Cholera toxin and cytolethal distending toxin

Expression, purification and stability with host machineries

by

Siri Luise Tveitan

Thesis submitted for the degree of
Master of Science in Molecular Bioscience
60 credits

Department of Biosciences
Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

July 2018

© Siri Luise Tveitan
2018

Cholera toxin and cytolethal distending toxin:
Expression, purification and stability with host machineries

Siri Luise Tveitan

http://www.duo.uio.no/
Print: Reprosentralen, University of Oslo

Acknowledgements

This thesis concludes my master's degree at the Department of Biosciences, University of Oslo. The work was carried out in the laboratory of Professor Ute Krengel at the Department of Chemistry, University of Oslo, between August 2016 and July 2018.

First and foremost, I would like to thank my main supervisor Ute for always following my progress and giving me sound advice whenever I was at a loss. Your high expectations have pushed me to work harder and be more organized in the lab. Thank you for keeping me on my toes while at the same time making sure that I do not tire myself out. I want to thank my co-supervisor Joël for his extensive assistance and support in the lab. You always have an answer, a suggestion or a solution to my problems. Furthermore, I want to show my gratitude to the rest of the members of our group. Thank you for all the coffee and ice cream breaks, birthday cakes, group lunches and section beers. You made the Protein Dungeon feel so much like home that it was sometimes difficult to focus on the lab work.

Even though most of my friends have no idea what I've been working on for the last two years, I want to thank them for keeping my spirits up and my stress levels down. A special thank you goes to Taran and Vigdis for always being there for me, despite our busy schedules.

Finally, my wonderful mom deserves her own paragraph of appreciation. You have fed with both love and lasagna through these last few months, and I don't think I could have finished this degree without all the support that you give me every single day. Thank you.

Siri,
July 2018

Abstract

Pathogenic bacteria can induce symptoms of disease in a host organism, such as a light fever or a localized inflammation. Highly invasive pathogens secrete potent toxins that induce symptoms often resulting in long-term health complications or a deadly outcome. A specialized form of intracellular toxins enters host cells in a complex with their membrane-binding components and work their way through the cells until they dissociate as free subunits. They exert their enzymatic activity once they reach their final destination in a stable form, which is often achieved by binding to host machineries to remain stable. One such toxin is the active subunit of cholera toxin (CTA1) of *Vibrio cholerae* (*V. cholerae*), which depends on the host heat shock protein 90 (HSP90) for translocation to the cytosol, proper folding, as well as maintaining a folded and active state. Without HSP90, CTA1 would not be able to induce its severe symptom of explosive diarrhea. Another toxin that possibly depends on a host machinery for its activation is the cytolethal distending toxin (CDT) of *Haemophilus ducreyi* (*H. ducreyi*). Its B subunit (CdtB) exerts DNA damage in the cell nucleus, inducing a range of diseases, but no host machinery-dependence have previously been documented for this protein. New insights into this protein from the collaborating group of Ken Teter at UCF now indicate that it might be adenosine triphosphate (ATP)-dependent. The structures of both toxins, as well as their binding partners, are well established, but their interacting residues are still a mystery.

This thesis focuses on the expression and purification of CTA1 and CdtB constructs and their stability both with and without host machineries. Established protocols were refined to obtain more stable conditions in which to perform the experiments presented in this thesis as well as future experiments. CTA1 and HSP90 constructs both with and without affinity tags were successfully produced and incubated under physiological conditions to test for CTA1 stability and possible complex formation. Solutions were analyzed with gel electrophoresis, as well as with a novel method of affinity chromatography with the potentially formed complex. Gel electrophoresis analysis of incubated tagged CTA1 with a full-length HSP90 construct proved to stabilize CTA1, possibly in a refolded complex that could be crystallized. CdtB was successfully purified, but a new buffer system was determined for this protein for stabilizing purposes. Optimized conditions provided insights into possible or impossible conditions for crystal screening, which were performed in parallel with buffer optimization. Addition of ATP provided a more stable protein at room temperature, which is could be useful for future purifications and attempts of crystal formation.

Abstract in Norwegian

Patogene bakterier kan indusere symptomer av sykdom i en vertsorganisme, for eksempel ved en lett feber eller en lokalisert betennelse. Svært invaderende patogener skiller ut potente toksiner (giftstoffer) og induserer symptomer som ofte resulterer i langsiktige helse-komplikasjoner eller et dødelig utfall for vertsorganismen. En spesialisert form av intracellulære toksiner går inn i vertsceller i et kompleks med deres membranbindende komponenter og arbeider seg gjennom cellene til de dissosieres som frie enheter. De utøver sin enzymatiske aktivitet når de når sin endelige destinasjon i en stabil form, som ofte oppnås ved å binde til vertsmaskiner for å forbli stabil. Ett slikt toksin er den aktive enheten til koleratoksin (CTA1) av *Vibrio cholerae* (*V. cholerae*), som avhenger av vertsproteinet heat shock protein 90 (HSP90) for translokasjon til cytosol, korrekt folding, samt opprettholdelse av sin aktive tilstand. Uten HSP90 ville CTA1 ikke kunne fremkalle den eksplosive diaréen som kolera er kjent for. Et annet toksin som muligens avhenger av en vertsmaskin for aktivering er cytolethal distending toxin (CDT) av *Haemophilus ducreyi* (*H. ducreyi*). Dens aktive enhet (CdtB) utøver skade på DNA i cellekjernen, men ingen vertsmaskinavhengighet har tidligere blitt dokumentert for dette proteinet. Ny innsikt fra samarbeidsgruppen til Ken Teter ved UCF viser nå at proteinet kan være adenosin trifosfat (ATP)-avhengig. Strukturene til begge toksinene, så vel som deres bindingspartnere, er godt etablerte, men deres interagerende residuene er fortsatt et mysterium.

Denne oppgaven fokuserer på proteinekspresjon og -rensing av CTA1- og CdtB-konstrukter og deres stabilitet både med og uten vertsmaskiner. Etablerte protokoller ble raffinert for å oppnå mer stabile forhold for utførelse av forsøkene presentert i denne oppgaven, samt fremtidige eksperimenter. CTA1- og HSP90-konstrukter ble produsert og inkubert under fysiologiske forhold for å teste stabiliteten til CTA1 og mulig kompleksdannelse med HSP90. Prøvene ble analysert med gelelektroforese, så vel som med en metode for affinitets-kromatografi av mulig dannet kompleks. Analyse med gelelektroforese av inkubert CTA1 med et HSP90-konstrukt av full lengde viste seg å stabilisere CTA1, muligens i et foldet kompleks som kan krystalliseres. CdtB ble produsert vellykket, men et nytt buffersystem ble bestemt for dette protein med et formål om høyere stabilitet. Optimaliseringen ga innsikt i mulige eller umulige forhold for krystall-screening, som ble utført parallelt med bufferoptimalisering. Tilsetning av ATP ga også et mer stabilt protein ved romtemperatur, hvilket kan være nyttig for fremtidige rensninger og forsøk på krystalldannelse.

Abbreviations

A_{280}	absorbance at 280 nm
A_{260}	absorbance at 260 nm
Amp	ampicillin
ARF	adenosine diphosphate ribosylation factor
ATP	adenosine triphosphate
Bis-Tris	bis(hydroxyethyl)iminotris(hydroxymethyl)methane
Cam	chloramphenicol
CDT	cytolethal distending toxin
CdtA	A subunit of CDT
CdtB	B subunit of CDT
CdtC	C subunit of CDT
CT	cholera toxin
CTA1	A1 fragment of CT
CTA2	A2 fragment of CT
CTB_5	B pentamer of CT
cAMP	cyclic adenosine monophosphate
DNase	deoxyribonuclease
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
ERAD	ER-associated degradation
GFP	green fluorescent protein
HdCDT	*Haemophilus ducreyi* CDT
His_6	tag of six histidines
HSP90	heat shock protein 90
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
Kan	kanamycin
LB	lysogeny broth
LDS	lithium dodecyl sulfate
MES	morpholino ethanesulfate

MME	monomethyl ether
MS	mass spectrometry
Mw	molecular weight
Ni-NTA	nickel-nitriloacetic acid
OD_{600}	optical density measured at 600 nm
ON	overnight
PAGE	polyacrylamide gel electrophoresis
PDI	protein-disulfide isomerase
PEG	polyethylene glycol
pI	isoelectric point
PMSF	phenylmethanesulfonyl fluoride
RT	room temperature
SD200	Superdex 200 10/30 GL
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SPR	surface plasmon resonance
TB	terrific broth
TEV	Tobacco Etch Virus
Tris	tris(hydroxymethyl)aminomethane
UV-vis	ultraviolet-visible

Table of Contents

1 Introduction

1.1 Infectious diseases

Pathogenic microorganisms such as bacteria can cause disease in a host organism by disrupting homeostatic processes or by stimulating an immune response. These diseases may be passed from one host to another through ingestion, inhalation, or contact with bodily fluids. Although most infections only give minor and easily treatable discomforts, some can be life-threatening, and others are even linked to long-term health complications [1, 2].

Once bacterial cells have successfully infected host tissues, the impaired cells may recognize their conserved pathogenic features and secrete cytokines to try to eliminate them, quickly initiating innate immune responses causing symptoms such as fever and localized tissue inflammation [3]. This response is meant to both kill the bacteria and repair the damaged tissue. If the infection still continues to spread, an adaptive immune response may be activated depending on the antigenicity, mode of action, and composition of the bacterial strain. Some strains also exert toxic effects on host cells by the use of toxins, causing a variety of endemic diseases such as cholera, anthrax and pneumonia [4].

For the last century, the leading therapy for both prevention and treatment of bacterial infections has been the use of antimicrobial agents [5]. One of the most widely used therapeutics is penicillin, which breaks down the peptidoglycan layer in the bacterial cell wall by disrupting the cross-linking of peptides by the enzyme transpeptidase. Antibiotics that break down bacterial cell walls are more effective against Gram-positive bacteria, as Gram-negative bacteria have a thinner peptidoglycan layer as well as an additional protective barrier around their cell wall [6]. Mutations within the active site of bacterial enzymes like transpeptidase can make bacteria less sensitive to antibiotics, leaving them to grow and multiply even after treatment. Intensive use and misuse of antibiotics in both humans and animals have increased the accumulation of antibiotic-resistant genes in bacteria, and a growing number of infections are becoming more difficult to treat as a result of this [7]. If this trend continues, there will be a desperate need for alternative ways of targeting pathogenic bacteria.

1.2 Bacterial virulence factors

As single-celled living organisms, bacteria can rapidly reproduce and travel through host systems such as the blood stream and the digestive tract. This ability to spread and infect host tissues is assisted by secretory, membrane-associated and cytosolic molecules known as virulence factors, which are produced by the bacteria themselves [8].

1.2.1 Adhesion factors

Adhesion and colonization of bacterial cells to a host surface involves virulence factors that efficiently adhere the bacteria to epithelial or mucosal tissue. These surface-recognizing molecules, known as *adhesins*, can be membrane-associated components or secreted molecules. They may recognize various classes of host molecules, including transmembrane proteins and components of the extracellular matrix, and are often crucial for bacterial survival in host environments [9]. One example of adhesins are the hair-like appendages known as pili. They are able to bind to receptors in a variety of host tissues and increase the efficiency of colonization. Among these are the toxin-co-regulated type IV pili produced by *Vibrio cholerae* (*V. cholerae*), which bind to cell-receptors of the human small intestine [10]. The same bacterium also secretes an *N*-acetyl glucosamine-binding protein that has been reported to be an important colonization factor [11-13].

1.2.2 Toxins

More invasive bacteria also produce highly potent virulence factors that induce symptoms resulting in the death of both infected cells and the host organism [14]. These types of toxigenic molecules can be classified as either exotoxins or endotoxins. Clinically, both induce common innate immune responses, but their structural properties make them otherwise very different in their effects [15, 16]. Endotoxins are mainly defined in bacteriology as non-specific lipopolysaccharides in the bacterial outer membrane only released upon cell lysis [17]. Exotoxins are secreted proteins that often exert enzymatic activity after entering the host cell. As proteins, these toxins also have high receptor-specificity, resulting in symptoms depending on the subtype of exotoxins.

After secretion from colonized bacteria, exotoxins may cause a variety of damaging or dysfunctional effects based on their mode of action in a host cell. Although there are many ways of classifying these molecules, it is common to group them as cell surface-activating (type I), membrane-damaging (type II) and intracellular (type III) systems. Some of the most potent and virulent bacterial toxins are among the latter, including a superfamily known as AB toxins [18]. These proteins have one or more binding B component that binds the holotoxin to host membranes and one or more enzymatically active A component that displays toxic activity. They are referred to as A_xB_x, where the x represents the number of A and B components. Among the many variants are the enterotoxin cholera toxin (AB_5) and the genotoxin cytolethal distending toxin (AB_2) which exert their effects on the small intestine and DNA, respectively.

1.3 The major virulence factor of *V. cholerae*

V. cholerae is a Gram-negative bacterium that can be found in contaminated waters or food reservoirs [19]. During an epidemic, the contamination usually originates from the feces of an already infected person, but brackish and marine waters are also natural environments for the bacterium [20]. In areas where the purify of sewage and drinking water is poor, the disease can be difficult to contain. Although oral cholera vaccines are available, evidence of long-term protection is limited, and re-vaccination is needed for continuous protection [21].

1.3.1 Cholera toxin

The major virulence factor of *V. cholerae* is the cholera toxin (CT), which is an AB_5 enterotoxin. Sharing 80 % sequence homology with the heat-labile enterotoxin of *Escherichia coli* (*E. coli*), the three-dimensional structure of CT was solved and refined to 2.13 Ångström (Å) in 2004 [22]. The holotoxin consists of five B chains in a ring-shaped B pentamer (CTB_5) and one A chain composed of two fragments designated CTA1 and CTA2. The alpha-helical CTA2 fragment tethers the less structured CTA1 fragment to CTB_5 through mostly non-covalent interactions and one proteolytically cleavable disulfide bond. The crystal structure and the schematic drawing of CT is presented in Figure 1.

The A1 fragment is considered the active component of the AB_5 toxin, as it exerts its toxic effect without the A2 chain. Except for having a tethering effect, CTA2 is believed to stabilize the holotoxin, keeping CTA1 in complex with CTB_5 in its folded state until it reaches the endoplasmic reticulum (ER). In the ER, a protein-disulfide isomerase (PDI) displaces CTA1 by reducing the disulfide bridge than links it with CTA2, resulting in an unfolded CTA1 [23, 24]. PDI is not believed to be the cause of this unfolding, as CTA1 tends to rapidly unfold due to conformational instability as a free subunit [25, 26].

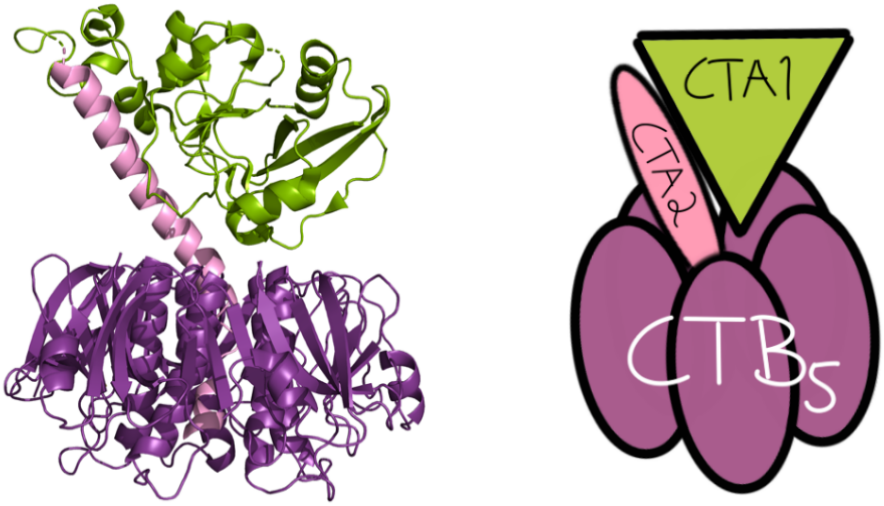

Figure 1 Crystal structure of CT (left) and a schematic drawing of CT (right). The catalytic CTA1 fragment (green) is anchored to the CTB_5 subunit (purple) by the helical CTA2 fragment (pink). PDB: 1S5E [22].

1.3.2 The CT activation pathway

Binding of CTB_5 to cholesterol-rich cell membrane regions (lipid rafts) in epithelial human small intestine cells triggers the uptake of the holotoxin [27, 28]. After endocytosis, vesicular trafficking transports the whole toxin in a retrograde manner through the Golgi apparatus to the ER [29-31]. Here, CTA1 dissociates from the rest of the complex with the help of PDI. The rapidly unfolded CTA1 is left susceptible for ER-associated degradation (ERAD), but it escapes degradation and is translocated to the cytosol. It is proposed that the ERAD escape is due to its lack of sufficient lysine residues necessary for marking a protein for proteasomal degradation by ubiquitination [32].

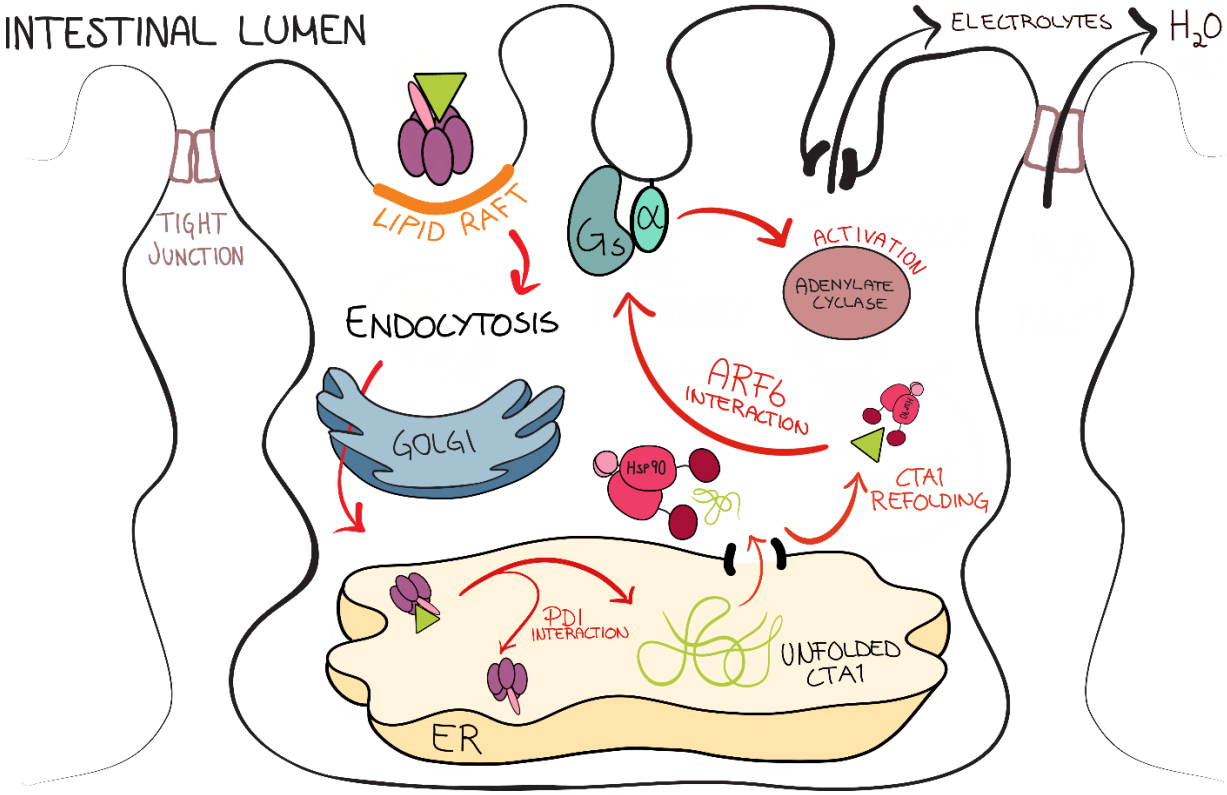

Figure 2 Schematic of the CT activation pathway in an epithelial human small intestine cell. The direction of the pathway and some cellular effects are indicated with red arrows, while efflux of electrolytes and water is indicated with black arrows.

Once inside the cytosol, CTA1 refolds, regaining its three-dimensional structure and in turn its toxic activity. Recent studies show that the adenosine triphosphate (ATP)-dependent heat shock protein 90 (HSP90) is required to both pull CTA1 through the ER membrane and refold it in the process [33, 34], as illustrated in Figure 2. The specific interactions between the two proteins remain unclear.

Refolded, CTA1 interacts with an ADP-ribosylation factor (ARF) called ARF6, further enhancing the activity of the toxin by eliciting dramatic changes in its loop regions [35]. CTA1 hence catalyzes the ADP-ribosylation of the membrane-bound G protein G_S, which in turn activates adenylate cyclase (marked activation in Figure 2), raising the cyclic adenosine monophosphate (cAMP) levels in the cells. Ultimately, this leads to an increased efflux of electrolytes and water into the small intestine, resulting in the severe diarrhea that cholera and other enteric diseases are known for [36].

1.3.2.1 Assisted refolding by the molecular chaperone HSP90

The correct folding and assembly of proteins *in vivo* is often accomplished with the aid of molecular chaperones. These chaperones have multiple domains that can bind a variety of both folded and unfolded proteins, and they are often overly expressed in cells exposed to high levels of stress. One such protein is the heat shock protein HSP90, whose folding abilities are exploited by CTA1. Its expression is connected with the folding of several partially folded proteins involved in cell growth and has been found overexpressed in numerous cancerous cells [37-39]. For those reasons, the structure and function of HSP90 have been extensively studied over the years [40-44].

Hsp90s are expressed by both prokaryotic and eukaryotic organisms, but there are structural variants between and within species. The two major classes in humans, HSP90A and HSP90B, function largely in the same way, but have small structural differences that enable them to operate in different cell types and compartments [45]. Both isoforms share the same conserved domains: An N-terminal domain responsible for nucleotide binding and ATP hydrolysis, a less conserved middle domain with several client protein binding sites, and a C-terminal domain with a hydrophobic patch for dimer formation. The crystal structure of the dimeric form of *Saccharomyces cerevisiae* (*S. cerevisiae*) HSP90, named HSP82, is presented in Figure 3.

