advice during meetings and talks.

I would like to acknowledge my supervisor MSc. Oliver Gajsek for his valuable guidance and encouragement throughout this project. I deeply appreciate the opportunity to have learned from such an experienced and motivated scientist and I am grateful for his patience and support.

I want to thank the whole Becker research group at the Institute of Biochemistry for their kind demeanor and the great time inside and outside of the laboratory.

Furthermore, I want to thank my friends for their continuing support, inspiration and motivation.

Finally, I would like to thank my family and my partner for their love and for offering comfort in difficult times. I will always be thankful.