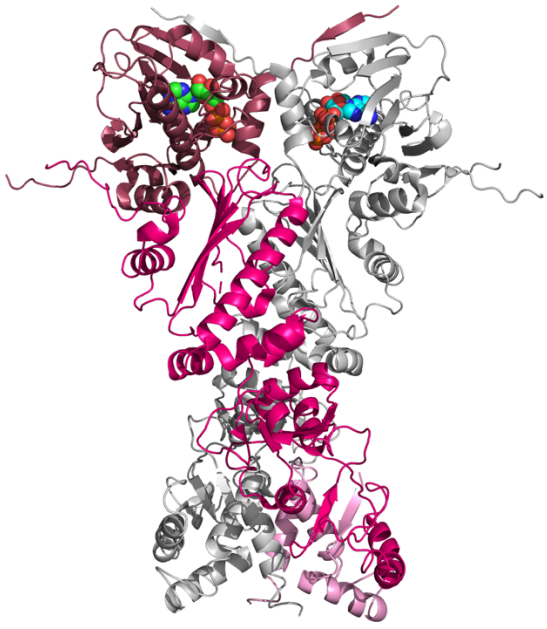

Figure 3 Crystal structure of *S. cerevisiae* HSP90 (HSP82) in a closed dimer conformation. ATP (spheres) is bound to the N-terminal domain (raspberry). Middle domain is presented in hot pink and C-terminal domain as light pink. Only one protomer is displayed in colors to clearly see the dimeric form. The domains are presented in the same colors in Figure 4. (PDB code: 2CG9) [46]

Each HSP90 subunit has a molecular weight of approximately 90 kDa, but it exists primarily as a homodimer in the cytosol of mammalian cells, as C-terminal dimerization is vital for its function [47]. As a dimer, its subunits open and close in response to binding and hydrolysis of ATP, moving in and out of a "fingers crossed" motion [48]. The binding of ATP during N-terminal domain dimerization illustrated in Figure 4.

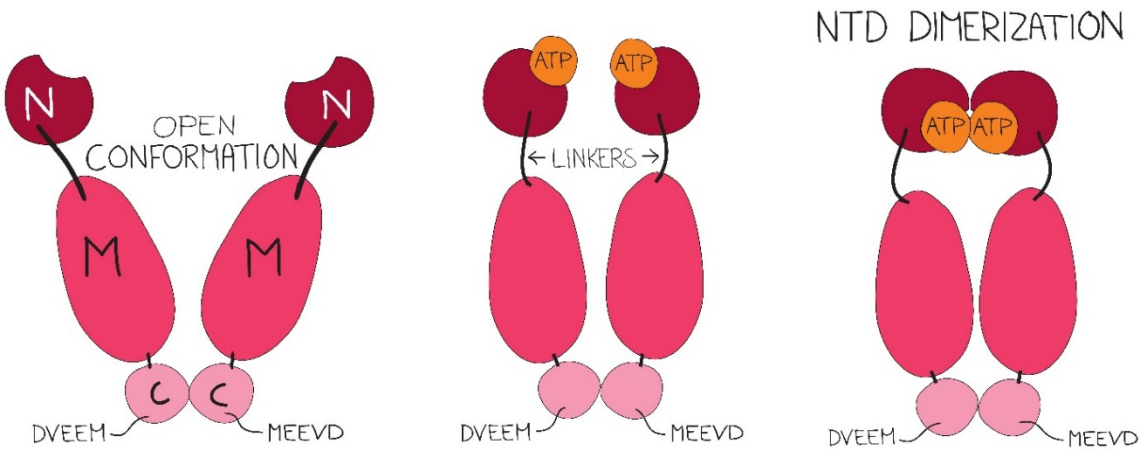

Figure 4 Schematic of ATP binding in HSP90. N-terminal, middle and C-terminal domains are here abbreviated N, M and C, respectively. HSP90 is in its open conformation (far left) when it binds two molecules of ATP (middle), leading to dimerization of the two N-terminal domains (far right) abbreviated NTD dimerization in drawing. After ATP hydrolysis, the two NTDs separate from each other, and the remaining ADP molecules are released. HSP90 will then regain its open conformation, readying it for binding of two new ATP molecules.

The middle domain might be the main discriminator in client protein binding [41, 49], but the other two domains are also known to bind molecules and impact structural changes [50, 51]. C-terminal domains harbor co-chaperone binding sites, specifically the conserved M-E-E-V-D peptide motifs, which are known for binding chaperone proteins containing a tetratricopeptide repeat (TPR) domain [52]. Binding of such molecules to HSP90 can impact its overall activity and conformation. For these reasons, the interactions between HSP90 and CTA1 are difficult to predict without further analyses.

In eukaryotes, a highly charged and conserved linker region connects the N-terminal and the middle domain of each HSP90 subunit. The charge of this region comes from its many charged amino acids, where lysine and glutamic acid residues are the most frequent (Appendix: Amino acid sequence of full-length HSP90). This "charged linker" (marked in Figure 4) has proven to be essential for regulating ATPase activity and conformational flexibility, which impacts the availability of potential binding sites for client proteins [43, 53]. It has yet to be determined if the presence or length of this "charged linker" has any functional effect on its interactions with CTA1.

1.4 Cytolethal distending toxins

One particular class of exotoxins has DNA-damaging effects. These toxins, known as cytolethal distending toxins (CDTs), are genotoxins widely distributed among many Gram-negative bacteria [54]. The first CDT was reported in the late 20th century, where a pathogenic strain of *E. coli* was found to produce a toxin causing cell expansion [55], and similar activities were later identified in strains of *Shigella* and *Campylobacter* [56, 57]. It is now clear that it is an AB toxin associated with a range of diseases [58-60].

1.4.1 Gene products and enzymatic activity

The CDTs are produced from three adjacent or overlapping genes encoding the proteins CdtA, CdtB and CdtC, where the former and latter make up the binding components of the AB_2 toxin. CdtB, the catalytically active component, carries out the enzymatic activity. Using PSI-BLAST, deoxyribonuclease (DNase) I-like homologous residues were first identified in the CdtB subunit of CDT in both *E. coli* and *Campylobacter jejuni* [61, 62]. Several catalytic residues corresponding to mammalian DNase I, as well as critical DNA-binding residues, have later been confirmed to be conserved for all CdtBs [63]. The conservation of these residues was consistent with the later established DNA-damaging role of *Haemophilus ducreyi* CDT (HdCDT) in mammalian cells [64]. CdtB was officially classified as an enzyme belonging to the DNase I family when the crystal structure of HdCDT was solved and refined to 2.0 Å in 2004 [63] presented in Figure 5.

Figure 5 Crystal structure of HdCDT holotoxin. CdtA (dark green) and CdtC (light green) constitute the binding (B) components, while CdtB (red) constitutes the active (A) component of the AB_2 toxin. The structure is presented schematically in the same colors in Figure 6. (PDB code: 1SR4) [63]

As a DNase I-like enzyme, CdtB causes breaks in double-stranded DNA, activating what is called a DNA damage response. The DNA damage is first recognized by sensor proteins, which activates specific protein kinases that amplifies and activates the cellular response of effector proteins such as gene-regulating p53 [65]. This induces an irreversible cell cycle block at a critical mitotic transition state, leaving the affected cells distended (expanded) and incapable of cellular division and proliferation for which CDTs are characterized [66-68]. The combination of DNA degradation and cell cycle blockage have proven to induce cell death (apoptosis) in mammalian cells [69].

1.4.2 From internalization to cell death

Due to the many variants of CDT produced by different bacterial species, their activation pathway is not nearly as straight-forward as that of CT. It is clear that CDT adheres to the target cell and that CdtB is internalized and trafficked to the nucleus to exert its function. The binding components are dissociated from CdtB somewhere along this route. Studies have demonstrated that there is more than one mode of adherence for CDTs [70-73] and that the time and place of toxin dissociation are not uniform among species [74-77]. A few extensively studied CDTs, including HdCDT, share the same dependence on retrograde transportation to the ER [78, 79]. The movement of CdtB into the nucleus is also poorly understood, but the CDT of *Actinobacillus actinomycetemcomitans* has demonstrated nuclear import through a functional domain in its N-terminal segment, indicating that CdtB specifically targets the cell nucleus [80].

1.4.3 How and when does CdtB dissociate from the holotoxin?

The individual roles of CdtA and CdtC in CDT toxicity are poorly understood, and there are contradictory reports of whether or not one or more binding subunits are retained in the membrane during toxin internalization [54, 78, 81]. However, a recent study of two different CDTs, including HdCDT, have found CdtC to be 1) localized in the cytosol of intoxicated cells as well as 2) affecting intracellular trafficking of CdtB [76]. One study also indicated that CdtA likely modifies the other two subunits to form the active CDT complex [74]. If these results represent a common internalization pathway for CDTs, it raises the question of how and when they dissociate from and thereby activate CdtB. Results from the collaborating group of Ken Teter at the University of Central Florida indicate that HdCdtB has an ATP binding site that may be crucial for toxin activation (personal communication). Investigating the structural changes induced in CdtB following ATP binding and locating this ATP binding site could help answer the unanswered question of holotoxin disassembly and activation.

2 Aims of thesis

Toxigenic bacteria such as *V. cholerae* and *H. ducreyi* rapidly replicate and spread in a variety of reservoirs. Except for the use of antibiotics, there are very few ways of battling them, and an observed increase in cases involving drug-resistant bacteria leaves us with a desperate need for alternative solutions. Investigating secreted toxins by bacteria can give valuable insight into their stability and actions within host cells. This thesis focuses on the exotoxins of *V. cholerae* and *H. ducreyi* and their interactions with the host machineries HSP90 and ATP, respectively. By the means of chromatography and interaction experiments, the final goal for the two toxin projects is to produce crystal structures of both toxins with their binding partners.

In vivo, CTA1 relies on HSP90 for refolding, but this has yet to be established *in vitro* due to the instability of CTA1 in higher temperatures and as a free subunit. One of the main goals for the cholera toxin project is to express and purify stable CTA1 and HSP90 protein constructs and induce complex formation under physiological conditions *in vitro*. In this thesis, three possible methods for determining if complex formation has been achieved are described using two constructs of both CTA1 and HSP90. Factors impacting unfolding and aggregation of CTA1 are investigated in the presence and absence of both truncated and full-length HSP90.

With no binding partners previously reported for CdtB, its dissociation from CDT is poorly understood. The collaborating group of K. Teter has now found that HdCdtB binds to ATP, but the binding site was not established. It is possible that binding of ATP induces toxin dissociation, subsequently activating CdtB. One of the main goals for the CDT project is to locate this binding site by the means of protein purification and crystallization. The aim for the thesis was to purify CdtB and find the optimal conditions for crystallization after incubation with ATP.

3 Method-related theory

3.1 Chromatography

3.1.1 Affinity chromatography

In this project, all constructs were initially designed with a tag of six histidine residues (His_6) on either their N- or C-terminal tails. Using an immobilized metal affinity chromatography (IMAC) resin charged with target-specific ions, molecules with repetitive histidines will bind to the resin material.

The two most frequently used IMAC resins available are the nickel-nitrilotriacetic acid (Ni-NTA) resin, coated with nickel ions, and the TALON resin, coated with cobalt ions. Purifying a tagged protein with a TALON resin usually results in less contaminants due to its lower affinity for polyhistidines. For that reason, it is often used if the tag binds weakly to a column. Ni-NTA is the preferred choice of resin for strong tags, as it has a higher affinity for binding histidines. In both resin materials, the protein of interest can be eluted from the column by applying increasing concentrations of imidazole. Before applying a protein extract, an Ni-NTA resin needs to be equilibrated with a binding buffer containing small amounts of imidazole to remove weakly interacting proteins. As a TALON resin is more selective for polyhistidines, the buffer used for resin equilibration should not contain any imidazole. The concentration of imidazole required for eluting a His-tagged protein varies depending on the choice of resin material and the accessibility of the tag. [82]

3.1.1.1 Enzymatic cleavage of affinity tags

Additional amino acids on polypeptides can impact the solubility and proper folding of a protein. Sequence-specific proteases such as Tobacco Etch Virus (TEV) and Thrombin can be used to enzymatically remove affinity tags after purification if the protein sequence contains a recognition site for that enzyme. The main cleavage sites for TEV and Thrombin proteases are ENLYFQ/G and LVPR/GS, respectively, but Thrombin can also perform off-target cleavage at non-specific sites [83].

3.1.2 Size exclusion chromatography

Size exclusion chromatography (SEC) separates proteins based on their molecular size. The stationary phase of a column material is composed of porous beads by which smaller molecules will be slowed down. Larger molecules pass through the material without entering the beads, eluting faster from the column. The column is equilibrated with the sample buffer and the concentrated solution of the target protein is injected using a sample loop. Buffer applied before and after protein injection may or may not be the same, making this an excellent desalting or buffer exchange step if needed. [84, 85]

3.2 Polyacrylamide gel electrophoresis

Polyacrylamide gel electrophoresis (PAGE) is a widely used technique for estimating the molecular weight (Mw) of a purified protein. A protein solution is run through a polyacrylamide gel in conditions that denature the proteins and give them a net negative charge relative to their size. This is obtained by 1) adding a reducing agent such as dithiothreitol (DTT) to reduce disulfide bridges, 2) adding detergents such as sodium or lithium dodecyl sulfate (SDS/LDS) to disrupt non-covalent bonds important for higher levels of structure, and 3) heating the sample at 70°C for 10 minutes to ensure complete denaturation of the protein. Negatively charged phosphate groups of SDS or LDS are exposed on the primary polypeptide chain, giving it a net negative charge in relation to their size. Applying a pre-stained protein standard to the gel for comparison makes it possible to estimate the Mw of all proteins in a sample from the relative migration of the protein standard. [86]

3.3 Protein crystallization and crystal screening

Crystallization is the process of solidifying a molecule into a highly organized crystal. As proteins rarely crystallize in nature, this can be artificially induced by precipitating them slowly from a super-saturated solution under controlled conditions. Individual protein molecules will pack together in a repeating array, forming a well-ordered crystal that can be analyzed using X-ray crystallography. A beam of X-rays is shot at the crystal, which bends the rays slightly depending on the arranged crystal pattern. This gives a diffraction pattern that can produce an image of how its atoms are arranged, and the three-dimensional structure of the protein can be determined. The conditions in which a protein crystallizes are rarely predictable and depend on its physiochemical characteristics.

Empirical testing of a variety of conditions must be performed to determine the appropriate crystallization condition is found. The vapor diffusion technique is a popular method of protein crystallization. A drop composed of protein solution (5-40 mg/mL) and reagent are set up in a sealed system either hanging or sitting next to a reservoir solution with reagent. Water evaporates from the drop to the reservoir, increasing protein and reagent concentration in the drop, eventually reaching an equilibrium between the drop and the reservoir. Commercial crystal screens are usually applied to evaluate the variables that affect protein crystallization. If there is no data indicating that a certain condition should be tested, sparse matrix screens composed of reagents that have previously proven to crystallize a protein are chosen as primary screens. One such screen is the JCSG-*plus*, which screens for a broad range of buffers and precipitants in a total of 96 conditions. It is highly effective when accompanied by the systematic grid screen PACT *premier*, which screens for the effects of pH/polyethylene glycol (PEG), anion/PEG and cation/PEG in solution.

4 Methods

4.1 Expression and purification of protein constructs

Chemicals, solutions and equipment referred to in-text can be found in the Appendix. Before use, all media and solutions were checked for visible contamination and bacterial growth. All media were autoclaved. All solutions were filtered and degassed, except for those used for dialysis (only filtered). Affinity columns were equilibrated and washed with a minimum of five column volumes of buffer.

4.1.1 TEV protease

BL21 (DE3) *E. coli* cells transformed with a plasmid encoding TEV-His_6 with green fluorescent protein (GFP) were stored in 50% (v/v) glycerol at -80 °C. The plasmid carries a gene for ampicillin resistance (Amp), and so 1:1000 Amp was added to all media before inoculation.

100 mL lysogeny broth (LB) medium were inoculated with the transformed cells to grow a preculture overnight (ON) at 30 °C (120 rpm) for 20 hours. For each main culture, 1:100 preculture was added and incubated at 37 °C (140 rpm) until the optical density measured at 600 nm (OD_{600}) was 0.78. At this point, 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added for induction of protein expression. The cultures were incubated ON at 20 °C (140 rpm) for 20 hours. Cells were harvested by centrifugation (7000 rpm) for 15 minutes at 4 °C. The pellet was stored at -20 °C.

The pellet was thawed and resuspended on ice at room temperature (RT) in 150 mL His AC buffer, flakes of DNase I, 1 M $MgCl_2$, 1 M $CaCl_2$ and 1:100 phenylmethanesulfonyl fluoride (PMSF). The mixture was lysed with a BioSpec BeadBeater (5x30 seconds, 1-minute breaks) and the protein extract was collected after centrifugation (17 000 rpm) and filtered with 0.45 μm syringe filters. A 5 mL Ni-NTA affinity column was equilibrated with His AC buffer before loading the filtered protein extract. After protein loading, the column was washed with His WB buffer. The protein was eluted with His EC buffer. A total of 20 mL of 7 mg/mL GFP-TEV-His_6 was diluted to a final concentration of 3.3 mg/mL in His S buffer and stored at -20 °C in 1.5 mL aliquots.

4.1.2 CTA1

BL21 (DE3) *E. coli* cells transformed with plasmid pET-22b+ encoding CTA1-His_6 or CTA1 were stored in 50% (v/v) glycerol at -80 °C. The plasmid carries a gene for Amp resistance, and so 1:1000 Amp was added to all media before inoculation. 1:1000 chloramphenicol (Cam) was added in addition to Amp when using BL21 (DE3) pLysE cells.

40 mL LB medium were inoculated with the transformed cells to grow a preculture ON at 30 °C (120 rpm) for 16-20 hours. For each main culture, 1:100 preculture was added and incubated at 37 °C (130 rpm) until OD_{600} was between 0.4 and 0.6. At this point, 1 mM IPTG was added for induction of protein expression. The cultures were incubated ON at 20 °C (130 rpm) for 20-22 hours. Cells were harvested by centrifugation for 20 minutes at 4 °C (6000 rpm) and pellets were stored at -80 °C.

Each pellet was thawed and resuspended on ice at RT in 40 mL His A.1 buffer, 2 µL benzonase, 1 M $MgCl_2$, 1 M $CaCl_2$, 1:100 PMSF, 1:10 Triton X-100 and 1:100 lysozyme. The mixture was centrifuged (12 000 rpm) for 30 minutes at 4 °C. The 1st pellet was resuspended in 40 mL His A.2 with 0.5 % (v/v) Triton X and centrifuged (10 000 rpm) for 15 minutes at 4 °C. The 2nd pellet was resuspended in pure His A.2 and centrifuged (10 000 rpm) for 15 minutes at 4 °C. The 3rd pellet was resuspended in His A.3 and incubated at RT for 1 hour before centrifugation (20 000 rpm) for 35 minutes at 4 °C. The supernatant was collected and filtered with 0.45 µm syringe filters. A 5 mL Ni-NTA affinity column was used for purification of the His-tagged construct. The column was equilibrated with buffer His A.3 before applying the filtered protein extract on the column. The column was then washed with buffer His A.3 and the protein was eluted using increasing concentrations of His B.1 buffer.

The protein solution was filtered using 0.45 µm syringe filters and dialyzed against dialysis buffers with decreasing concentrations of urea (6M to 0M), all performed at 4 °C with stirring. First against D1 for 2 hours, then ON against D2. The next day, the solution was dialyzed against D3 to D5 for 2 hours each time. The protein was concentrated using 50K MWCO Ultra-15-Centrifugal Filter Device. The concentrated protein was either used immediately or snap-frozen with 10% (v/v) glycerol and stored at -80 °C in 200-400 µL aliquots.

4.1.3 HSP90

BL21 (C43) *E. coli* cells transformed with plasmid pETM-11 encoding HSP90*-His_6 or HSP90-His_6 were stored in 50% (v/v) glycerol at -80 °C. The plasmid carries a gene for kanamycin (Kan) resistance, and so 1:2000 Kan was added to all media before inoculation.

150 mL LB medium were inoculated with the transformed cells to grow a preculture overnight (ON) at 30 °C (110 rpm) for 16-20 hours. For each main culture, 1:100 preculture was added and incubated at 37 °C (120 rpm) until OD_{600} was between 0.6 and 1. At this point, 0.5 mM IPTG was added for induction of protein expression. The cultures were incubated ON at 20 °C (110 rpm) for 18-22 hours. Cells were harvested by centrifugation (6000 rpm) for 20 minutes at 4 °C and pellets were stored at -80 °C.

Each pellet was thawed and resuspended on ice at RT in 150 mL His A buffer, 3 µL benzonase, 1 M $MgCl_2$, 1 M $CaCl_2$ and 1:100 PMSF. After a two-hour incubation, the mixture was lysed with a BioSpec BeadBeater (5x30 seconds, 1-minute breaks) and the protein extract was collected after centrifugation (17 000 rpm) for 2x30 minutes at 4 °C and filtered with 0.45 µm syringe filters. A 5 mL Ni-NTA affinity column was equilibrated with buffer His A before loading the protein extract. The column was then washed with buffer His A and the protein was eluted using buffer His B.

GFP-TEV-His_6 was thawed in hand and kept on ice. The protein solution was filtered with 0.45 µm syringe filters before adding 1:25 GFP-TEV-His_6. The protein mixture was dialyzed against TEV-D buffer ON at 4 °C with stirring using 10K MWCO SnakeSkin dialysis tubing. Another round of affinity purification was performed, leaving untagged HSP90* in the flow-through. The protein was dialyzed against HSP-D buffer ON at 4 °C with stirring using 10 MWCO SnakeSkin dialysis tubing. If SDS-PAGE analysis indicated strong impurities or oligomerization of HSP90*, SEC was performed using a Superdex 200 10/30 GL (SD200) column equilibrated and washed with His C buffer. The protein was concentrated using 50K MWCO Ultra-15-Centrifugal Filter Device. The concentrated protein was snap-frozen and stored at -80 °C in 200-400 µL aliquots.

4.1.4 CdtB

BL21 (C43) *E. coli* cells transformed with plasmid pET-28b+ encoding CdtB-His_6 were stored in 50% (v/v) glycerol at -80 °C. The plasmid carries a gene for Kan resistance, and so 1:1000 Kan was added to all media before inoculation.

40 mL LB medium were inoculated with the transformed cells to grow a preculture ON at 30 °C (120 rpm) for 16-20 hours. For each main culture, 1:100 preculture was added and incubated at 37 °C (120 rpm) until OD_{600} was between 0.6 and 0.8. At this point, 1 mM IPTG was added for induction of protein expression. The cultures were incubated ON at 20 °C (110 rpm) for 18 hours. Cells were harvested by centrifugation (6000 rpm) for 20 minutes at 4 °C and pellets were stored at -80 °C. Resuspension of pellet, purification of inclusion bodies, extraction of protein, NiNTA purification, and dialysis steps were performed as described for CTA1 constructs in sections **4.1.3**.

For enzymatic cleavage of the His-tag, the protein solution was dialyzed against Thrombin dialysis buffer. 1 mg/mL of protein solution was mixed with 15 µL of resuspended Thrombin Sepharose Beads per 1 mL of volume. Mixture was incubated at 10 °C ON in Thrombin buffer with stirring for optimal cleavage reaction and then gently centrifuged (3000 rpm) for 5 at minutes at 10 °C to separate beads from cleavage product. Beads were washed with Thrombin buffer and spun down (3000 rpm) three times to maximize recovery of target protein. Protein was dialyzed against Talon A buffer ON at 10 °C. Cleaved and uncleaved protein were separated using a 5 mL TALON affinity column where the flow-through (untagged protein) was collected and concentrated using Ultra-15-Centrifugal Filter Device with a 10K MW cut-off. Uncleaved protein was eluted with Talon B buffer. Samples were kept at 10 °C for short-term storage.

4.2 SEC analysis

An SD200 column was equilibrated with two column volumes of the appropriate buffer. In the case of HSP90 constructs, it was equilibrated with 20 mM Tris-HCl and 150 mM NaCl (pH 7.5). In the case of CdtB constructs, it was equilibrated with 20 mM NaH_2PO_4 and 150 mM NaCl (pH 8.0). 250 µL of concentrated protein (5-10 mg/mL) was applied to the column using a 500 µL sample loop and the column was washed with one column volume of the same buffer as it was equilibrated with. All steps were performed at 4 °C. Fractions corresponding to the largest protein peaks were collected and analyzed using SDS-PAGE (section **4.3**).

4.3 SDS-PAGE analysis

A sample reducing buffer (SRB) with a 10:1 ratio of Bolt LDS Sample Buffer (4X) and Bolt Sample Reducing agent (10X) was made prior to sample preparation. 15 µL of protein sample were mixed with 5 µL of SRB and heated at 70 °C for 10 minutes. Electrophoresis chamber was filled with 1X Bolt MES SDS Running Buffer. Samples were loaded into pre-casted Bolt 4-12 % Bis-Tris Plus gels. 5 µL of SeeBlue® Plus2 Pre-stained Standard were loaded as a molecular weight reference. Electrophoresis was carried out at RT with a constant voltage of 150-200V until the dye front reached end of gel (25-40 minutes). Gels were stained with Coomassie Blue and de-stained with acetic acid and ethanol. All steps were performed at RT.

4.4 CTA1/HSP90 complex formation

4.4.1 Incubation tests under physiological conditions

Purified CTA1-His_6 was dialyzed against HSP-D buffer at 4°C with stirring overnight. Stored HSP90* or HSP90-His_6 samples were thawed in hand and kept on ice during sample preparation. CTA1-His_6 was diluted to 1 mg/mL and kept on ice. HSP90* or HSP90*-His_6 samples were mixed with CTA1-His_6 in a 3:1 ratio with 1 mM ATP and incubated at 37°C for a total of 3 hours. 50 µL samples were taken after every hour and frozen at -20°C. All samples were analyzed on an SDS-PAGE gel.

4.4.2 TALON pull-down of CTA1 with HSP90

Purified CTA1 was dialyzed against Talon A buffer at 4°C with stirring overnight. Stored HSP90*-His_6 or HSP90-His_6 samples were thawed in hand and dialyzed against Talon A buffer at RT. CTA1 was diluted to 0.5 mg/mL and kept on ice. HSP90*-His_6 or HSP90-His_6 samples were mixed with CTA1 in a 3:1 ratio with 2 mM ATP and incubated at 37 °C for a total of 1 hour. A 1 mL TALON column was equilibrated at RT with Talon A buffer and the incubated protein mixture was loaded onto the column. The column was washed with increasing concentrations of Talon B buffer. 15 µL samples were taken from the incubated protein mixture, column flow-through and elutions from each wash step. All samples were analyzed on an SDS-PAGE gel.

4.5 ATP-binding in CdtB

4.5.1 Incubation with ATP

Purified CdtB was dialyzed against 20 mM Tris-HCl and 200 mM NaCl (pH 7.5) at 10°C with stirring overnight and concentrated to 5-8 mg/mL using Ultra-15-Centrifugal Filter Device with a 10K MW cut-off. Concentrated protein was mixed with 1 mM ATP and incubated at RT for 1 hour. Samples were check for visible precipitation.

4.5.2 Crystal screening of CdtB-ATP

For sitting-drop vapor diffusion, JCSG *plus*, PACT premier and MORPHEUS screens were applied. All crystal screens were spun down (3000 rpm) for a few minutes at 4 °C prior to use. For each screen, 80 µL of the different screening solutions were added to their respective wells (A1-H12) in a 96-wells Swissci 2 Lens Crystallization plate. Two drops per condition with 1:1 and 1:4 protein solution and reservoir solution were distributed using a protein crystallization robot (Oryx4) with an evaporation shield. Plates were placed in cooled crystal hotel for storage and imaging. For hanging-drop vapor diffusion, reservoir solutions described in Table 5, **section 5.5.3**, were distributed in their respective wells in a 1 mL volume. Drops of 1:1 protein solution and reservoir solution were set by hand on siliconized cover slides and sealed with grease. Drops were set at RT and stored in a cooled room with minimal exposure to light.

5 Results and discussion

Performed experiments and proposed strategies are in accordance with the ethical guidelines that apply at the University of Oslo and do not involve the use of laboratory animals at this facility. Written consent to publish results confirmed through personal communication was obtained by collaborating partners prior to publication. All mentioned sequences and protein parameters can be found in the Appendix.

5.1 Initial status of project

Due to the unstable nature of CTA1, its interacting residues with HSP90 have yet to be determined. At the start of this project, crystal structures of both proteins had already been solved and purification protocols for His-tagged constructs of CTA1 and HSP90 had been established. CTA1 protocol was established by Teter group, UCF, and HSP90 protocol was established by Joël B. Heim, University of Oslo. The first HSP90 construct had a truncated charged linker and a deleted C-terminal tail, as well as a TEV cleavage site for His_6 removal after purification. It was designed by Joël B. Heim and ordered through GenScript. As it is a very flexible protein full-length, these measures were meant to reduce its mobility in solution to increase the chances of crystallization. The tagged CTA1 construct was not designed with a protease cleavage site for His tag removal, leaving it more flexible than the wildtype toxin. As the project progressed, an untagged CTA1 constructs were designed and produced, and a full-length His-tagged HSP90 was designed by Joël B. Heim and also GenScript.

The crystal structure of HdCdtB (hereafter CdtB) was solved in 2004, but ATP-dependence in this toxin has never been reported. Since we know CdtB can be crystallized, this method of structure determination can also be applied to its ATP-bound form. However, ATP-binding might impact the overall structure and solubility of CdtB, leaving little basis for the new conditions of crystallization. Screening for these conditions with a final goal of crystallization is the first step to answering one of the many unanswered questions regarding CdtB activation. The His-tagged CdtB construct proven to bind ATP by the collaborating group of K. Teter was also used in this project. CdtB purification protocol had been established by Teter group, but adjustments were made to increase protein stability. It also had a Thrombin cleavage site for removal of His_6, but the protocol for this was not established prior to project start and proved to be challenging for this construct.

5.2 CTA1

5.2.1 Production and stability of His-tagged CTA1

Unless inactivated, CTA1 gene products are toxic, and their expression can interfere with the physiology and growth rate of *E. coli* host cells. At the start of this project, the CTA1-His_6 construct had already been successfully transformed into BL21(DE3) pLysE cells. This strain was chosen due to its high expression of T7 lysozyme from its pLysE plasmids, which allowed for reduced activity of the T7 RNA polymerase to control the transcription of toxic CTA1. Despite the reduced toxic activity, these cells grew slowly, and protein expression was for that reason induced at an OD_{600} between 0.2 and 0.4. This OD_{600} was considered a good bacterial density for protein expression in pLysE cells, as they sometimes never even reached this value.

CTA1 readily unfolds in higher concentrations and at temperatures higher than 4 °C, forming aggregates, and the CTA1-His_6 construct was expected to behave in the same way regardless of an extended C-terminal domain. Aggregates of unfolded protein, inclusion bodies, were in aqueous solutions and could not be extracted from the cell lysate in the same way as soluble proteins. Following the established protocol, a purely biochemical rather than mechanical cell lysis was performed with lysozyme and detergent in addition to PMSF and nucleases. Components of the bacterial cell wall and membrane were disrupted by these substances, allowing the inclusion bodies to seep out without being exposed to mechanical stress.

The purification protocol was from this point on quite extensive, as the inclusion bodies obtained after centrifugation were sedimented together with cellular debris and possibly intact cells. Once the inclusion bodies had been properly washed, they were dissolved in high concentration of denaturing urea to re-solubilize the aggregated proteins. It was then very important to treat this extract with care, maintaining a constant temperature of no more than 4 °C due to the unstable nature of CTA1. Even under denaturing conditions, the His-tagged construct easily precipitated when left at RT by mistake, indicating that the tag did not increase the solubility of unfolded CTA1, at least not in any significant capacity.

Few impurities were expected, as soluble cytosolic proteins had been removed already during the extensive preparation of the protein extract. The following Ni-NTA purification would ideally be performed with an imidazole gradient using an ÄKTA purifier, but the 8 M urea buffers proved corrosive to this equipment, which is why the purifications were performed manually with a peristaltic pump. The expectation of few impurities was confirmed after loading the protein extract on the column. Figure 6, showing an SDS-PAGE gel of the Ni-NTA purification, shows that the only over-expressed protein in the protein extract (well 1) has a size of approximately 22 kDa, corresponding to the size of CTA1-His_6.

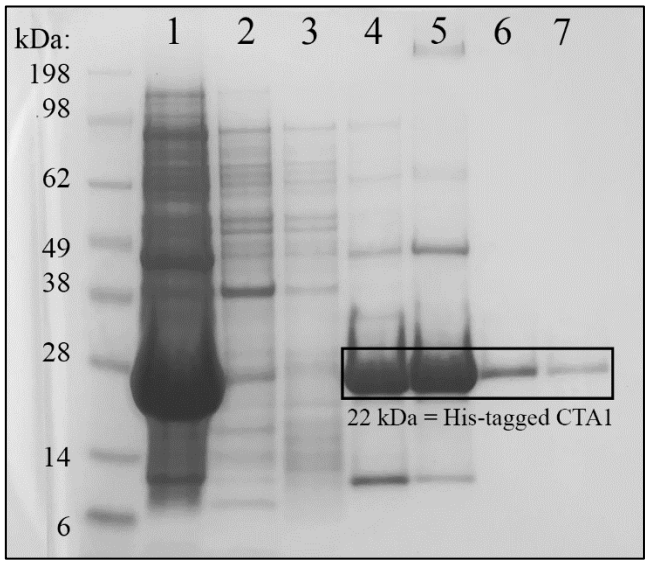

Figure 6 SDS-PAGE analysis of fractions eluted from CTA1-His_6 Ni-NTA purification. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Wells marked 1-5 were loaded with 1) protein extract, 2) column flow-through, 3) 20 mM imidazole wash, 4) 100 mM imidazole wash, 5) 150 mM imidazole wash, 6) 200 mM imidazole wash and 7) 2nd 200 mM imidazole wash.

It is somewhat difficult to compare to the protein standard due to the heavily over-loaded sample, but it is clear in well 2 that almost no target protein was left in the flow-through of the column. A thorough 20 mM imidazole wash (well 3) removed almost all weakly bound impurities from the sample and target protein elution was successful with a 100-150 mM imidazole concentration. This supports that 1) CTA1- His_6 can be overexpressed in sub-optimal bacterial densities of BL21 (DE3) pLysE cells and that 2) the affinity tag binds to Ni-NTA resin material, effectively separating CTA1-His_6 from the few impurities present.

5.2.2 Optimization of refolding steps

After Ni-NTA elution, CTA1-His_6 is left unfolded in highly denaturing 8 M urea. Its concentration before refolding varied from 1 to 3 mg/mL in 25 mL elution. Dialyzing the protein against buffers with gradually decreasing urea concentrations, in combination with low temperatures at which these steps should be performed, gave CTA1 time to refold properly so as to achieve a minimal level of unwanted precipitation or aggregation. However, the refolding protocol of CTA1-His_6 was refined many times during this project due to reoccurring visible precipitation. After re-measuring protein concentrations, it was clear that up to 50 % of protein was lost to misfolding. The step at which this was the most evident was unsurprisingly when decreasing the urea concentration from 2 M to 0 M, where the protein is supposed to refold completely. Misfolded proteins will at this point clump together, as well as react to the removal of all urea salt from solution. To avoid loss of protein, additional time was first added to each step of dialysis. As this did not change the degree of precipitation, an additional dialysis step with only 1 M urea was introduced, but also with no visible change. Finally, a variety of protein concentrations were tested for dialysis, and it was confirmed that keeping the concentration to less than 0.5 mg/mL significantly reduced visible precipitation of CTA1-His_6.

5.2.3 Removal of C-terminal His_6

As already indicated, the extended C-terminal His tag in CTA1-His_6 did not significantly increase its solubility, but this did not mean that it had no impact on this. It is possible that the tag decreases the solubility, as well as affecting binding to other proteins, which could prove to be a factor in its binding to HSP90 *in vitro*. Given the high purity of CTA1-His_6 in inclusion bodies, it seemed fully possible to purify CTA1 without a tag, but the construct contained no recognition sites for enzymatic cleavage. Primers presented in the Appendix were designed to introduce a stop codon immediately prior to the six histidines of CTA1-His_6 C-terminal tail, and the construct was sequenced through GATC Biotech (now Eurofins Genomics). Sequencing confirmed that the tag had been successfully removed from CTA1-His_6, creating a new construct (hereafter CTA1). Although the absence of a tag could prove otherwise favorable, this construct was designed and produced mainly for the intention of testing a novel method of determining if complex formation had occurred between CTA1 and HSP90. This method is described and discussed in section **5.4.2**.

5.2.4 Production and stability of untagged CTA1

The new CTA1 construct, also transformed into BL21 (DE3) pLysE cells, was expressed and produced by copying the CTA1-His_6 protocol except for the Ni-NTA purification step. CTA1 no longer had a His tag, but the method was still employed to remove some of its impurities. An affinity chromatography step was also necessary to ensure that the protein indeed lacked the removed affinity tag. The target protein would, in this case, flow right through the column material, while weakly interacting proteins were left in the resin material. As an Ni-NTA resin has a low specificity for polyhistidines, possibly binding CTA1 despite the absence of a His tag, a TALON resin was chosen instead for this purification step. To clearly show the successful removal of His_6 and purification of CTA1, CTA1-His_6 was produced simultaneously. Figure 7 shows an SDS-PAGE analysis of the TALON purification of CTA1-His_6 (well 1-4) and the Ni-NTA purification of CTA1 (well 5-7). The protein extracts (well 1 and 5) both show one strong band of approximately 22 kDa, indicating a successful induction of protein expression for both constructs. CTA1-His_6 has a size of 22.3 kDa, while CTA1 is closer to 21.5 kDa, but this small size difference of almost 1 kDa can be very difficult to visualize on a gel. After loading these extracts onto their respective columns, it was evident that CTA1-His_6 bound to the Ni-NTA column (well 2), while most of CTA1 passed through the TALON column (well 6). The small difference in size were, therefore, not relevant to determine if the construct had lost its His tag.

Impurities observed in the 200 mM elution of CTA1-His6 (well 4) are very similar to that of the CTA1 flow-through. One stronger band of approximately 40 kDa is present in the CTA1 fraction, while almost completely separated from CTA1-His6 with a 20 mM and 50 mM imidazole wash. This might be the only advantage of this construct regarding the level of purity obtained. Although some CTA1 was retained by the resin material and could possibly have been collected with very little imidazole, it was instead eluted with a high imidazole concentration (well 7) and discarded. This was a conscious choice, as it was important that the untagged CTA1 did not bind at all after incubation with HSP90 in the experiment discussed in section **5.4.2**.

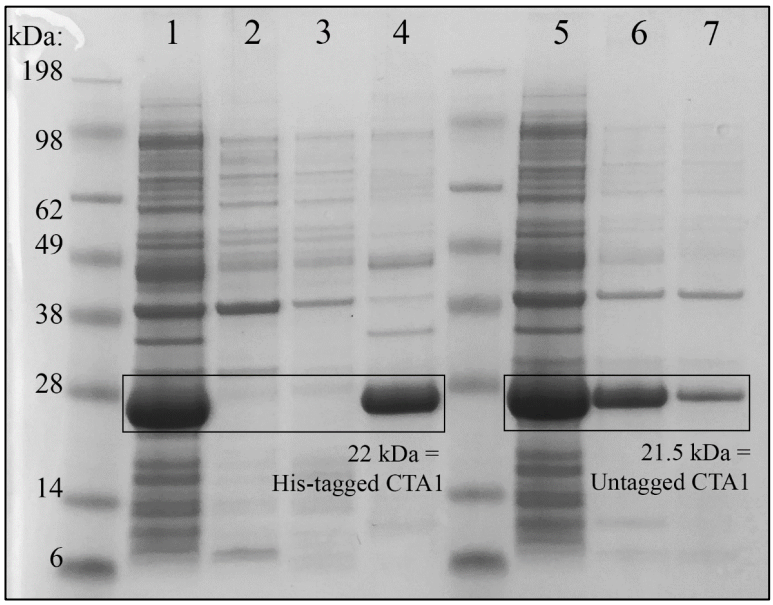

Figure 7 Comparison of SDS-PAGE analysis of fractions eluted from CTA1-His_6 Ni-NTA purification (left) and CTA1 TALON purification (right). SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Well 1-4 (CTA1-His_6) were loaded with 1) protein extract, 2) column flow-through, 3) 20 mM imidazole wash, 4) 200 mM imidazole wash. Well 5-7 (CTA1) were loaded with 5) protein extract, 6) column flow-through and 7) 200 mM imidazole wash.

The yield measured prior to CTA1 refolding was significantly higher than that of CTA1-His_6. Where CTA1-His_6 rarely exceeded 3 mg/mL in a 25 mL elution, CTA1 was close to 4 mg/mL in the same fraction volume. This is not clear in the SDS-PAGE analysis, as the concentrations were diluted to clearly see the size of the construct. The second difference from CTA1-His_6 was its capacity to refold properly at concentrations up to 1 mg/mL. This construct proved more soluble and stable, but since CTA1-His_6 refolding had already been optimized, the difference in yield before and after refolding was minimal.

5.3 HSP90

5.3.1 Production of truncated His-tagged HSP90

The flexible nature of HSP90 in solution, as well as its ATPase activity and large size, is problematic for CTA1-binding *in vitro*. A truncated and less flexible HSP90 construct with an N-terminal His tag and a TEV recognition site (denoted HSP90*-His_6) was designed by Joël B. Heim to overcome these obstacles and ordered through GenScript. The charged linker-region proven to be is essential for its flexibility was reduced by 45 amino acids and its C-terminal tail deleted by 35 amino acids. As a result, the new HSP90*-His_6 construct was reduced in both movement and size compared to full-length HSP90 without affecting its ATPase activity. The construct was designed from a cytosolic HSP90 isoform (AB).

HSP90*-His_6 had been successfully transformed into an expression strain and purified prior to the experiments performed in this thesis, and the established purification protocol by Joël B. Heim had proven to give a high yield. The BL21 (C43) cells grew very fast at 37 °C, which was problematic when trying to induce protein expression at an OD_{600} of 0.8. More often than not, induction occurred when the OD_{600} measured between 0.6 and 1.0, but this did not seem to affect the level of target protein over-expression. However, higher bacterial densities also produced more off-target protein products, illustrated by SDS-PAGE analysis of undiluted protein extract later introduced in Figure 10. The added endonuclease benzonase proved ineffective during cell lysis and separation of protein extract. The lysate often remained viscous even after spinning down the debris, which was most likely due to an excess of DNA in solution. This could also be a result of the high bacterial density. As enzymes do not function optimally at lower temperatures, the time of incubation with nucleases was increased from 30 minutes to 2 hours, which finally reduced the lysate viscosity and made it possible to filter the solution prior to affinity purification.

Initially, the protein extract was loaded manually on the column using a peristaltic pump and then connected to an ÄKTA purifier for column wash and protein elution. This was necessary at first to find the optimal imidazole concentration for elution of target-protein, but it was not a continued necessity after this was already established. A chromatogram of a typical HSP90*-His_6 elution from an Ni-NTA column is presented in Figure 8. The column was extensively washed with 20 mM imidazole before a gradient of 50 to 200 mM imidazole elution was applied, where marked elution peaks (1 and 2) always presented themselves in some degree and target protein elution occurred at an imidazole elution around 140 mM (peak 3).

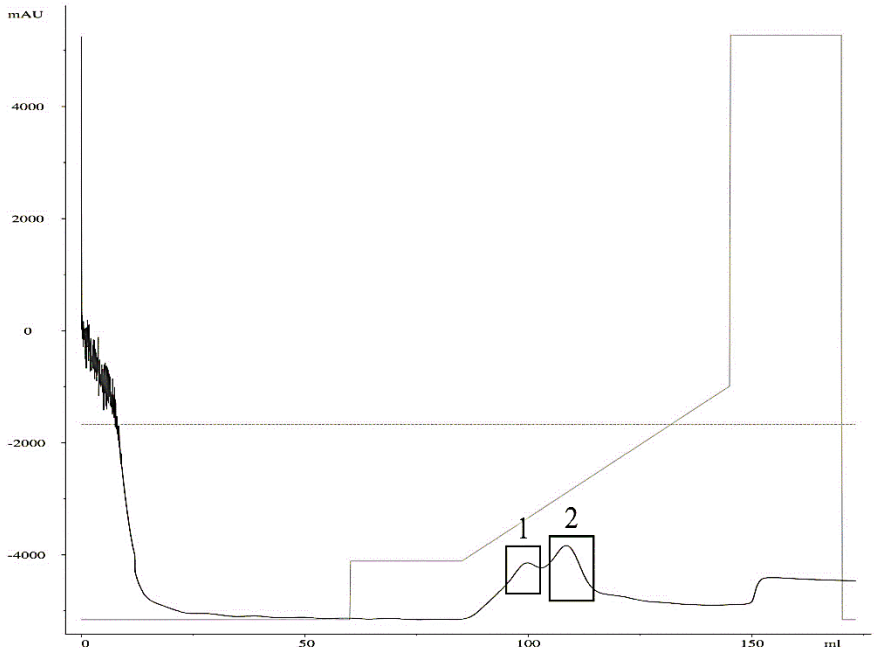

Figure 8 Chromatogram of HSP90*-His_6 Ni-NTA purification with an ÄKTA purifier. The marked peaks were reoccurring in all HSP90*-His_6 chromatograms. Axes y and x are measurements of A_{280} and mL, respectively. Slope represents gradient of increasing percentage of 500 mM imidazole applied.

Peak 1 was impossible to separate from peak 2 due to their similar affinity to the resin material, which was a concern in this purification step. Only proteins with an introduced His tag were expected to bind this strongly to the column. Although not possible to separate by Ni-NTA purification, the fractions from the two peaks were analyzed on an SDS-PAGE gel (Figure 9) in the hope of possibly being able to separate the two proteins by size exclusion. Due to its truncated regions, HSP90*-His_6 was expected to be reduced to a size of 78 kDa.

SDS-PAGE analysis showed that the two proteins share a similar size, which excluded the possibility of separation by size exclusion. It was indeed one protein with the appropriate size in the two peaks, but also what appeared to be a cleavage product of the target protein. Most of this was in the fraction sample from peak 1, but some was also present in the fraction sample from peak 2. The protein of similar size and binding capacity as HSP90*His_6 was referred to as “the weird peak” for a long time, as it was reoccurring in the Ni-NTA chromatograms. If it was a cleavage product, it had to be C-terminal, as the His tag was obviously still present and functional. The protein band from an SDS-PAGE gel was analyzed by mass spectrometry (MS) and confirmed as a C-terminal cleavage product by the group of Bernd Thiede at the Department of Biosciences, University of Oslo. This cleavage could be a result of its already cleaved C-terminal tail, but this was yet to be determined.

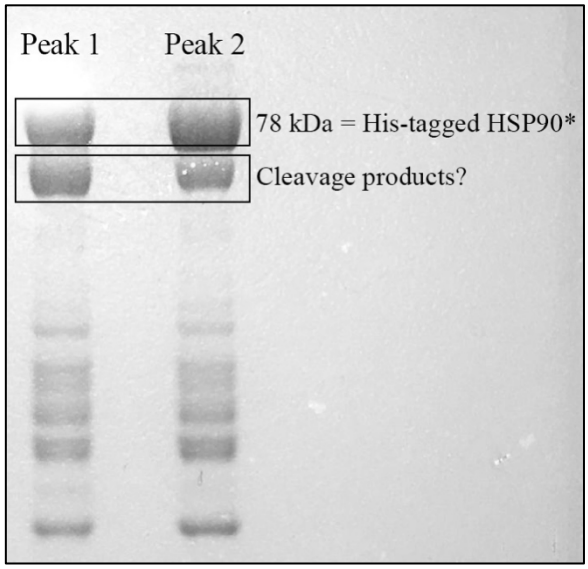

Figure 9 SDS-PAGE analysis of fractions from peaks of chromatogram in Figure 8.

The predictable imidazole elution of HSP90*-His_6 was performed manually with a peristaltic pump when the ÄKTA purifier was not available, which proved equally effective. Figure 11 presents an SDS-PAGE analysis of fractions collected from a manual Ni-NTA purification. It is also an excellent example of how an expression test of this construct (well 1) is often misleading, showing no indication of induced expression when, in fact, there was. Thick and smudged bands for HSP90*-His_6 protein extracts and column flow-throughs were not uncommon, exemplified in Figure 10 (well 2 and 3). This was a continued trend that mirrored the rapid cell growth of the expression strain, but it did not affect the overall yield or purity of the target protein.

Figure 10 shows that many of these impurities could be removed with a 20 and 50 mM imidazole wash (well 4 and 5). As expected, a protein of approximately 78 kDa elutes with an imidazole concentration of 150 mM (well 6) and it also appears to be free of the expected C-terminal cleavage product, which possibly eluted in the fraction of well 5. HSP90*-His_6 was successfully over-expressed and purified both manually and with an ÄKTA purifier. Using a NanoPhotometer (theoretical absorbance at 280 nm (A_{280}) = 1.280), the target protein concentration wash measured between 0.5 to 1.5 mg/mL in 25 mL of volume, giving an average of 25 mg per 9 L of culture (2.78 mg/L).

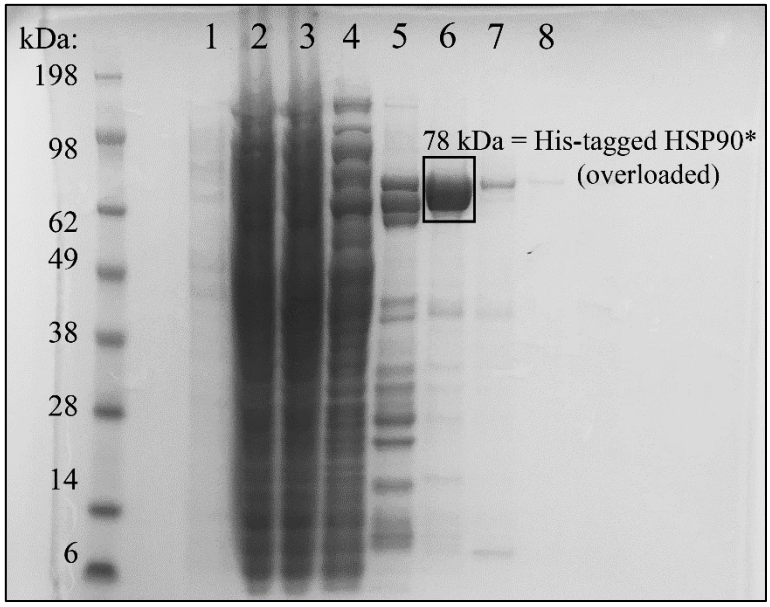

Figure 10 SDS-PAGE analysis of HSP90*-His_6 expression test and manual Ni-NTA purification. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Well 1 was loaded with culture diluted 1:10 after a 24-hour induction of protein expression. Wells marked 2-8 (Ni-NTA purification) were loaded with 2) protein extract, 3) column flow-through, 4) 20 mM imidazole wash, 5) 50 mM imidazole wash, 6) 150 mM imidazole wash, 7) 200 mM imidazole wash and 8) 500 mM imidazole wash.

5.3.2 TEV digest for His tag removal

HSP90*-His_6 had a TEV recognition site downstream of its N-terminal His tag. Proteolytic cleavage of this tag was preferable, as the construct did not need any additional flexibility. Small cleavage tests (undocumented) with thawed HSP90-His_6 samples were performed by Joël B. Heim to ensure that the TEV protease produced in the lab had an acceptable enzymatic activity. It was also important to confirm that the tag was exposed and cleavable in the construct. Production of this protease is discussed in section **5.6**.

The recommended ratio of commercial TEV protease is 1 mg per 100 mg of target protein. Since the protease was not commercial and the production yield in the lab was very high, the HSP90*-His_6 elution from Figure 10 was subjected to GFP-TEV-His_6 (referred to as TEV protease in this section) in a 25:1 mg ratio to be certain all target protein was cleaved. Following the established protocol, an Ni-NTA resin material was used for separation of tagged and cleaved product. SDS-PAGE analysis of the digestion (Figure 11) showed successful cleavage after digestion.

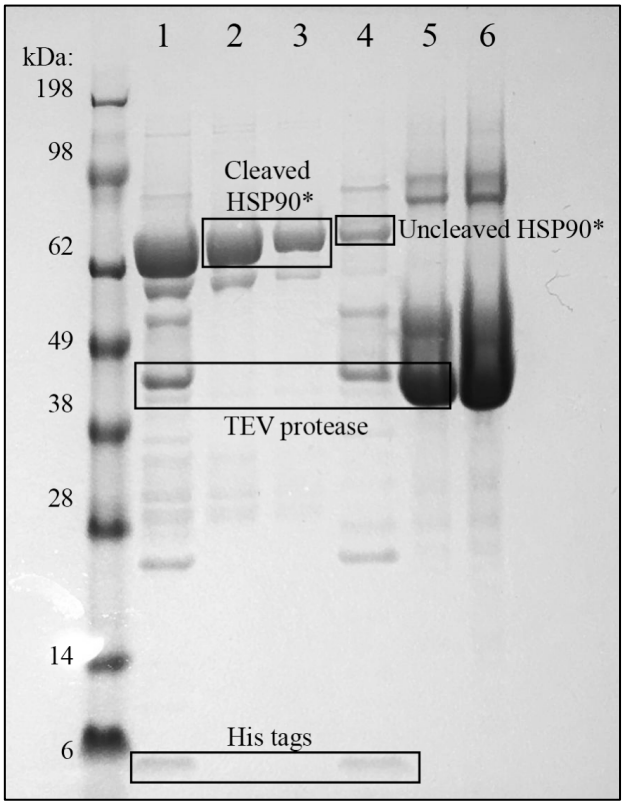

Figure 11 SDS-PAGE analysis of HSP*-His_6 digestion with TEV protease before and after Ni-NTA separation. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Wells marked 1-6 were loaded with 1) digest product, 2) column flow-through, 3) 20 mM imidazole wash, 4) 500 mM imidazole wash, 5) 1 µg TEV protease and 6) 5 µg TEV protease.

The cleavage product (well 1) showed that the target protein and the protease were both present in the sample with no additional products after digestion, only showing strong bands of sizes corresponding to HSP90*-His_6 and TEV. This was important to confirm, as off-target cleaving can sometimes occur, but the high sequence-specificity of TEV ensured that this did not happen. This well also had a small band that could possibly be cleaved His tags, as well as the reoccurring C-terminal cleavage product. Cleaved HSP90* was not retained by the resin material, but rather present in the flow-through (2) and the 20 mM imidazole wash (3). Cleavage of His_6 from HSP90* was successful in both these fractions, as His-tagged HSP90* would not have eluted from the Ni-NTA resin with such a low imidazole concentration. A 500 mM imidazole wash (well 4) released uncleaved HSP90*-His_6 from the column and His-tagged TEV protease. The 500 mM imidazole wash also eluted the same small peptide observed in the digestion sample assumed to be the cleaved His tags.

A different sample of HSP90* was incubated with 1 mM ATP at 37 °C for 1 hour and analyzed with SEC to ensure that the protein did not aggregate or oligomerize. Figure 12 presents the resulting chromatogram with a single peak of approximately 2500 mAU at A_{280}, indicating that HSP90* can be subjected to physiological temperatures without oligomerizing. Two peaks would have shown if some protein had formed oligomers, representing two elutions of different-sized proteins, while a completely oligomerized sample would elute much earlier due to its large size. The high peak observed here (A_{280}) does not represent a large protein, but rather a high concentration of a protein of one size. Absorbance of ATP at 260 nm (A_{260}) could have affected the size of this peak, but HSP90* was still proven to survive both the 37 °C-incubation and SEC without oligomerizing. However, SEC was not an appropriate indication of how much HSP90 is left in a sample after incubation with ATP.

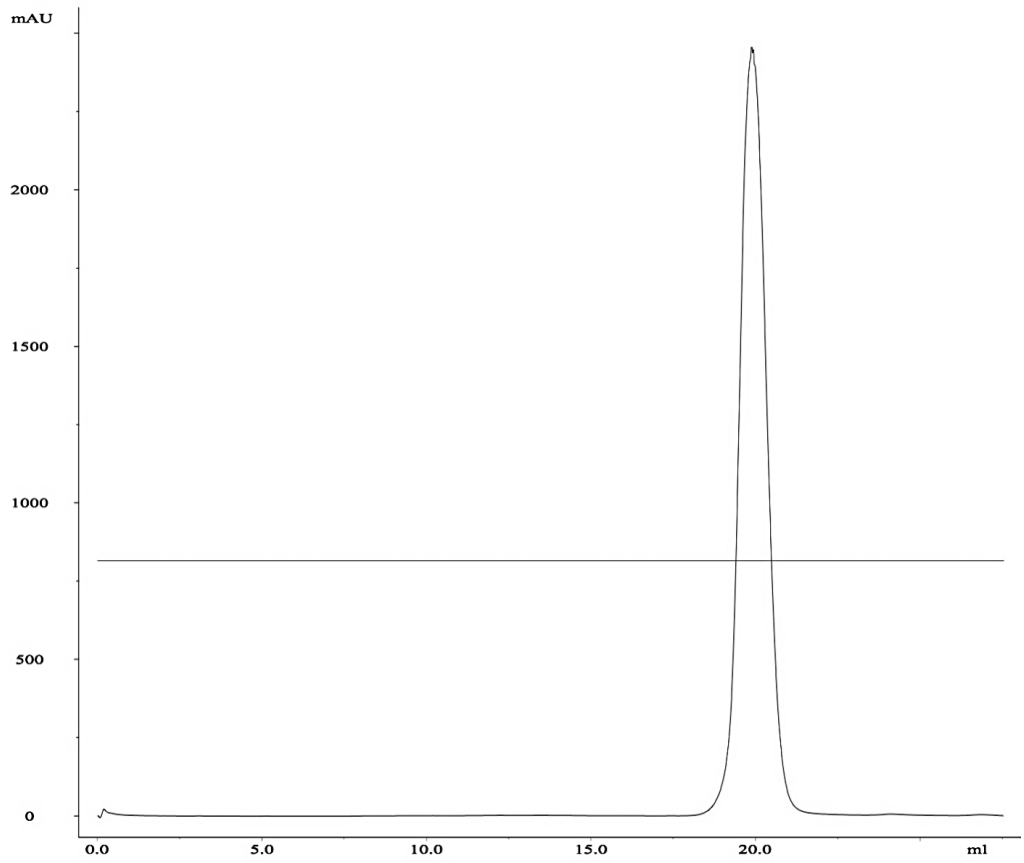

Figure 12 Chromatogram from SEC of concentrated HSP90* incubated with ATP. Axes y and x are measurements of A_{280} and mL, respectively.

5.3.3 Trouble-shooting with full-length HSP90

The N-terminal, middle and C-terminal domain of HSP90 are all possible binding sites for CTA1. As the first two have proven to depend on flexibility for function, any changes made to decrease overall protein flexibility could have a significant impact on this. The HSP90* construct could be incapable of binding to and, subsequently, refolding CTA1, as the deletions in this construct most definitely affects its flexibility. The previously deleted amino acids from the "charged linker" and C-terminal tail were for that reason reintroduced. This provided a full-length His-tagged HSP90 construct with a TEV recognition site (hereafter HSP90-His_6) that could be used in parallel experiments, as well as provide positive or negative controls for the effects of a linker region to CTA1-binding. The construct, which was also designed by Joël B. Heim and ordered through GenScript, still had an additional alanine in its N-terminal tail compared to the original full-length sequence due to cloning into the NcoI-XhoI cloning site of the pETM-11 vector, but this small and uncharged amino acid was not expected to have any impact on its binding capacity.

HSP90-His_6 proved more difficult to overexpress despite producing acceptable levels of the truncated construct with the previously established purification protocol. This might be due to the fact that this construct resembles the wildtype mammalian HSP90 to a greater extent, rather than bacterial HSP90 paralogs, which do not have extended linkers between their N-terminal and middle domains. The environment in which a eukaryotic gene is expressed in bacteria can affect protein production and stability, leading to a decreased production of the target protein than that of its own bacterial proteins. HSP90* could have been expressed in such great quantities because of its structural similarity to bacterial HSP90 compared to the more flexible eukaryotic protein. Another impacting factor on HSP90-His_6 over-expression was the choice of bacterial culture medium, as transcription of genes and bacterial growth are altered by the level of nutrients available to the bacteria. The LB medium used for HSP90*His_6 is commonly used for both growth and maintenance of recombinant *E. coli* strains, but it has a medium nutrition level in contrast to the highly nutrient-rich TB medium. Growth and expression tests were, therefore, performed with both LB and TB medium to find the optimal medium for growth and over-expression.

Figure 13 presents the average growth pattern of three BL21 (C43) *E. coli* cell culture in LB medium (grey) and TB medium (black) after induction of protein expression in the *lac* operon. Although the cultures were in their exponential growth phase when induced, the OD_{600} in the TB medium was 0.7 compared to 0.6 in LB medium. A difference in bacterial density is not ideal, possibly a pitfall, but a bacterial growth pattern could still be determined from the experiment. It was apparent after no more than two hours that the cell cultures grew significantly faster in TB medium rather than LB medium. Cells in TB medium continued to grow at an accelerated rate up until the 7-hour mark, where growth seemed to stagnate. In contrast, all cultures grown in LB medium seemed to have had a growth optimum at this point. The fact that cultures in TB medium reach a stationary phase prior to those in LB medium could be a result of the cells exhausting nutrients more rapidly as they grow at a faster rate. The maximum mean OD_{600} measurements for cells grown in LB medium and TB medium were 8.98 and 13.32, respectively. There was a tendency for death 20 hours after induction, but it was not determined at which hour this decline in growth started or the level of impact this had on protein expression. It is possible that nutrients released by dead and lysed cells are utilized by surviving cells to maintain a longer stationary phase where the target protein can still be produced.

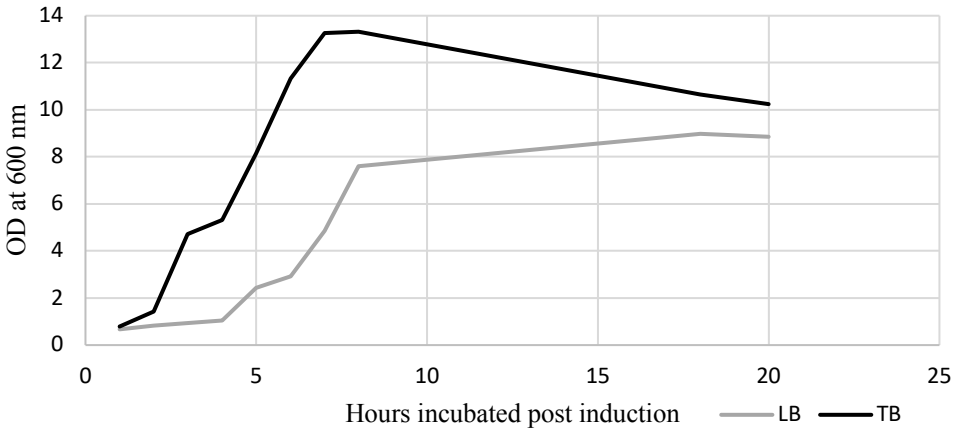

Figure 13 Mean growth curves for C43 cells expressing HSP90-His_6 in TB medium (black) and LB medium (grey) post induction from mean OD_{600} values measured for three incubated cultures each.

High bacterial densities do not correlate directly with a high expression of target genes, but instead serve as regulators for gene expression. A rapid increase in bacterial growth will amplify expression of genes closest to the origin of replication. Also, cells in nutrient-rich environments will coordinate expression of certain genes differently than those who are depleted of nutrients. The increased bacterial growth observed in TB medium was for those reasons not a deciding factor for choice of medium for production of HSP90-His_6.

Since TB medium was the most ideal medium for cellular growth of the IPTG-induced cells, growth tests were repeated in TB medium with added pre-induction OD_{600} measurements. This was performed to clarify whether or not over-expression of the *lac* operon had a significant impact on bacterial growth. If cellular growth stagnated after induction, it would imply that bacterial resources were re-distributed to express the target gene rather than genes associated with metabolic fluxes. The new measurements are presented in Figure 14 (black) where they are compared to those of Figure 13 (grey). Cells showed a similar increase in growth as the first test, presenting a clear tendency of rapid growth for BL21 (C43) cells in TB medium.

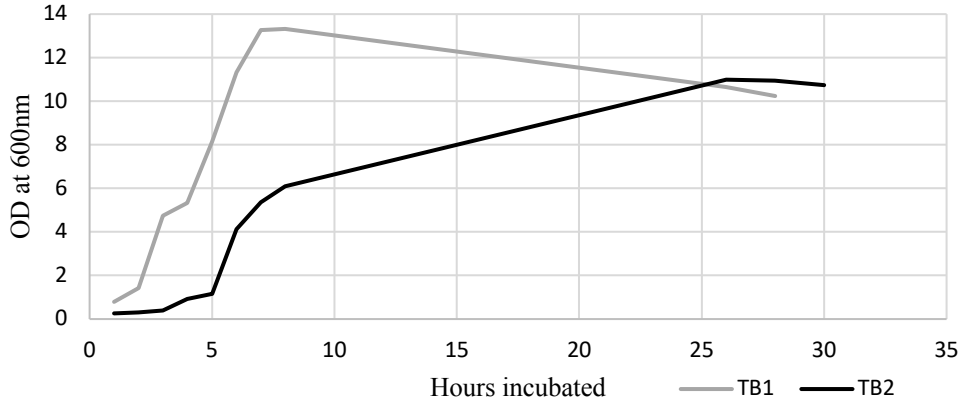

Figure 14 Mean growth curves for C43 cells expressing HSP90-His_6 in TB medium both pre and post induction (black) compared to TB medium graph from Figure 13 (grey). The cells (black) were induced after 4 hours.

There was an indication of bacterial growth stagnation 2 hours after induction (6 hours incubated), which was not clear in the original post-induction OD_{600} measurements. This could be due to the IPTG-induced over-expression of the target protein, and so HSP90-His_6 expression tests were performed in both LB and TB medium for the first 3 hours of induction and analyzed on an SDS-PAGE gel (Figure 15). Because of the re-introduced "charged linker" and C-terminal tail, a band of 86 kDa was expected for HSP90*-His_6, as well as an increased thickness of this band for each hour of induction. Cultures were diluted 1:10 to more clearly see separate bands with SDS-PAGE analysis, as they were often over-loaded when analyzing the protein extract and flow-through of HSP90*-His_6. The dilution successfully showed separate bands, but the cells did not show a significant increase in production of HSP90-His_6 in TB medium compared to LB. The thicker bands from the 3-hour induction (both well 3) is most likely due to the higher bacterial density in the samples and should, in retrospect, be diluted based on their respective OD_{600} measurements.

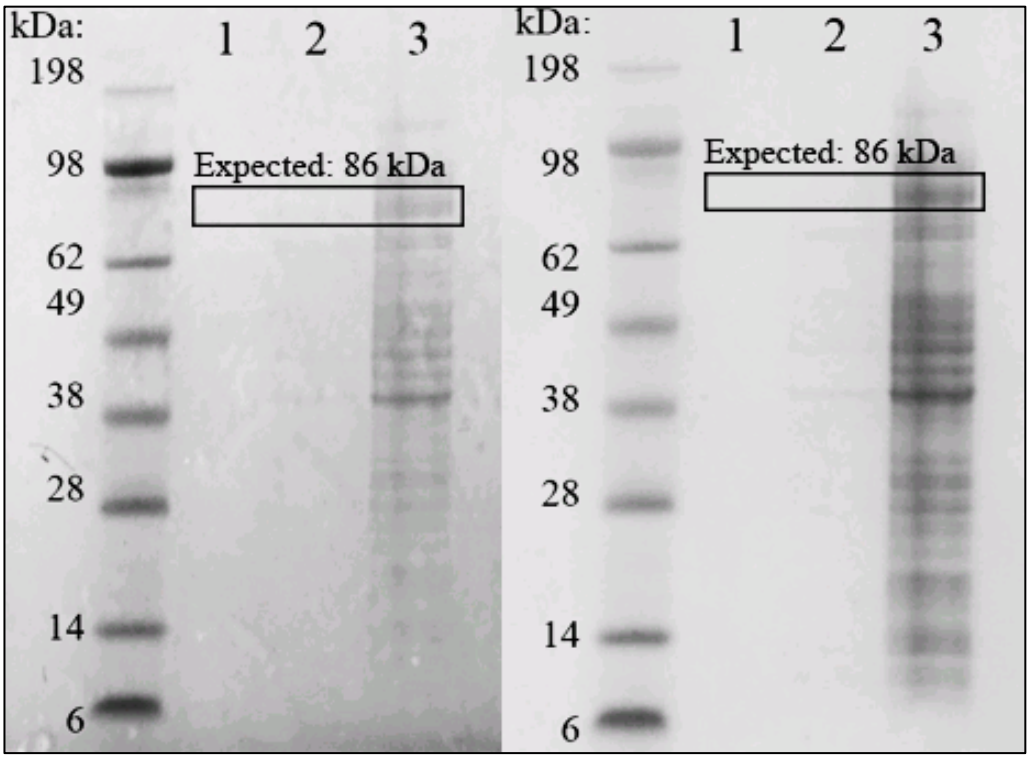

Figure 15 SDS-PAGE analysis of HSP90-His_6 expression in LB (left) and TB medium (right). SeeBlue Plus2 protein standard was loaded as a molecular weight reference in both gels. Wells marked 1-3 were loaded with 1) 1-hour induction, 2) 2-hour induction and 3) 3-hour induction. It was expected to see an increased band thickness of 86 kDa for HSP90-His_6 for each hour of induction.

The expected band of 86 kDa was more visible in TB medium cultures, indicating that TB is a superior medium for HSP90-His_6 over-expression. In addition to the increased level of nutrients, this could be due to the addition of phosphate, which maintains a physiological pH during exponential growth. The longer bacteria grow and the more slowly they are lysed, the more plasmid DNA can be expressed.

5.3.4 Purification of His-tagged HSP90

Purification tests of HSP90-His_6 were performed with the superior TB medium to verify target protein expression. It was important to check if the purification protocol of HSP90*-His_6 could be copied for this construct, as a difference in bacterial protein expression could impact the number of impurities eluted with the target protein. The most pressing impurity was that of the C-terminal cleavage product, which was expected to be gone now that the C-terminal tail was intact. The His tag could also have a lower affinity for the Ni-NTA column due to the additional amino acids in the HSP90-His_6 construct, which would make a change to a TALON resin material an option.

Figure 16 (left) presents an SDS-PAGE analysis of the fractions collected from an Ni-NTA column elution. Column wash and elutions were performed manually. None of the samples were diluted, but they still clearly show separate bands in contrast to SDS-PAGE analysis of purifications of the truncated construct. This is due to the decreased bacterial growth and protein expression observed for this construct. Although favorable when analyzing on a gel, protein concentrations were overall significantly lower in these tests. The protein was expected to elute with an imidazole concentration close to that of HSP90*-His_6 (140 mM), but a band of 86 kDa, the expected size of HSP90-His_6, is present in the 100 mM imidazole wash (well 5). This indicates that the additional amino acids have an effect on the binding affinity of the tag. The purification should be repeated using an ÄKTA purifier to determine the exact imidazole concentration for elution. Using a NanoPhotometer (A_{280} = 1.420), the target protein concentration was measured to 0.1 mg/mL in a 25 mL volume, giving a yield of 2.5 mg of protein per 2 L of cell culture (1.25 mg/L). This shows that the new full-length construct can be purified using the HSP90*His_6 protocol, but that it will produce a signifantly lower protein concentration.

In addition to the 86-kDa band, there is a hint of what appears to be a 70-kDa band. The size is similar to the C-terminal cleavage product observed in HSP90*-His_6 purifications. This could be the same product, despite the re-introduced amino acids in the C-terminal tail. The band should be analyzed by MS to confirm this if it continues to show in future purifications, and if confirmed, deletions in the C-terminal tail can be excluded as the cleavage culprit. The elution fraction from well 5 was collected and concentrated (Figure 16, right) to more clearly see how much of this impurity was present in the elution. Although very thin, the assumed cleavage product is still present and would not be able to separate by size exclusion. Other chromatographic methods could have been applied to remove this impurity, but it could also be separated from the target protein by adjusting the imidazole elution. The leap from a 50 to 100 mM imidazole elution was quite large, and so a second batch of HSP90-His_6 was purified with an 80 mM imidazole wash step instead of 50 mM (Figure 17). Although more target protein was lost in the wash, it successfully eluted all of the unwanted 70-kDa impurity (well 3) from the 100 mM imidazole elution (well 4). Due to an unfortunate mistake, self-produced TEV protease was not available and the His tag was as a result of this not removed prior to further experiments. A TEV digest should be performed to ensure that the tag is still cleavable for the purpose of crystallization.

Figure 16 SDS-PAGE analysis of HSP90-His_6 elution tests with an Ni-NTA column (left) and concentrated eluted target protein (right). SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Wells marked 1-6 (left) were loaded with 1) protein extract, 2) column flow-through, 3) 20 mM imidazole wash 4), 50 mM imidazole wash, 5) 100 mM imidazole wash 6) 150 mM imidazole wash. Wells marked 1-3 (right) were loaded with 1) 1 µg, 2) 5 µg and 3) 10 µg of concentrated 100 mM wash fraction.

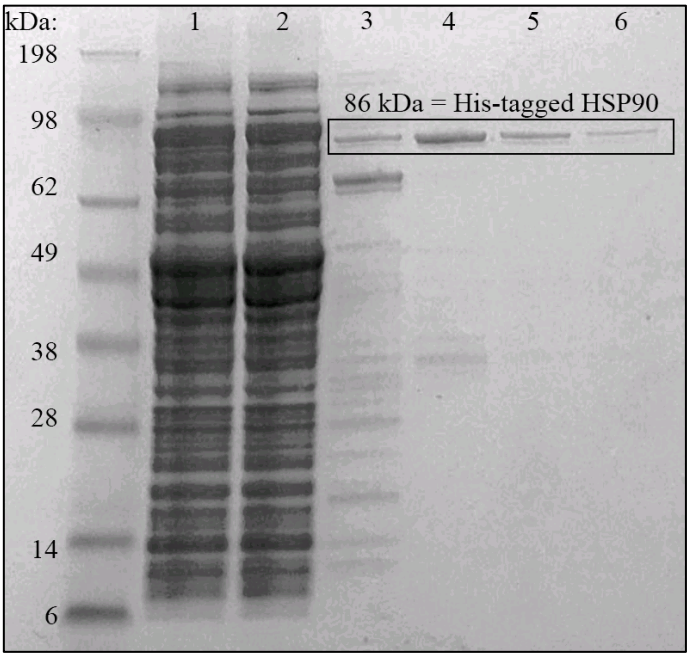

Figure 17 SDS-PAGE analysis of HSP90-His_6 Ni-NTA purification with an Ni-NTA column. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Wells marked 1-6 were loaded with 1) protein extract, 2) column flow-through, 3) 80 mM imidazole wash 4), 100 mM imidazole wash, 5) 150 mM imidazole wash and 6) 200 mM imidazole wash.

5.4 CTA1/HSP90 complex formation

5.4.1 Incubation tests under physiological conditions

Purified constructs of CTA1 and HSP90 were used for incubation tests under physiological conditions. CTA1 has proven unstable at temperatures higher than 4 C° during and after purifications discussed in this thesis, but it is believed to stabilize once bound to HSP90. To simulate the cellular mechanism of HSP90 *in vitro*, incubation tests were performed with CTA1-His_6 and both HSP90 constructs in the presence of ATP at 37 °C. ATP is essential for the dimerization of HSP90 N-terminal domains and subsequently its function, and so is believed to be a necessary additive to achieve complex formation. CTA1 will unfold again when subjected to higher temperatures, leaving it open for binding and refolding. The idea is that, if bound to HSP90, CTA1 will not be degraded by the physiological temperatures as it otherwise would in its unfolded state. As a heat shock protein, HSP90 is expected to remain folded and stable under the same conditions. Clear bands for both HSP90 and CTA1 on an SDS-PAGE gel would indicate that this is possible *in vitro*.

All incubation tests were performed with a higher ratio of HSP90 than CTA1 constructs due to the abundance of HSP90 in a eukaryotic cell and the tendency of CTA1 to aggregate when unfolded in higher concentrations. The proteins were as a starting point incubated in a 20 mM Tris-HCl and 150 mM NaCl buffer solution with a pH of 7.5. It was essential to have this concentration of NaCl in the solution, as it is the concentration under physiological conditions. Tris was the buffer of choice based on its buffering range at RT but was later discarded due to its increased pKa at physiological temperatures. Tris has a pKa of 8.07 at 25 °C, buffering effectively in the pH range of 7.1 to 9.1, but it is a highly temperature-sensitive buffer. Its pKa increases around 0.03 units per increased °C, meaning that its pKa at 37 °C changes to 7.43 to 9.43. This is borderline to the physiological pH used in these experiments and use of this buffer system is a possible pitfall in case of unsuccessful complex formation. Both HSP90 and CTA1 constructs have a pI considerably lower than the pH used in these experiments (4.98-6.21) and should not be considered a factor for protein precipitation in this buffer system.

HSP90* was first incubated with CTA1-His_6 in a 3:1 ratio. ATP was also added at a 1 mM concentration to ensure that HSP90* did not hydrolyze all ATP prior to CTA1-binding. Several incubation attempts were performed with varying concentrations of CTA1 and HSP90 constructs, but CTA1 kept falling out of solution after only 1 hour of incubation. Figure 18 presents an SDS-PAGE analysis of the loss of CTA1-His_6, but not HSP90*, in solution after 1, 2 and 3 hours of incubation (well 3-5) in the same test tube. Both constructs were loaded on the gel for size reference (well 1 and 2). HSP90* was not diluted as CTA1-His_6 and the bands do not represent the concentration at which the proteins started in the incubation solution as a result of this. It was clear that CTA1 was present in solution at the start of the experiment but fell out of solution already after 1 hour. Complex formation most likely did not occur, as CTA1 is not released by HSP90 once refolded and should stay in a stable complex if formed. There was no visible precipitation in the samples, but the protein concentration might have been too low to see this with the naked eye. Concentrations should have been measured before and after performing these experiments. It is difficult to say at which point CTA1 fell out of solution, but it is likely due to lack of binding to HSP90*. There are three major factors that could have caused HSP90* not to bind CTA1-His_6: 1) the high temperature led to instant CTA1 aggregation, 2) all ATP was hydrolyzed by HSP90* prior to CTA1-binding or 3) the less flexible HSP90* is incapable of or have difficulties binding CTA1.

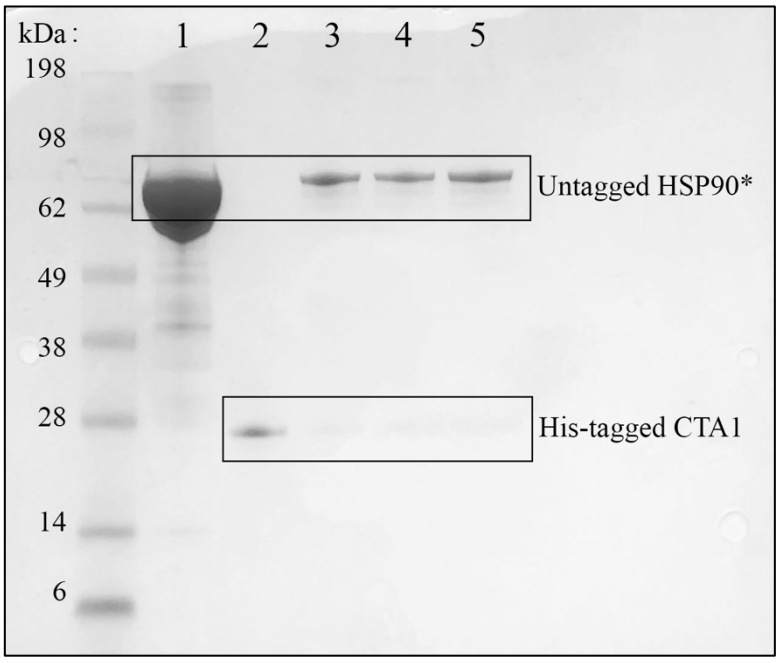

Figure 18 SDS-PAGE analysis of HSP90*-His_6/CTA1-His_6 incubation test. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Wells marked 1-5 were loaded with 1) concentrated HSP90*-His6, 2) diluted CTA1-His_6, 3) 1-hour incubation, 4) 2-hour incubation, 5) 3-hour incubation.

A new incubation test was performed with HSP90-His_6 and CTA1-His_6 in a repeated 3:1 ratio, but other changes were made to the protocol to optimize the conditions. Possible causative factor number 3 from lack of complex formation in Figure 18 was in these experiments no longer a factor, as the HSP90 construct had regained its native flexibility. Factor number 2 is difficult to correct without knowing exactly how much ATP is hydrolyzed by HSP90 in a given time interval, but the concentration was increased from 1 mM to 2 mM. Temperature-induced aggregation of CTA1 cannot be completely avoided if being consistent with a physiological temperature, but the hope was that the other changes would decrease the likelihood of this happening at a fast rate. The last change made to the protocol was that of the buffer system from Tris to phosphate (NaH_2PO_4), which has a pKa of 7.20, buffering in the pH range of 6.2 to 8.2 at RT. It is less temperature-sensitive than Tris with a pKa increase of only 0.0336 at 37 °C. This gives the system approximately the same buffering range as it has at RT, which is much more ideal for the pH used in this incubation test (7.4). Phosphate and sodium were already in solution after ATP hydrolysis by HSP90 and the added NaCl, respectively, and so the phosphate and sodium ions from this buffer were not believed to disrupt any interactions. An SDS-PAGE analysis of incubated HSP90-His_6 and CTA1-His_6 after 1, 2 and 3 hours is presented in Figure 19 together with concentrated HSP90-His_6 and diluted CTA1-His_6 for size reference.

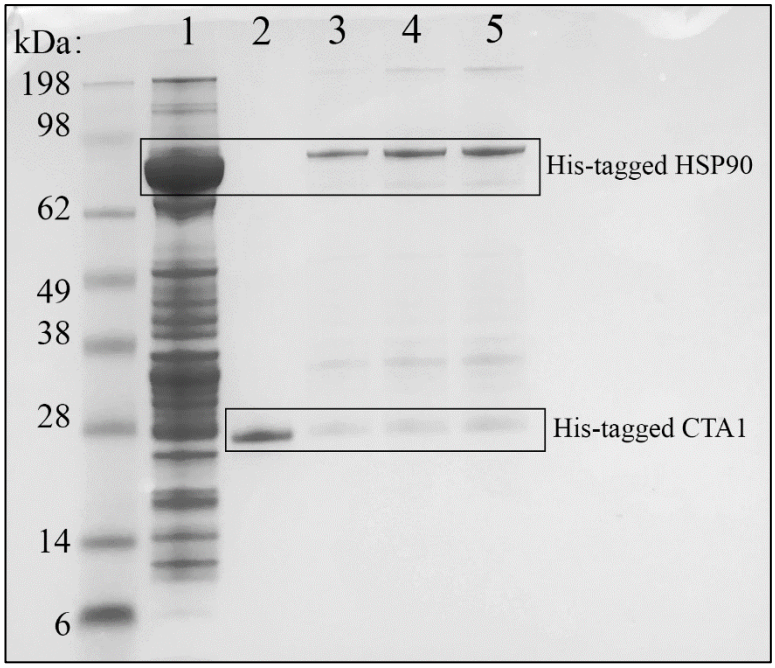

Figure 19 SDS-PAGE analysis of HSP90-His_6/CTA1-His_6 incubation test. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Wells marked 1-5 were loaded with 1) concentrated HSP90-His_6, 2) diluted CTA1-His_6, 3) 1-hour incubation, 4) 2-hour incubation, 5) 3-hour incubation.

HSP90-His_6 loaded as a size reference (well 1) was again taken from a concentrated sample and this should, in retrospect, be avoided. It gives the impression that most of the protein was lost during incubation, which was neither expected nor the case. Compared to HSP90*, the HSP90-His_6 sample had more impurities, but these were expected to not withstand the high incubation temperature. However, a sample of higher purity should be used in repeating experiments to rule out possible CTA1 interactions with the contaminants. CTA1-His_6 showed signs of a decreased concentration from the non-incubated sample (well 2), but it was still visible on the SDS-PAGE gel (well 3-5). Increased time of incubation also did not factor in on the strength of the band, indicating that some CTA1-His_6 was successfully stabilized by HSP90-His_6. One possible scenario is that the CTA1 construct did not instantly fall out of solution but was partially bound to HSP90-His_6, interacting weakly. While most CTA1-HSP90 interactions were lost, some might have increased in strength, stabilizing CTA1-His_6. It was unlikely that CTA1-His_6 had stayed in solution without interacting with HSP90-His_6 in some ways, but it was not possible to determine whether complete refolding occurred from the incubation test alone. This indicates that CTA1-His_6 can and will bind to full-length HSP90 under physiological conditions despite the introduced His tag. Whether or not the changes made to the protocol had any impact on this should be tested in repeating experiments. Samples are currently stored at -20 °C and could be further analyzed or possibly screened for crystal conditions. In retrospect, the samples should have been snap-frozen and stored at -80 °C to reduce protein damage.

5.4.2 TALON pull-down of CTA1 with HSP90

A new method of determining possible complex formation was introduced following the incubation tests with HSP90* and CTA1-His_6. It was called the "pull-down method". The idea was to pull down a complex of His-tagged HSP90 and untagged CTA1 from an affinity column, leaving all unbound CTA1 in the flow-through of the column. If complex formation had occurred after incubation, the hope was that CTA1 would elute together with the His-tagged HSP90 construct at an imidazole concentration above 100 mM. The untagged CTA1 construct was created mainly for the purpose of performing this experiment. A TALON resin was chosen as the column material for two reasons: 1) the incubated proteins were already in phosphate buffer and 2) a high selectivity for polyhistidines was preferable. An Ni-NTA resin might give an illusion of successful complex formation due to weakly bound CTA1. The buffers used for TALON purification normally had a pH of 8.0, but they were in this case adjusted to of 7.4 to simulate a physiological pH.

The first test with this "pull-down method" was performed with stored HSP90*His_6 and fresh CTA1. CTA1 and 1 mM ATP were incubated both with and without HSP90*-His_6, but the ratio was adjusted from the initial incubation tests to a 1:1 ratio. The idea was that if the truncated HSP90 construct had trouble binding to CTA1, an equally high concentration of the two proteins might increase the possibility of complex formation. This was, however, far from the expected protein ratio *in vivo* and possibly a factor for CTA1 aggregation. Figure 20 presents the incubation samples after 1 and 2 hours with ATP, which were the only samples showing any bands on this gel, despite also loading CTA1/ATP samples and TALON elutions after loading sample 1 and 2 on the column. In retrospect, the flow-through should also have been loaded, as this is the fraction where unbound CTA1 would appear. The analysis method was poorly executed, but it was at the time considered an indication of unsuccessful complex formation during incubation due to the close to non-existing 21.5 kDa band for CTA1 after both 1 and 2 hours. The method was put on hold, as CTA1 should ideally be purified the same day as the experiment and other purifications were prioritized at the time. His-tagged full-length HSP90 was at a later point tested with this method hoping to recover the believed stabilized CTA1 construct presented in Figure 19. An SDS-PAGE analysis of the fractions recovered from the TALON purification are shown in Figure 21.

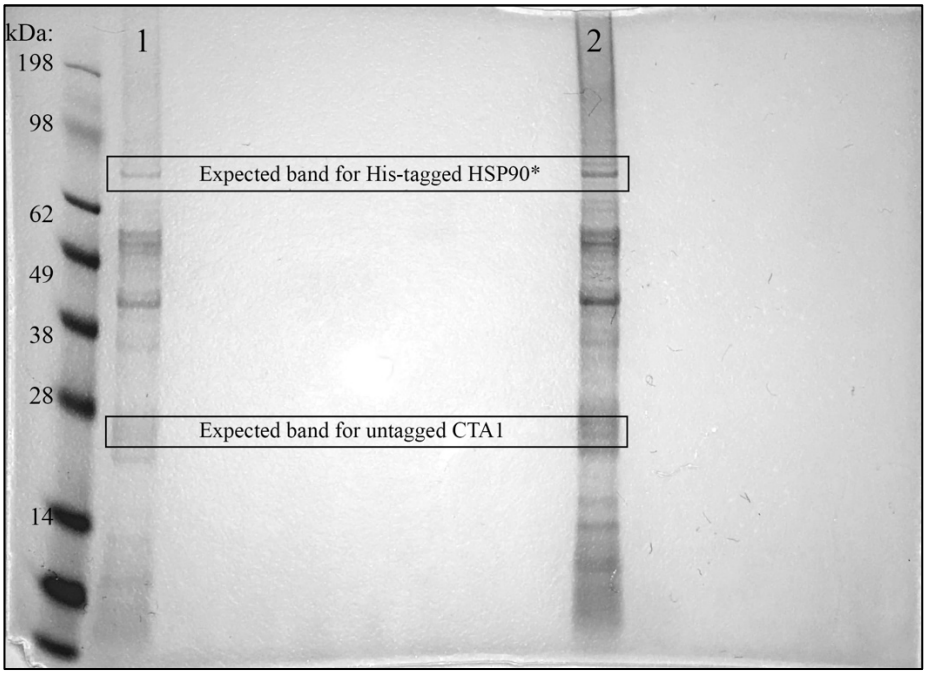

Figure 20 SDS-PAGE analysis of HSP90*-His_6/CTA1 pull-method after 1-hour (well 1) and 2-hour (well 2) incubations. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. No other wells are marked, as the proteins did not survive the TALON column.

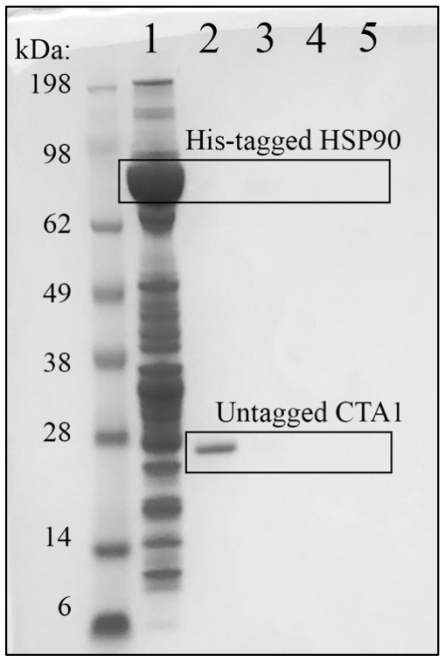

Figure 21 SDS-PAGE analysis of HSP90-His_6/CTA1 pull-down method after a 1-hour incubation. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Wells marked 1-5 were loaded with 1) concentrated HSP90*-His_6, 2) diluted CTA1, 3) column flow-through, 4) 20 mM imidazole wash, 5) 200 mM imidazole wash.

Samples taken from the flow-through (well 3), imidazole wash (well 4) and imidazole solution (well 5) did not show any bands on the gel. No target proteins were still in solution, but also no impurities. The first thought was that the protein concentration in the elution volumes (10 mL) was too small for any bands to be detected on a gel, as they were also unreadable with a NanoPhotometer. This machine can give incorrect measurements when concentrations are especially low, and so A_{280} for both constructs was measured manually with a quartz cuvette and the protein concentrations were calculated manually with their theoretical molar extinction co-efficient to double-check the lack of eluted protein. Unfortunately, the concentration of HSP90-His_6 and CTA1 were calculated to be 0.002 mg/mL and 0.009 mg/mL, respectively, which is an extreme decrease from the initial 1 mg/mL and 0.3 mg/mL that were incubated. The small protein concentrations measured are most likely due to a slightly incorrect adjustment of background noise with the buffer reference. If there really was no protein that survived the TALON column, there is a possibility that they precipitated on the column despite no change in pH or NaCl concentration. All metals will at some degree leak from an IMAC column, which could damage the loaded proteins, but this is less frequently observed with a TALON column. In retrospect, the incubated solution should have been loaded on the gel as to ensure that the proteins were not lost during the 1-hour incubation. The pull-down method should not be discarded, but instead repeated with an optimized SDS-PAGE analysis.

5.4.3 Preliminary experiments

Expression of CTA1 and HSP90 in the same cell is another possible method for obtaining a stable complex *in vitro*. Co-expression can prove advantageous to CTA1 solubility as well as increase the chance of refolding by HSP90 during and after protein expression. This would require inactivation of CTA1 toxic activity, which can be obtained by a glutamate-to-aspartate substitution in residues 110 and 112 [87] underlined in the amino acid sequence of CTA1. An inactivated CTA1 construct (hereafter x-CTA1) was designed and confirmed with sequencing by Joël B. Heim. Co-transformation was attempted with this construct together with both HSP90*-His_6 and HSP90-His_6 into competent BL21 (C43) cells, since they had already proven effective for cell growth and HSP90*-His_6 protein expression. Cells expressing pLysE plasmids were not tested, as they were proven to grow very slowly, and repression of toxic gene expression was also no longer necessary for CTA1. Since the plasmids encoding x-CTA1 and HSP90 constructs also encode genes for Amp and Kan resistance, respectively, there was a double-selection for co-transformed cells on Amp + Kan LB agar plates.

Colonies have yet to be observed with both selection markers after four attempted co-transformations but is still believed to be obtainable. New co-transformation attempts should be repeated using a less expensive *E. coli* strain such as BL21 (DE3). When successful, the transformed cells should be induced for over-expression of the target proteins in small-scale LB medium expression tests and analyzed on an SDS-PAGE gel to confirm co-expression. Both proteins are expected to be soluble in this case. Induction should ideally be performed at 37 °C. Parallel HSP90 expression tests should be performed at 37 °C to ensure that this is also an appropriate temperature for over-expression of this protein. If both proteins are expressed in the cell cultures, this would be a major break-through for stabilizing x-CTA1 in the presence of either truncated or full-length HSP90. One pitfall of this method is that any additional ATP in the medium would not be available for hydrolysis by HSP90 in the bacterial cytosol, but the ATP produced by the bacteria might be sufficient for dimerization of its N-terminal domains. If it is not, ATP could easily be added after cell lysis.

5.5 CdtB

5.5.1 Production of His-tagged CdtB

A CdtB construct had to be produced and purified in the lab for crystal screening with ATP. It was important that the construct had a removable His tag and was proven to bind ATP, and so the His-tagged CdtB construct with a Thrombin cleavage site received from the collaborating group of K. Teter was used. It was sent to GATC Biotech (now Eurofins genomics) for sequencing prior to any purifications to ensure that it was identical to the given sequence information. Residue number 53 was found to be an aspartic acid rather than the listed asparagine, which was corrected by mutagenesis using the given primers in the Appendix. The protein was re-sequenced after mutagenesis and this confirmed the successful point mutation underlined in the amino acid sequence of CdtB-His_6. The purification protocol for this protein was established by the Teter group and was similar to that of CTA1. Extraction of unfolded CdtB-His_6 from inclusion bodies gave an unexpectedly high yield compared to that of CTA1. The yield prior to refolding was on average 2 mg/mL in a 30 mL volume, giving a total of 60 mg per 2 L of culture (30 mg/L). Figure 22 illustrates how the Ni-NTA column in this case did not bind all target protein, as a band of around 32.5 kDa, the expected size of CdtB-His_6, elutes with an imidazole wash of only 20 mM (well 4) and continues to elute in high amounts with increasing imidazole concentrations (well 5-6).

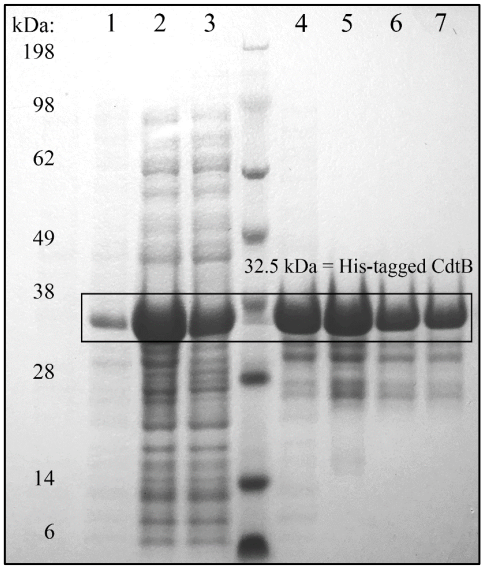

Figure 22 SDS-PAGE analysis of CdtB-His_6 expression test and Ni-NTA purification. SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Well 1 was loaded with culture after a 20-hour induction of protein expression. Wells marked 2-7 were loaded with 2) protein extract, 3) column flow-through, 4) 20 mM imidazole wash, 5) 50 mM imidazole wash, 6) 100 mM imidazole wash, 7) 200 mM imidazole wash.

The presence of the target protein in the column flow-through (well 3) indicated that the column might have been over-saturated with His-tagged protein, or possibly that the protein extract and wash buffers should have been loaded at a slower rate than the recommended 1 mL per minute for these columns. If not carefully washed, the fractions eluted from each wash step will not be separated properly. The exact imidazole concentration at which elution occurred was not established with an ÄKTA purifier due to the corrosive urea in the buffers. The purification was repeated manually for a new batch with the same target protein yield, but with a loading rate of 0.5 mL per minute and increased volumes of imidazole wash. Figure 23 presents the successful binding of all target protein from the second batch to the Ni-NTA column. There was again an over-loaded band of approximately 32.5 kDa in the protein extract (well 1), but this band was no longer visible in the column flow-through or the 20 mM wash (well 2 and 3). CdtB-His_6 was eluted (well 4) and collected for refolding. Refolded CdtB-His_6 was run through an SD200 column (Figure 24) to remove residual impurities and the fractions were analyzed on an SDS-PAGE gel (Figure 25). The main peak (well 2-5) was clearly separated from the shoulder peak (well 1), successfully obtaining a highly pure sample of CdtB-His_6.

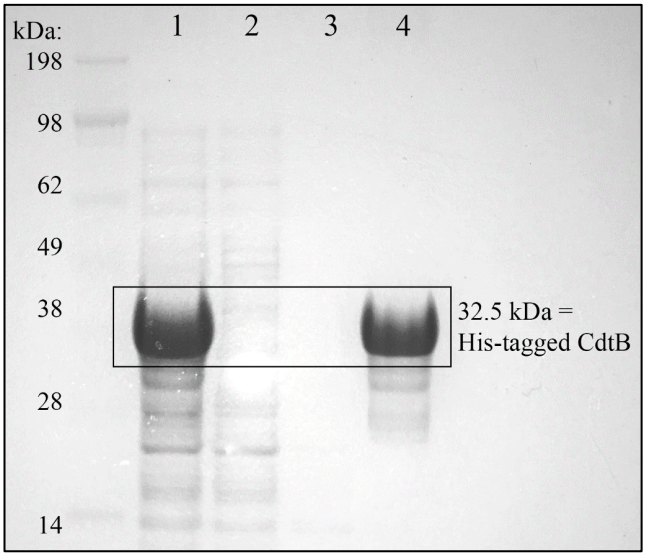

Figure 23 SDS-PAGE analysis of CdtB-His_6 Ni-NTA purification. SeeBlue Plus2 protein standard was loaded as a molecular weight reference, but bands were too weak to visualize. An approximate interpretation has been added. Wells marked 1-4 were loaded with 1) protein extract, 2) column flow-through, 3) 20 mM imidazole wash, 4) 100 mM imidazole wash.

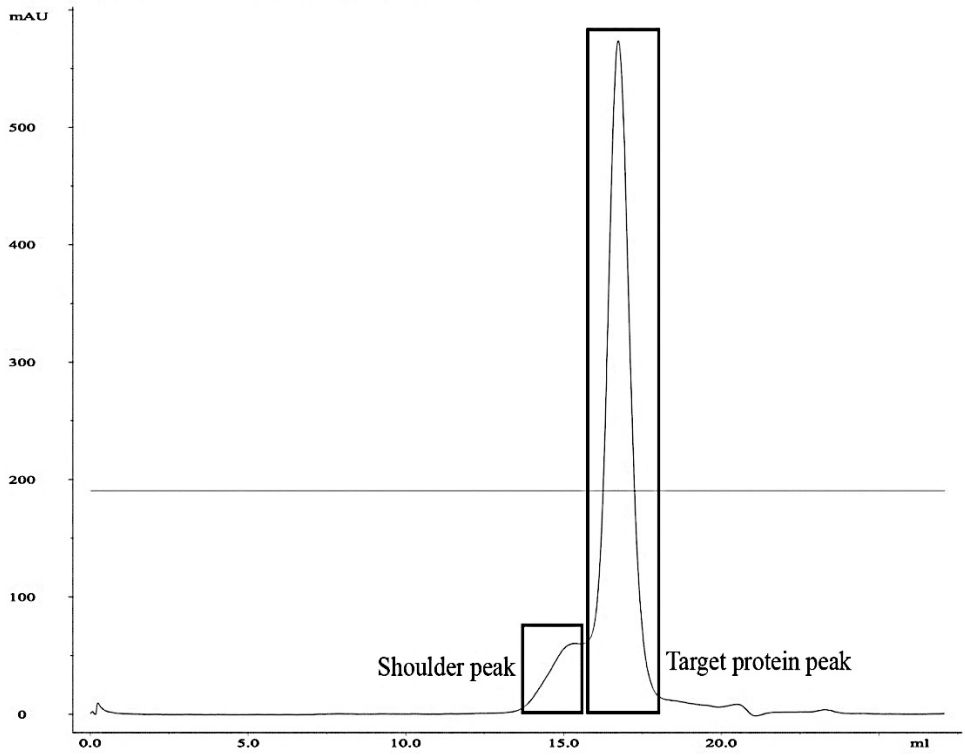

Figure 24 Chromatogram of CdtB-His_6 SEC. Absorbance was measured at 280 nm and peaks represent eluted protein. The target protein peak and its shoulder peak were analyzed on and SDS-PAGE gel (Figure 25) to confirm size and purity of CdtB-His_6.

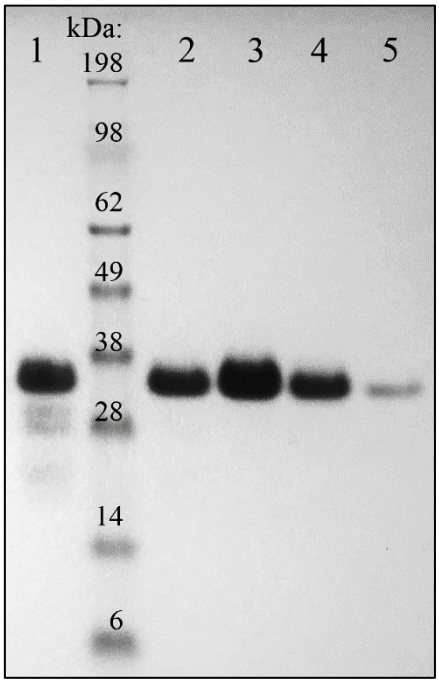

Figure 25 SDS-PAGE analysis of CdtB SEC fractions (Figure 24). SeeBlue Plus2 protein standard was loaded as a molecular weight reference. Well 1 was loaded with a fraction collected from the shoulder peak. Well 2-5 were loaded with fractions collected from the target protein peak.

5.5.2 Trouble-shooting with protein stability and Thrombin cleavage

The refolding of CdtB-His_6 proved surprisingly more unstable than that of both CTA1 constructs. Up to 60 % of the protein misfolded and aggregated after the urea decrease from 2 M to 0 M at 4 °C, but loss of protein was not limited to this dialysis step. It was apparent after a few purifications that storage of CdtB, both with and without its His tag, for more than two days at 4°C led to visible protein precipitation. Protein concentrations were also measured using a NanoPhotometer (A_{280} = 1.050) both before and after storage overnight and were reduced by 5-10 %. In contrast to the instability of CTA1 constructs at temperatures higher than 4 °C, CdtB-His_6 was unstable at both higher and lower temperatures. This became especially problematic when following the Thrombin Sepharose Beads protocol for cleaving of its His tag, which proved to be another significant step that resulted in loss of protein due to precipitation. The recommended Thrombin cleavage test involves a low target protein concentration (maximum 1 mg/mL) and a maximum 4-hour incubation period to obtain successful cleavage without any off-target cleaving. It was expected that the protein was fully cleaved after 4 hours, but as presented in Figure 26, this was not the case. Samples were taken after each hour and loaded on the gel in volumes of 5 and 10 μL to ensure that the size difference between cleaved and uncleaved CdtB would be visible. A lot of CdtB-His_6 still remained in the sample after 4 hours, and additional cleavage tests had to be performed. After several tests with various temperatures and time intervals, it was concluded that cleavage at 10 °C overnight with stirring gave the highest amount of cleavage product and the least protein precipitation. The Thrombin digest was still not 100 % successful, and an additional purification step was introduced to provide the most homogenous protein solution possible for crystal screening.

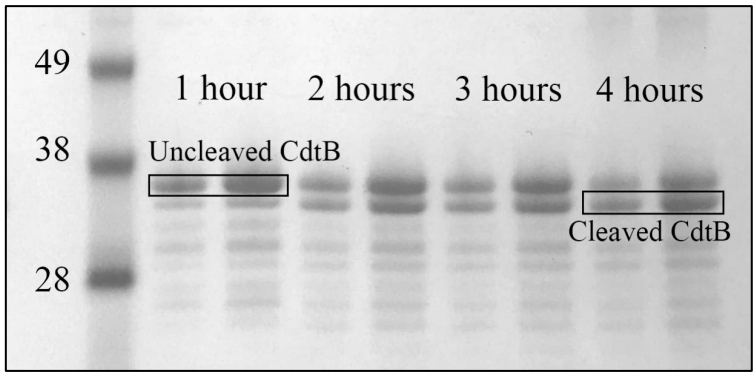

Figure 26 SDS-PAGE analysis of CdtB/Thrombin digest test. SeeBlue Plus2 protein standard was loaded as a molecular weight reference, but the gel is cropped to the relevant area. Wells were loaded with samples from every hour of a 4-hour incubation period at RT. 5 and 10 μL volumes were loaded from each hour.

A TALON resin material was chosen to separate tagged and untagged CdtB due to its high selectivity for polyhistidines. CdtB-His_6 had already proven to elute with a low imidazole concentration, so the use of an Ni-NTA resin could possibly bind cleaved CdtB by mistake and results in a much lower yield. Cleaved CdtB has an expected size of either 29 kDa or 30 kDa due to the possible promiscuity of Thrombin. The main cleavage site, LVPR/GS, had been introduced intentionally, while the secondary cleavage site, GR/GS, was possibly introduced unintentionally downstream of the main site when the gene was cloned into pET-28b+. This introduced a possible problem for crystal screening, as a sample with two different cleavage products could not produce crystals with a repeating pattern of just one construct. An overnight digest would give Thrombin time to cleave at both of these sites. It was expected that a mixed solution would show as separate bands or one larger blurred band if analyzed on an SDS-PAGE gel. Figure 27 shows how the use of a TALON column successfully separates the digest product (well 2) into purely cleaved CdtB (well 3) and uncleaved CdtB (well 5). The 0 mM imidazole wash (well 4) only show one thin band of cleaved CdtB, indicating that only one cleavage site has been recognized by Thrombin, most likely LVPR/GS that produces CdtB with a size of 30 kDa. Another possible scenario is that CdtB-His_6 was first cleaved at the main cleavage site and then the secondary site, but this would also produce a homogenous sample of 29-kDa sized protein as long as all molecules were cleaved at both sites.

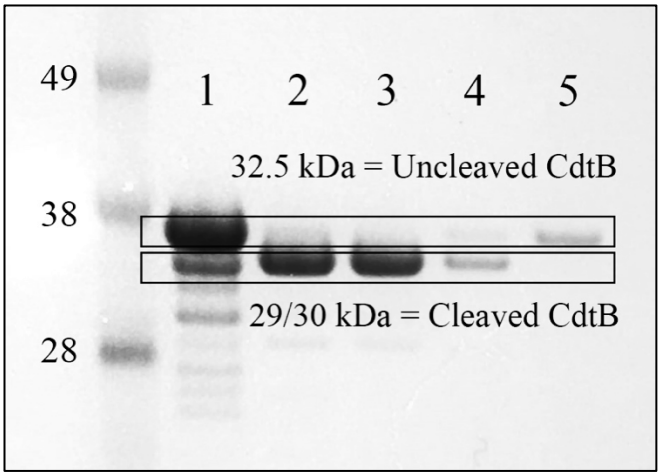

Figure 27 SDS-PAGE analysis of CdtB/Thrombin digest test. SeeBlue Plus2 protein standard was loaded as a molecular weight reference, but the gel is cropped to the relevant area. Well 1-5 were loaded with 1) undigested product, 2) digest product, 3) column flow-through, 4) 0 mM imidazole wash and 5) 200 mM imidazole wash.

Protein precipitation during storage and His tag cleaving was significant enough to question whether the buffers and pH used were the most optimal for this protein, especially since cleavage of the His tag did not improve protein solubility. The protocol established for CdtB purification by the Teter group was almost identical to that of CTA1, but the toxins have very different isoelectric points (pI). Proteins are less soluble in buffer systems with a pH closer to their pI due to their less charged surfaces, meaning that the pH in the buffers used for purifications has to be adjusted to the pI of individual proteins. As a rule, the pH should not be less than +/- 1 unit from its theoretical pI. CdtB and CdtB-His_6 have a pI of 9.15 and 9.50, respectively, and were purified in a buffer system with a pH of 8. This is borderline to the stated rule and could be the reason for the continuing precipitation of both CdtB constructs. It would explain why cleavage of the His tag did not improve solubility, as this construct has the lowest pI of the two.

To determine if the pH of the system was indeed the cause for precipitation, CdtB-His_6 were dialyzed against Tris and phosphate buffers with pH 6.5, 7.0 and 7.5. The temperatures were held at a constant 10 °C, as this was already established to be the most stable temperature. Samples were stored overnight to simulate the conditions in which the protein had previously precipitated, and the protein concentration was measured three times for each sample prior to and after storage. The mean values were calculated and given for storage in phosphate buffer (Table 1) and Tris buffer (Table 2).

Table 1 Uncleaved CdtB pH precipitation test with 20 mM phosphate buffer and 200 mM NaCl and varying pH. Each 500-µL sample was dialyzed against a 250 mL buffer system. All tests were executed at 10°C.

Sample	**1**	**2**	**3**
pH	6.5	7.0	7.5
Concentration before dialysis	1.20 mg/mL	1.20 mg/mL	1.20 mg/mL
Concentration after storage	1,17 mg/mL	1,02 mg/mL	0,90 mg/mL

Table 2 Uncleaved CdtB pH precipitation test with 50 mM Tris-HCl buffer and 200 mM NaCl. Each 500-µL sample was dialyzed against a 250 mL buffer system. All tests were executed at 10°C.

Sample	**4**	**5**	**6**
pH	6.5	7.0	7.5
Concentration before dialysis	1.20 mg/mL	1.20 mg/mL	1.20 mg/mL
Concentration after storage	1,13 mg/mL	1,19 mg/mL	0,80 mg/mL

The originally used phosphate buffer system was tested with a 1.20 mg/mL concentration of the uncleaved construct. This resembled the average protein concentration during refolding, storage and Thrombin digest. It was very clear after overnight storage and re-measuring the concentration that even a pH of just 7.5 resulted in precipitation of more protein than it did in a buffer with a lower pH. Almost no protein was lost when stored at pH 6.5, which clearly shows that the pH used for purification of CTA1 constructs cannot be used for CdtB purifications. A Tris buffer system was also tested with the same concentrations of the protein and with the same pH range as that of the phosphate buffers. Precipitation was again evident at a pH of 7.5, but here with the least loss of protein in a solution with a pH of 7.0 instead of 6.5. This could be due to the fact that the buffering capacity of Tris is quite poor at this pH. The two stability tests combined prove that the buffer system in the CdtB purification protocol should be completely changed from that of CTA1. One possible buffer optimal for a lower pH is the Bis-Tris (bis(hydroxyethyl)iminotris(hydroxymethyl)methane) buffer, which is effective in the pH range of 5.8 to 7.2. Stability and precipitation tests should be performed in this buffer parallel to repeated tests in phosphate and Tris buffers to confirm its possibly stabilizing effects.

CdtB-His_6 samples were also tested for possible protein precipitation after freezing, as one sample had previously visibly precipitated when thawed after storage at -20 °C without snap-freezing. Sample 4 presented in Table 2 was further tested for the effects of long-term storage on protein concentration with and without glycerol at -20 °C and -80 °C. The results, presented in Table 3, show that the addition of 10 % (v/v) glycerol to the sample prior to freezing resulted in the least loss of protein. Glycerol is an alcohol that increases the solubility of proteins, and so help keep CdtB-His_6 in solution when thawed. However, glycerol in the sample could be problematic when trying to crystallize the protein, as it might affect crystal nucleation growth rate. The difference between loss of protein with and without glycerol was fairly insignificant and should be avoided if crystallization is the final goal, as in the case of CdtB.

Table 3 CdtB-His_6 freeze-and-thaw precipitation test of sample 4 from Table 2. The total aliquots were 500 µL.

Sample	**7**	**8**	**9**	**10**
Freezing temperature	-20°C	-20°C	-80°C	-80°C
Glycerol		10%		10%
Snap-frozen			X	X
Concentration before thawing	1.13 mg/mL	1.13 mg/mL	1.13 mg/mL	1.13 mg/mL
Concentration after thawing	0.81 mg/mL	1.09 mg/mL	1.00 mg/mL	1.05 mg/mL

5.5.3 Incubation and crystal screening of CdB with ATP

CdtB constructs were purified for the main purpose of forming a complex with ATP. It has already proven to bind ATP and had to be screened for optimal conditions for crystallization. If protein crystals were obtained with some of the screened conditions, it would lay the ground work for setting up crystal plates to determine its interacting residues with ATP. To accomplish this, the protein was screened for optimal conditions with JCSG *plus*, PACT *premier* and MORPHEUS. This ensured that a range of pH, buffers, precipitants and ligands were tested. In addition to the composition of the reservoir solution, the protein concentration was also important to adjust so that protein precipitation occurred at a slow rate necessary for formation of an ordered protein crystal.

Purified CdtB-His_6 and CdtB were both successfully incubated with 1 mM ATP in 20 mM Tris-HCl pH 7.5 at RT without visible precipitation, despite the proven precipitation of this protein in that buffer system and at higher temperatures. It is possible that ATP stabilizes the protein in a similar way to that of HSP90 to CTA1, which would explain the role of an ATP binding site. Stabilization of CdtB by ATP could prove to be the cause of, or at least a factor in, its dissociation from CdtA and CdtC. Crystal screening of incubated CdtB constructs were performed four times during this toxin project, setting up 2-3 of the primary screens each time with drops of 25 % and 50 % protein concentration to ensure that enough protein was present in the drop to achieve some level of precipitation. Drops were applied with an Oryx4 crystallization robot and ranged in a total volume of 0.5 to 1.5 µL based on the amount of protein available for screening. Smaller drops in several screens were preferred to larger drops in a single screen, as it was important to test as many formulations as possible.

JCSG *plus* produced what a appeared to be a crystal in two separate screening attempts, but in the same condition (B11). The drops were 25 % (v/v) of 5 mg/mL and 7.5 mg/mL protein in a reservoir solution of 1.6 M sodium citrate tribasic dihydrate (pH 6.5) with no added buffer or precipitant. Crystals indicated presence of protein with an ultraviolet-visible (UV-vis) filter, but fishing of these two crystals was not successful. They powdered very easily, which was another indication that the crystals were formed by protein since salt crystals require more force as they are denser. In retrospect, images of these crystals should have been saved prior to crystal fishing. PACT *premier* screening gave tiny crystals in conditions with and without PEGs, but these were discarded as salt crystals after adding a UV-vis filter over the images collected.

The MORPHEUS screen with cryo-protected formulations included a range of low molecular weight ligands such as monosaccharides and amino acids, but none of the formulations with these ligands produced any crystals in the three months of storage in a crystal hotel. However, one formulation without a ligand produced what appeared to be three protein crystals that were also attempted to fish from the drop. The drop was 25 % (v/v) 5.0 mg/mL protein in a reservoir solution of 0.03 M of 10% (w/v) PEG 20 0000 and 20% (w/v) PEG MME (polyethylene glycol monomethyl ether) 550 in a 0.1 M MES (N-morpholino ethanesulfate)/imidazole (pH 6.5) solution. Fishing from this drop required a cryo-protectant (10% (v/v) glycerol). Two out of three crystals were crushed during fishing attempts due to the formation of a disrupting membrane over the drop, and so the last crystal was fished by Joël B. Heim to ensure that at least one crystal was spared for X-ray diffraction. The crystal was mounted on a loop, stored in liquid nitrogen, and tested with the BioMax beamline of MAX IV Laboratory. No data was recovered from this crystal and is assumed to have been a salt crystal.

Crystal formation of *E. coli* CdtB was achieved in 2006 with high-throughput crystallization screening in three conditions [88]. These conditions, presented in Table 4, could possibly also crystallize the HdCdtB construct with bound ATP. The PEG additives are favorable for crystal growth, but these conditions have very high concentrations, making the reservoir solutions highly viscous. The pH in which crystal form 1 grew is surprisingly low, but the high pI of CdtB might make a lower pH more ideal for crystallization to avoid immediate precipitation. A hanging-drop crystal screening was performed with 1 mL drops of 50 % (v/v) 10 mg/mL CdtB incubated with 1 mM ATP in six formulations. Table 5 presents the six adjusted formulations from Table 4 for which incubated CdtB/ATP samples were screened.

Table 4 Growth conditions for the formation of three *E. coli* CdtB crystals established by high-throughput crystallization in 2006. New formulations were made for hanging-drop crystal screening in the lab based on these conditions and presented in Table 5.

Crystal form 1	**Crystal form 2**	**Crystal form 3**
100 mM ammonium thiocyanate	100 mM calcium chloride hydrate	100 mM calcium chloride hydrate
100 mM sodium citrate pH 4	100 mM HEPES pH 7.0	100 mM HEPES pH 7.0
40 % (w/v) PEG 8000	20 % (w/v) PEG 4000	80 % (w/v) PEG 400

Table 5 CdtB 10 mg/ml, 6 µL 50/50 hanging drop, with and without ATP incubation, 15/5.

	1	**2**
A	100 mM sodium thiocyanate 100 mM ammonium citrate pH 4.5 80 % (v/v) PEG 400	100 mM sodium thiocyanate 100 mM ammonium citrate pH 5.0 80 % (v/v) PEG 400
B	100 mM HEPES pH 6.5 100 mM calcium chloride dihydrate 20 % (v/v) PEG 4000	100 mM HEPES pH 7.5 100 mM calcium chloride dihydrate 20 % (v/v) PEG 4000
C	100 mM HEPES pH 6.5 100 mM calcium chloride dihydrate 40 % (v/v) PEG 8000	100 mM HEPES pH 7.5 100 mM calcium chloride dihydrate 40 % (v/v) PEG 8000

Level of precipitation did not correlate with percentage of PEGs added, but rather with a decreased pH. This was not surprising, as pH is a very important parameters in the process of protein crystallization due to its effect on solubility and electrostatic interactions of the protein in solution. The actual pH of each drop is not the one stated in Table 5, but rather a combination of the contributing salts and buffers. However, all drops that contained salts or buffers with a pH between 4.5 and 6.5 showed immediate heavy precipitation after adding protein to the drops, while those with a pH of 7.5 showed no signs of precipitation at all. Based on the pH stability tests presented in Table 1 and Table 2 in section **5.5.2**, the effect of pH was expected to go the other way, as in increased visible precipitation with an increased pH. It is not a good starting point for crystallization when almost all protein falls out of solution, but this does not mean that a lower pH should be discarded. The protein concentration in the applied drops (10 mg/mL) might have been too high, which is easily adjustable in repeating experiments. This hanging-drop experiment should also be adjusted and repeated with an increased range of buffers, salts and PEGs to clearly see their effects on crystallization. A lower pH (< 6.5) is still believed to be ideal for crystallization.

Crystal formation can take many months to achieve and formulations should not be discarded if protein precipitation is too low or too high in a drop. Concentrations of CdtB were possibly too high in the hanging-drop experiment and sitting-drop crystal screening with JCSG *plus* and MORPHEUS confirmed that a protein concentration between 5-7.5 mg/mL can give an appropriate amount of precipitate for crystal formation. Any of the stored screens could still produce a crystal with ATP-bound CdtB.

5.6 TEV protease

5.6.1 Expression and production

The sequences of both HSP90 constructs in this thesis have a TEV cleavage site, ENLYFQG, for removal of their N-terminal His tags. As the tags could increase the already high flexibility of HSP90, TEV protease had to be produced in high quantities for this project. The construct was a superfolder (*sf)* GFP-TEV-His_6 variant and the purification was based on an already established protocol by X. Wu *et al.* in 2009 [89]. Since the TEV fusion protein contained a His tag itself, IMAC was performed to remove unwanted impurities. The tag was expected to interact strongly with an affinity column, and as a high yield was preferred to a high purity, the protein extract was loaded onto an Ni-NTA resin. Figure 28 (well 1-4) shows an SDS-PAGE analysis of this purification. Well 1 and 2, loaded with protein extract and flow-through of the column, respectively, show the variety of proteins extracted, as well as the successful over-expression and binding of the target protein. Both TEV and GFP have a molecular mass of approximately 27 kDa each, and will, therefore, as a fusion protein have an approximate size of 54 kDa. A protein of this size is over-expressed in the extract, missing in the flow-through, and clearly eluted from the column at a 500 mM concentration of imidazole. Some, but not much, is lost in the 40 mM imidazole wash, where it is clear that the protein is slightly larger than 49 kDa. The fluorescent GFP provided an easily recognizable protein marker in the elution.

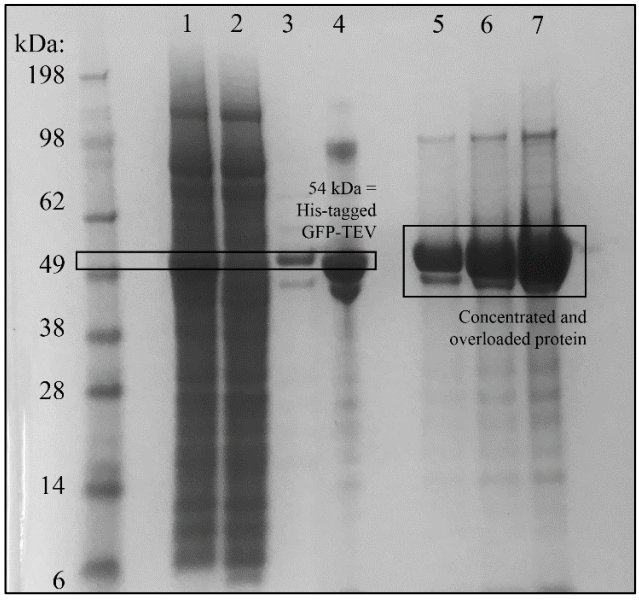

Figure 28 SDS-PAGE analysis of fractions eluted from His_6-TEV-GFP Ni-NTA purification and concentrated protein. SeeBlue Plus2 protein standard was loaded a molecular weight reference. Wells marked 1-4 were loaded with 1) protein extract, 2) column flow-through, 3) 40 mM imidazole wash, 4) 500 mM imidazole wash. The 500 mM elution fraction was concentrated and loaded in wells marked 5-7 as 5) 1 µg, 6) 11 µg and 7) 25 µg of protein.

5.6.2 Purity and stability

The residual impurities in the target protein elution could have been removed if a gradient of increasing imidazole concentrations had been applied. Well 5-6 in Figure 7 show 1 μg, 11 μg, and 25 μg of loaded protein to clearly see these impurities. However, this was not an important factor when purifying TEV. As this protease was to be used in great quantities over a long period of time, it was more important to gain a high yield than a high purity. As an enzyme, its activity is reduced if not stored properly, and so ethylenediaminetetraacetic acid (EDTA) and glycerol were added to stabilize the protein before storage at -20 °C. Despite these precautions, reduced activity was observed by other group members after 10-12 months of storage. Storing the enzyme at -80 °C instead could prove to be a solution to this problem. The activity of TEV protease was tested and discussed for HSP90*-His_6 in section **5.3.2**, where it demonstrated successful cleavage in its recognition sequence.

Summary and future perspectives

The crystal structures of CTA1 (O'Neal *et al.*) and CdtB (D. Nešić *et al.*) were solved and refined in 2004, but questions still remain regarding their interactions with host machineries for toxin activation. CTA1 was proven by Taylor *et al.* in 2010 to depend on the molecular chaperone HSP90 for refolding and translocation to the cytosol, while CdtB of *H. ducreyi* has just recently been proven by the collaborating group of K. Teter at UCF to bind ATP as a free subunit. Interacting residues between the two toxins and the host machineries HSP90 and ATP have not been established due to their unstable nature. This thesis provides insight into factors affecting the stability of CTA1 and CdtB during protein purification, storage and incubation with their binding partners.

All constructs of CTA1, HSP90 and CdtB used in this project were successfully purified using the previously established protocols or adaptations of these protocols. Optimization was especially important for CdtB, proven to be more stable in a new buffer system with a pH of no higher than 7.0 and a working temperature of 10 °C. The second extensive protocol optimization was that of the later introduced full-length HSP90 and its poor expression, which indicated that a change in culture medium from LB to TB could provide a higher expression of this construct. Both full-length and truncated HSP90 also exhibited what appears to be self-cleavage in their C-terminal tails, initially thought to be due to sequence deletions, but now an unsolved mystery.

For complex formation, His-tagged CTA1 and full-length His-tagged HSP90 were proven to both remain stable in solution after a 3-hour incubation at 37 °C. As CTA1 is highly unstable as a single subunit at that temperature, HSP90 must have bound to some of the unfolded CTA1, possibly partly or completely refolded it. The experiment should be repeated to confirm this, and crystal screens could be set up immediately following incubation to achieve the final goal of co-crystallization. Crystal screening of ATP-incubated CdtB gave inconclusive results, but many nuclei are still growing and should be regularly checked for crystal formation and protein content. ATP also proved to be a surprisingly stabilizing factor for CdtB at room temperature. The dissociation constant should be determined for both CTA1/HSP90 and CdtB/ATP in the near future to provide insight into how strong their interactions are, possibly with isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR). The provided insight into toxin and complex stability in this thesis lays a baseline for the conditions in which future experiments should be performed, as well as those they should not.

Appendix

I. Materials

Table A: Chemicals

Chemical	
Agar-agar	Merck
Ammonium citrate	Sigma
Ampicillin	Merck
Benzonase nuclease	Novagen
BL21 (C43) *E. coli* cells	Sigma
BL21 (DE3) *E. coli* cells	Sigma
Bolt LDS 4-12% Plus	Invitrogen
Bolt LDS Sample Buffer (4X)	Invitrogen
Bolt Sample Reducing Agent (10X)	Invitrogen
Bolt MES SDS Running Buffer (20X)	Invitrogen
Calcium chloride	Merck
Chloramphenicol	Fluka Biochemika
Coomassie Brilliant Blue G250	Fluka Biochemika
DTT	AppliChem
EDTA	VWR
Glycerol	Merck
HEPES	Thermo Fischer
Hydrochloric acid	Merck
Imidazole	Sigma
IPTG	Sigma
Kanamycin	AppliChem
Lysozyme	Sigma
Magnesium chloride	Merck
Ni-NTA Superflow	QIAGEN

Pancreatic peptone	VWR
PEG solutions	Fluka
PMSF	Sigma
Potassium dihydrogen phosphate	VWR
SeeBlue Plus2 protein standard	Invitrogen
Sodium chloride (buffers)	Merck
Sodium chloride (medium)	Merck
Sodium dihydrogen phosphate	Merck
Sodium hydroxide	Merck
Sodium thiocyanate	Sigma
Tris-HCl	VWR
Triton 100X	Sigma
Tryptone (peptone)	VWR
Urea	VWR
Virkon	RelyOn
Yeast extract	VWR

Table B: Chromatographic column materials

Material	
HisTrap HP, 5 mL	GE healthcare
HiTrap TALON crude, 1 and 5 mL	GE healthcare
SD 200 R10/300	GE healthcare

Table D: Kits

Kit	
JCSG *plus*	Molecular Dimensions
Morpheus	Molecular Dimensions
PACT *premier*	Molecular Dimensions
Q5® Site-Directed Mutagenesis	NEB
QIAprep Spin Miniprep 250	QIAGEN

Table C: Equipment

Equipment	
2 Lens Crystallization plate, 96-wells	Swissci
Cell spreaders	VWR
Conical centrifuge tubes, 15 and 50 mL	FALCON
Disposable Cuvettes, 1.5 mL	BRAND/Sigma-Aldrich
Disposable Serological Pipets, 10 and 25 mL	VWR
Disposable Transfer Pipets, 5 mL	VWR
Filter Upper Cups	VWR
Glass beads, 1 mm	BioSpec
Microcapillary Pipet Tips	VWR
Nitrile gloves, AQL 1.5	VWR
Optifit Refill Tips	Sartorius
PCR tubes	VWR
Petri dishes, polystyrene	VWR
Safe-Lock Tubes, 1.5 and 2 mL	Eppendorf
Siliconized cover slides	Hampton Research
SnakeSkin Dialysis Tubing, 10K MWCO	Thermo Fischer
Syringe filters, 0.45 µm	VWR Chemicals
Syringes, 20 mL	SOFT-JECT
Thrombin Sepharose Beads, 1 mL	BioVision
Ultra-15-Centrifugal Filter Device, 10K and 50K MWCO	Millipore

Table E: Hardware

Hardware	
Avanti Centrifuge J-26 XP	Beckman Coulter
ÄKTA Purifier	GE Healthcare
BeadBeater	BioSpec
Biofuge Fresco	Heraeus
Biofuge Pico	Heraeus
Centrifuge 5810 R	Eppendorf
Electrophoresis Power Supply EPS 601	GE Healthcare
Eppendorf Thermomixer comfort	Eppendorf
Forma ULT Freezer -86C	Thermo Scientific
Jenway 3510 pH meter	Jenway
Kelvitron B6120 Incubator	Heraeus
Mini Orbital Shaker SSM1	Stuart
MINIPULS 3	Gilson
Multitron Shaking Incubator	VWR
NanoPhotometer	IMPLEN
Oryx4	Douglas instruments
RCT basic	IKA
Thermomixer comfort	Eppendorf
Wacom Intuos tablet	Wacom

Table F: Software

Software	
ApE	Wayne Davis
CorelDRAW	Corel Suites
Photoshop	Adobe
PyMol	Schrödinger
SketchBook	Sketch App
WaspRun	Douglas Instruments

II. Solutions

Table S1: Media used for bacterial culturing and agar plates. Weight of material is given per 1 L of solution.

	LB medium	**TB medium**	**Agar plates**
Yeast extract	5 g	24 g	—
Tryptone	10 g	12 g	15 g
NaCl	10 g	—	10 g
NaH_2PO_4	—	12 g	—
Agar-agar	—	—	12 g

Table S2: Buffers used for NiNTA purification and dialysis of TEV-GFP.

	His AC	**His WB**	**His EC**	**His D**	**His S**
Tris-HCl	50 mM	50 mM	50 mM	50 mM	50 mM
NaCl	150 mM	150 mM	150 mM	150 mM	150 mM
Imidazole	20 mM	40 mM	500 mM	—	—
Glycerol	10%	10%	10%	10%	80%
EDTA	—	—	—	—	0.5 mM
pH	8.0	8.0	8.0	8.0	8.0

Table S3: Buffers used for NiNTA purification of HSP90 constructs.

	His A	**His B**	**His C**	**HSP-D**	**TEV-D**
Tris-HCl	20 mM	20 mM	20 mM	20 mM	50 mM
NaCl	200 mM	200 mM	200 mM	200 mM	150 mM
Imidazole	20 mM	500 mM	—	20 mM	20 mM
pH	7.5	7.5	7.5	7.5	8.0

Table S4: Buffers used for NiNTA purification of CTA1 and CdtB constructs.

	His A.1	**His A.1**	**His A.3**	**His B.1**
Tris-HCl	20 mM	20 mM	20 mM	20 mM
NaCl	200 mM	200 mM	200 mM	200 mM
Imidazole	20 mM	20 mM	20 mM	500 mM
Urea	—	2 M	8 M	—
pH	8.0	8.0	8.0	8.0

Table S5: Buffers used for dialysis of CTA1 and CdtB constructs.

	D1	**D2**	**D3**	**D4**	**D5**
NaH_2PO_4	20 mM	20 mM	20 mM	20 mM	20 mM
NaCl	—	—	—	—	200 mM
Urea	6 M	4 M	2 M	—	—
pH	8.0	8.0	8.0	8.0	8.0

Table S6: Buffers used for TALON purifications of CdtB constructs and the CTA1/HSP90 "pull-down method".

	Talon A	**Talon B**	**Talon dialysis**	**Thrombin**
NaH_2PO_4	20 mM	20 mM	20 mM	—
Tris-HCl	—	—	—	50 mM
NaCl	200 mM	200 mM	200 mM	150 mM
Imidazole	—	200 mM	—	—
pH	8.0	8.0	8.0	7.5

III. Nucleic acid sequences

Hsp90*-His_6

```
1     CCATGGCGCC  GGAGGAAGTT  CACCACGGCG  AAGAAGAAGT  TGAGACCTTT  GCGTTTCAAG
61    CGGAGATTGC  GCAGCTGATG  AGCCTGATTA  TTAACACCTT  CTACAGCAAC  AAGGAAATTT
121   TTCTGCGTGA  GCTGATCAGC  AACGCGAGCG  ACGCGCTGGA  TAAGATTCGT  TATGAGAGCC
181   TGACCGACCC  GAGCAAACTG  GATAGCGGTA  AAGAACTGAA  GATTGACATC  ATTCCGAACC
241   CGCAGGAGCG  TACCCTGACC  CTGGTTGACA  CCGGTATCGG  CATGACCAAA  GCGGATCTGA
301   TTAACAACCT  GGGTACCATC  GCGAAGAGCG  GCACCAAAGC  GTTCATGGAA  GCGCTGCAGG
361   CGGGTGCGGA  TATCAGCATG  ATTGGCCAAT  TCGGTGTTGG  CTTTTACAGC  GCGTATCTGG
421   TGGCGGAGAA  GGTGGTTGTG  ATCACCAAAC  ACAACGACGA  TGAACAATAT  GCGTGGGAGA
481   GCAGCGCGGG  TGGCAGCTTC  ACCGTTCGTG  CGGACCACGG  CGAACCGATT  GGTCGTGGCA
541   CCAAGGTGAT  CCTGCACCTG  AAAGAAGATC  AGACCGAGTA  TCTGGAGGAA  CGTCGTGTTA
601   AAGAGGTGGT  GAAGAAACAC  AGCCAATTTA  TCGGTTACCC  GATTACCCTG  TATCTGGAAA
661   AGGAGCGTGA  AAAAGAGATT  AGCGACAAAA  TTAAAGAGAA  GTACATCGAT  CAGGAAGAGC
721   TGAACAAAAC  CAAGCCGATT  TGGACCCGTA  ACCCGGACGA  TATCACCCAA  GAGGAATACG
781   GTGAATTCTA  TAAGAGCCTG  ACCAACGACT  GGGAGGATCA  CCTGGCGGTT  AAACACTTTA
841   GCGTGGAAGG  CCAGCTGGAG  TTCCGTGCGC  TGCTGTTTAT  CCCGCGTCGT  GCGCCGTTCG
901   ACCTGTTTGA  AAACAAGAAA  AAGAAAAACA  ACATCAAACT  GTACGTTCGT  CGTGTGTTCA
961   TCATGGACAG  CTGCGATGAA  CTGATTCCGG  AGTATCTGAA  CTTTATCCGT  GGTGTTGTGG
1021  ACAGCGAAGA  TCTGCCGCTG  AACATTAGCC  GTGAGATGCT  GCAGCAAAGC  AAAATCCTGA
1081  AGGTTATTCG  TAAGAACATC  GTGAAGAAAT  GCCTGGAACT  GTTCAGCGAA  CTGGCGGAGG
1141  ACAAAGAAAA  CTACAAGAAA  TTCTATGAGG  CGTTTAGCAA  AAACCTGAAG  CTGGGTATCC
1201  ACGAAGACAG  CACCAACCGT  CGTCGTCTGA  GCGAGCTGCT  GCGTTACCAC  ACCAGCCAAA
1261  GCGGCGATGA  AATGACCAGC  CTGAGCGAGT  ATGTTAGCCG  TATGAAGGAA  ACCCAGAAAA
1321  GCATCTACTA  TATTACCGGC  GAGAGCAAGG  AACAAGTGGC  GAACAGCGCG  TTCGTTGAAC
1381  GTGTGCGTAA  ACGTGGCTTT  GAGGTTGTGT  ACATGACCGA  ACCGATCGAC  GAGTATTGCG
1441  TGCAGCAACT  GAAGGAATTC  GATGGTAAAA  GCCTGGTTAG  CGTGACCAAA  GAGGGTCTGG
1501  AGCTGCCGGA  GGACGAGGAA  GAGAAGAAAA  AGATGGAAGA  GAGCAAAGCG  AAGTTTGAGA
1561  ACCTGTGCAA  ACTGATGAAG  GAAATTCTGG  ATAAAAAGGT  TGAGAAGGTG  ACCATCAGCA
1621  ACCGTCTGGT  TAGCAGCCCG  TGCTGCATTG  TGACCAGCAC  CTACGGTTGG  ACCGCGAACA
1681  TGGAACGTAT  CATGAAAGCG  CAGGCGCTGC  GTGACAACAG  CACGATGGGT  TATATGATGG
1741  CGAAAAAGCA  CCTGGAGATT  AACCCGGATC  ACCCGATCGT  TGAAACCCTG  CGTCAAAAGG
1801  CGGAGGCGGA  CAAAAACGAT  AAGGCGGTGA  AAGACCTGGT  TGTGCTGCTG  TTCGAAACCG
1861  CGCTGCTGAG  CAGCGGTTTT  AGCCTGGAGG  ATCCGCAGAC  CCACAGCAAC  CGTATTTACC
1921  GTATGATCAA  ACTGGGTCTG  GGCTGAGACG  AAGATGAGGT  TGCGGCGGAA  GAGCCGAACG
1981  CGGCGGTGCC  GGATGAAATC  CCGCCGCTGG  AAGGTGACGA  AGATGCGAGC  CGTATGGAAG
2041  AGGTGGACTA  ACTCGAG
```

Hsp90-His_6

```
1     ATGAAACATC  ACCATCACCA  TCACCCCATG  AGCGATTACG  ACATCCCCAC  TACTGAGAAT
61    CTTTATTTTC  AGGGCGCCAT  GCCGGAGGAA  GTTCACCACG  GCGAAGAAGA  AGTTGAGACC
121   TTTGCGTTTC  AAGCGGAGAT  TGCGCAGCTG  ATGAGCCTGA  TTATTAACAC  CTTCTACAGC
181   AACAAGGAAA  TTTTTCTGCG  TGAGCTGATC  AGCAACGCGA  GCGACGCGCT  GGATAAGATT
241   CGTTATGAGA  GCCTGACCGA  CCCGAGCAAA  CTGGATAGCG  GTAAAGAACT  GAAGATTGAC
301   ATCATTCCGA  ACCCGCAGGA  GCGTACCCTG  ACCCTGGTTG  ACACCGGTAT  CGGCATGACC
361   AAAGCGGATC  TGATTAACAA  CCTGGGTACC  ATCGCGAAGA  GCGGCACCAA  AGCGTTCATG
421   GAAGCGCTGC  AGGCGGGTGC  GGATATCAGC  ATGATTGGCC  AATTCGGTGT  TGGCTTTTAC
481   AGCGCGTATC  TGGTGGCGGA  GAAGGTGGTT  GTGATCACCA  AACACAACGA  CGATGAACAA
541   TATGCGTGGG  AGAGCAGCGC  GGGTGGCAGC  TTCACCGTTC  GTGCGGACCA  CGGCGAACCG
601   ATTGGTCGTG  GCACCAAGGT  GATCCTGCAC  CTGAAAGAAG  ATCAGACCGA  GTATCTGGAG
661   GAACGTCGTG  TTAAAGAGGT  GGTGAAGAAA  CACAGCCAAT  TTATCGGTTA  CCCGATTACC
721   CTGTATCTGG  AAAAGGAGCG  TGAAAAAGAG  ATTAGCGACG  ACGAGGCGGA  GGAAGAGAAG
781   GGTGAAAAAG  AAGAGGAAGA  TAAGGACGAT  GAGGAAAAGC  CGAAAATCGA  GGACGTGGGT
841   AGCGATGAGG  AAGACGATAG  CGGCAAGGAC  AAGAAAAAGA  AAACCAAAAA  AATTAAAGAG
901   AAGTACATCG  ATCAGGAAGA  GCTGAACAAA  ACCAAGCCGA  TTTGGACCCG  TAACCCGGAC
961   GATATCACCC  AAGAGGAATA  CGGTGAATTC  TATAAGAGCC  TGACCAACGA  CTGGGAGGAT
1021  CACCTGGCGG  TTAAACACTT  TAGCGTGGAA  GGCCAGCTGG  AGTTCCGTGC  GCTGCTGTTT
1081  ATCCCGCGTC  GTGCGCCGTT  CGACCTGTTT  GAAAACAAGA  AAAAGAAAAA  CAACATCAAA
1141  CTGTACGTTC  GTCGTGTGTT  CATCATGGAC  AGCTGCGATG  AACTGATTCC  GGAGTATCTG
1201  AACTTTATCC  GTGGTGTTGT  GGACAGCGAA  GATCTGCCGC  TGAACATTAG  CCGTGAGATG
1261  CTGCAGCAAA  GCAAAATCCT  GAAGGTTATT  CGTAAGAACA  TCGTGAAGAA  ATGCCTGGAA
1321  CTGTTCAGCG  AACTGGCGGA  GGACAAAGAA  AACTACAAGA  AATTCTATGA  GGCGTTTAGC
1381  AAAAACCTGA  AGCTGGGTAT  CCACGAAGAC  AGCACCAACC  GTCGTCGTCT  GAGCGAGCTG
1441  CTGCGTTACC  ACACCAGCCA  AAGCGGCGAT  GAAATGACCA  GCCTGAGCGA  GTATGTTAGC
1501  CGTATGAAGG  AAACCCAGAA  AAGCATCTAC  TATATTACCG  GCGAGAGCAA  GGAACAAGTG
1561  GCGAACAGCG  CGTTCGTTGA  ACGTGTGCGT  AAACGTGGCT  TTGAGGTTGT  GTACATGACC
1621  GAACCGATCG  ACGAGTATTG  CGTGCAGCAA  CTGAAGGAAT  TCGATGGTAA  AAGCCTGGTT
1681  AGCGTGACCA  AAGAGGGTCT  GGAGCTGCCG  GAGGACGAGG  AAGAGAAGAA  AAAGATGGAA
1741  GAGAGCAAAG  CGAAGTTTGA  GAACCTGTGC  AAACTGATGA  AGGAAATTCT  GGATAAAAAG
1801  GTTGAGAAGG  TGACCATCAG  CAACCGTCTG  GTTAGCAGCC  CGTGCTGCAT  TGTGACCAGC
1861  ACCTACGGTT  GGACCGCGAA  CATGGAACGT  ATCATGAAAG  CGCAGGCGCT  GCGTGACAAC
1921  AGCACGATGG  GTTATATGAT  GGCGAAAAAG  CACCTGGAGA  TTAACCCGGA  TCACCCGATC
1981  GTTGAAACCC  TGCGTCAAAA  GGCGGAGGCG  GACAAAAACG  ATAAGGCGGT  GAAAGACCTG
2041  GTTGTGCTGC  TGTTCGAAAC  CGCGCTGCTG  AGCAGCGGTT  TTAGCCTGGA  GGATCCGCAG
2101  ACCCACAGCA  ACCGTATTTA  CCGTATGATC  AAACTGGGTC  TGGGCATTGA  CGAAGATGAG
2161  GTTGCGGCGG  AAGAGCCGAA  CGCGGCGGTG  CCGGATGAAA  TCCCGCCGCT  GGAAGGTGAC
2221  GAAGATGCGA  GCCGTATGGA  AGAGGTGGAC  TAA
```

CTA1-His_6

```
1     ATGAATGATG  ATAAGTTATA  TCGGGCAGAT  TCTAGACCTC  CTGATGAAAT  AAAGCAGTCA
61    GGTGGTCTTA  TGCCAAGAGG  ACAGAGTGAG  TACTTTGACC  GAGGTACTCA  AATGAATATC
121   AACCTTTATG  ATCATGCAAG  AGGAACTCAG  ACGGGATTTG  TTAGGCACGA  TGATGGATAT
181   GTTTCCACCT  CAATTAGTTT  GAGAAGTGCC  CACTTAGTGG  GTCAAACTAT  ATTGTCTGGT
241   CATTCTACTT  ATTATATATA  TGTTATAGCC  ACTGCACCCA  ACATGTTTAA  CGTTAATGAT
301   GTATTAGGGG  CATACAGTCC  TCATCCAGAT  GAACAAGAAG  TTTCTGCTTT  AGGTGGGATT
361   CCATACTCCC  AAATATATGG  ATGGTATCGA  GTTCATTTTG  GGGTGCTTGA  TGAACAATTA
421   CATCGTAATA  GGGGCTACAG  AGATAGATAT  TACAGTAACT  TAGATATTGC  TCCAGCAGCA
481   GATGGTTATG  GATTGGCAGG  TTTCCCTCCG  GAGCATAGAG  CTTGGAGGGA  AGAGCCGTGG
541   ATTCATCATG  CACCGCCGGG  TTGTGGGAAT  TCACACCATC  ACCATCACCA  TTGA
```

→ Introduction of stop codon for removal of His_6

	Primers (uppercase = target-specific)	T_m	T_a
Fw	TGGGAATTCAtgaCATCACCATCACCATTGATAATGCAGCC	73.2 °C	69.6 °C
Rv	CAACCCGGCGGTGCATGA	66.4 °C	

CTA1

```
1     ATGAATGATG  ATAAGTTATA  TCGGGCAGAT  TCTAGACCTC  CTGATGAAAT  AAAGCAGTCA
61    GGTGGTCTTA  TGCCAAGAGG  ACAGAGTGAG  TACTTTGACC  GAGGTACTCA  AATGAATATC
121   AACCTTTATG  ATCATGCAAG  AGGAACTCAG  ACGGGATTTG  TTAGGCACGA  TGATGGATAT
181   GTTTCCACCT  CAATTAGTTT  GAGAAGTGCC  CACTTAGTGG  GTCAAACTAT  ATTGTCTGGT
241   CATTCTACTT  ATTATATATA  TGTTATAGCC  ACTGCACCCA  ACATGTTTAA  CGTTAATGAT
301   GTATTAGGGG  CATACAGTCC  TCATCCAGAT  GAACAAGAAG  TTTCTGCTTT  AGGTGGGATT
361   CCATACTCCC  AAATATATGG  ATGGTATCGA  GTTCATTTTG  GGGTGCTTGA  TGAACAATTA
421   CATCGTAATA  GGGGCTACAG  AGATAGATAT  TACAGTAACT  TAGATATTGC  TCCAGCAGCA
481   GATGGTTATG  GATTGGCAGG  TTTCCCTCCG  GAGCATAGAG  CTTGGAGGGA  AGAGCCGTGG
541   ATTCATCATG  CACCGCCGGG  TTGTGGGAAT  TCATGA
```

Received CdtB-His_6

```
1     ATGGGCAGCA  GCCATCATCA  TCATCATCAC  AGCAGCGGCC  TGGTGCCGCG  CGGCAGCCAT
61    ATGGCTAGCA  TGACTGGTGG  ACAGCAAATG  GGTCGCGGAT  CCAACTTGAG  TGACTTCAAA
121   GTAGCAACTT  GGAATCTGCA  AGGTTCTTCA  GCAGTAGATG  AAAGTAAATG  GAATATTAAT
181   GTGCGCCAAT  TATTATCGGG  AGAACAAGGT  GCAGATATTT  TGATGGTACA  AGAAGCGGGT
241   TCCTTACCAA  GTTCGGCAGT  AAGAACCTCA  CGGGTAATTC  AACATGGGGG  AACGCCAATT
301   GAGGAATATA  CTTGGAATTT  AGGTACTCGT  TCCCGCCCAA  ATATGGTCTA  TATTTATTAT
361   TCTCGTTTAG  ATGTTGGGGC  AAACCGAGTG  AACTTAGCTA  TCGTGTCACG  CCGTCAAGCC
421   GATGAAGCTT  TTATCGTACA  TTCTGATTCT  TCTGTGCTTC  AATCTCGCCC  TGCAGTAGGT
481   ATCCGCATTG  GTACTGATGT  ATTTTTTACA  GTGCATGCTT  TGGCCACAGG  CGGTTCTGAT
541   GCGGTAAGTC  TGATTCGTAA  TATCTTCACT  ACTTTTAACT  CATCATCATC  CCCACCGGAA
601   AGACGAGTAT  ATAGCTGGAT  GGTTGTTGGT  GATTTCAATC  GTGCGCCGGC  TAATCTGGAA
661   GTTGCATTAA  GACAGGAGCC  CGCAGTGAGT  GAAAATACAA  TTATTATTGC  GCCAACAGAA
721   CCGACTCATC  GATCTGGTAA  TATTTTAGAT  TATGCAATTT  TACATGATGC  ACATTTACCA
781   CGTAGAGAAC  AGGCCCGTGA  ACGTATCGGT  GCAAGTTTAA  TGTTAAATCA  GTTACGCTCA
841   CAAATTACAT  CCGATCATTT  TCCTGTTAGT  TTTGTTCGTG  ATCGCTAA
```

→ Sequence correction for mutated amino acid residue (marked in amino acid sequence)

	Primers (uppercase = target-specific)	T_m	T_a
Fw	TTCAGCAGTAaATGAAAGTAAATG	58.0 °C	59.0 °C
Rv	GAACCTTGCAGATTCCAAG	60.0 °C	

CdtB-His_6

```
1     ATGGGCAGCA  GCCATCATCA  TCATCATCAC  AGCAGCGGCC  TGGTGCCGCG  CGGCAGCCAT
61    ATGGCTAGCA  TGACTGGTGG  ACAGCAAATG  GGTCGCGGAT  CCAACTTGAG  TGACTTCAAA
121   GTAGCAACTT  GGAATCTGCA  AGGTTCTTCA  GCAGTAGATG  AAAGTAAATG  GAATATTAAT
181   GTGCGCCAAT  TATTATCGGG  AGAACAAGGT  GCAGATATTT  TGATGGTACA  AGAAGCGGGT
241   TCCTTACCAA  GTTCGGCAGT  AAGAACCTCA  CGGGTAATTC  AACATGGGGG  AACGCCAATT
301   GAGGAATATA  CTTGGAATTT  AGGTACTCGT  TCCCGCCCAA  ATATGGTCTA  TATTTATTAT
361   TCTCGTTTAG  ATGTTGGGGC  AAACCGAGTG  AACTTAGCTA  TCGTGTCACG  CCGTCAAGCC
421   GATGAAGCTT  TTATCGTACA  TTCTGATTCT  TCTGTGCTTC  AATCTCGCCC  TGCAGTAGGT
481   ATCCGCATTG  GTACTGATGT  ATTTTTTACA  GTGCATGCTT  TGGCCACAGG  CGGTTCTGAT
541   GCGGTAAGTC  TGATTCGTAA  TATCTTCACT  ACTTTTAACT  CATCATCATC  CCCACCGGAA
601   AGACGAGTAT  ATAGCTGGAT  GGTTGTTGGT  GATTTCAATC  GTGCGCCGGC  TAATCTGGAA
661   GTTGCATTAA  GACAGGAGCC  CGCAGTGAGT  GAAAATACAA  TTATTATTGC  GCCAACAGAA
721   CCGACTCATC  GATCTGGTAA  TATTTTAGAT  TATGCAATTT  TACATGATGC  ACATTTACCA
781   CGTAGAGAAC  AGGCCCGTGA  ACGTATCGGT  GCAAGTTTAA  TGTTAAATCA  GTTACGCTCA
841   CAAATTACAT  CCGATCATTT  TCCTGTTAGT  TTTGTTCGTG  ATCGCTAA
```

IV. Amino acid sequences and theoretical parameters

HSP90*-His_6 TEV cleavage site

```
1     MKHHHHHHPM  SDYDIPTTEN  LYFQ|GAMAPE  EVHHGEEEVE  TFAFQAEIAQ  LMSLIINTFY
61    SNKEIFLREL  ISNASDALDK  IRYESLTDPS  KLDSGKELKI  DIIPNPQERT  LTLVDTGIGM
121   TKADLINNLG  TIAKSGTKAF  MEALQAGADI  SMIGQFGVGF  YSAYLVAEKV  VVITKHNDDE
181   QYAWESSAGG  SFTVRADHGE  PIGRGTKVIL  HLKEDQTEYL  EERRVKEVVK  KHSQFIGYPI
241   TLYLEKEREK  EISDKIKEKY  IDQEELNKTK  PIWTRNPDDI  TQEEYGEFYK  SLTNDWEDHL
301   AVKHFSVEGQ  LEFRALLFIP  RRAPFDLFEN  KKKKNNIKLY  VRRVFIMDSC  DELIPEYLNF
361   IRGVVDSEDL  PLNISREMLQ  QSKILKVIRK  NIVKKCLELF  SELAEDKENY  KKFYEAFSKN
421   LKLGIHEDST  NRRRLSELLR  YHTSQSGDEM  TSLSEYVSRM  KETQKSIYYI  TGESKEQVAN
481   SAFVERVRKR  GFEVVYMTEP  IDEYCVQQLK  EFDGKSLVSV  TKEGLELPED  EEEKKKMEES
541   KAKFENLCKL  MKEILDKKVE  KVTISNRLVS  SPCCIVTSTY  GWTANMERIM  KAQALRDNST
601   MGYMMAKKHL  EINPDHPIVE  TLRQKAEADK  NDKAVKDLVV  LLFETALLSS  GFSLEDPQTH
661   SNRIYRMIKL  GLG
```

Theoretical protein parameters

Mw: 77 736 Da

pI: 5.71

ε_{280}: 60 740 M^{-1} cm^{-1}

A_{280} = 1.280

HSP90*

```
1     GAMAPEEVHH  GEEEVETFAF  QAEIAQLMSL  IINTFYSNKE  IFLRELISNA  SDALDKIRYE
61    SLTDPSKLDS  GKELKIDIIP  NPQERTLTLV  DTGIGMTKAD  LINNLGTIAK  SGTKAFMEAL
121   QAGADISMIG  QFGVGFYSAY  LVAEKVVVIT  KHNDDEQYAW  ESSAGGSFTV  RADHGEPIGR
181   GTKVILHLKE  DQTEYLEERR  VKEVVKKHSQ  FIGYPITLYL  EKEREKEISD  KIKEKYIDQE
241   ELNKTKPIWT  RNPDDITQEE  YGEFYKSLTN  DWEDHLAVKH  FSVEGQLEFR  ALLFIPRRAP
301   FDLFENKKKK  NNIKLYVRRV  FIMDSCDELI  PEYLNFIRGV  VDSEDLPLNI  SREMLQQSKI
361   LKVIRKNIVK  KCLELFSELA  EDKENYKKFY  EAFSKNLKLG  IHEDSTNRRR  LSELLRYHTS
421   QSGDEMTSLS  EYVSRMKETQ  KSIYYITGES  KEQVANSAFV  ERVRKRGFEV  VYMTEPIDEY
481   CVQQLKEFDG  KSLVSVTKEG  LELPEDEEEK  KKMEESKAKF  ENLCKLMKEI  LDKKVEKVTI
541   SNRLVSSPCC  IVTSTYGWTA  NMERIMKAQA  LRDNSTMGYM  MAKKHLEINP  DHPIVETLRQ
601   KAEADKNDKA  VKDLVVLLFE  TALLSSGFSL  EDPQTHSNRI  YRMIKLGLG
```

Theoretical protein parameters

Mw: 74 734 Da

pI: 5.60

ε_{280}: 57 760 M^{-1} cm^{-1}

A_{280}: 1.294

HSP90-His_6 TEV cleavage site

```
1     MKHHHHHHPM  SDYDIPTTEN  LYFQ|GAMPEE  VHHGEEEVET  FAFQAEIAQL  MSLIINTFYS
61    NKEIFLRELI  SNASDALDKI  RYESLTDPSK  LDSGKELKID  IIPNPQERTL  TLVDTGIGMT
121   KADLINNLGT  IAKSGTKAFM  EALQAGADIS  MIGQFGVGFY  SAYLVAEKVV  VITKHNDDEQ
181   YAWESSAGGS  FTVRADHGEP  IGRGTKVILH  LKEDQTEYLE  ERRVKEVVKK  HSQFIGYPIT
241   LYLEKEREKE  ISDDEAEEEK  GEKEEEDKDD  EEKPKIEDVG  SDEEDDSGKK  KKKTKKIKEK
301   YIDQEELNKT  KPIWTRNPDD  ITQEEYGEFY  KSLTNDWEDH  LAVKHFSVEG  QLEFRALLFI
361   PRRAPFDLFE  NKKKKNNIKL  YVRRVFIMDS  CDELIPEYLN  FIRGVVDSED  LPLNISREML
421   QQSKILKVIR  KNIVKKCLEL  FSELAEDKEN  YKKFYEAFSK  NLKLGIHEDS  TNRRRLSELL
481   RYHTSQSGDE  MTSLSEYVSR  MKETQKSIYY  ITGESKEQVA  NSAFVERVRK  RGFEVVYMTE
541   PIDEYCVQQL  KEFDGKSLVS  VTKEGLELPE  DEEEKKKMEE  SKAKFENLCK  LMKEILDKKV
601   EKVTISNRLV  SSPCCIVTST  YGWTANMERI  MKAQALRDNS  TMGYMMAKKH  LEINPDHPIV
661   ETLRQKAEAD  KNDKAVKDLV  VLLFETALLS  SGFSLEDPQT  HSNRIYRMIK  LGLGIDEDEV
721   AAEEPNAAVP  DEIPPLEGDE  DASRMEEVD
```

Theoretical protein parameters

Mw: 86 276

pI: 5.05

ε_{280}: 60 740 M^{-1} cm^{-1}

A_{280}: 1.420

HSP90

```
1     GAMPEEVHHG  EEEVETFAFQ  AEIAQLMSLI  INTFYSNKEI  FLRELISNAS  DALDKIRYES
61    LTDPSKLDSG  KELKIDIIPN  PQERTLTLVD  TGIGMTKADL  INNLGTIAKS  GTKAFMEALQ
121   AGADISMIGQ  FGVGFYSAYL  VAEKVVVITK  HNDDEQYAWE  SSAGGSFTVR  ADHGEPIGRG
181   TKVILHLKED  QTEYLEERRV  KEVVKKHSQF  IGYPITLYLE  KEREKEISDD  EAEEEKGEKE
241   EEDKDDEEKP  KIEDVGSDEE  DDSGKKKKKT  KKIKEKYIDQ  EELNKTKPIW  TRNPDDITQE
301   EYGEFYKSLT  NDWEDHLAVK  HFSVEGQLEF  RALLFIPRRA  PFDLFENKKK  KNNIKLYVRR
361   VFIMDSCDEL  IPEYLNFIRG  VVDSEDLPLN  ISREMLQQSK  ILKVIRKNIV  KKCLELFSEL
421   AEDKENYKKF  YEAFSKNLKL  GIHEDSTNRR  RLSELLRYHT  SQSGDEMTSL  SEYVSRMKET
481   QKSIYYITGE  SKEQVANSAF  VERVRKRGFE  VVYMTEPIDE  YCVQQLKEFD  GKSLVSVTKE
541   GLELPEDEEE  KKKMEESKAK  FENLCKLMKE  ILDKKVEKVT  ISNRLVSSPC  CIVTSTYGWT
601   ANMERIMKAQ  ALRDNSTMGY  MMAKKHLEIN  PDHPIVETLR  QKAEADKNDK  AVKDLVVLLF
661   ETALLSSGFS  LEDPQTHSNR  IYRMIKLGLG  IDEDEVAAEE  PNAAVPDEIP  PLEGDEDASR
721   MEEVD
```

Theoretical protein parameters

Mw: 83 277 Da

pI: 4.98

ε_{280}: 57 760 M^{-1} cm^{-1}

A_{280}: 1.441

CTA1-His_6

```
1     MNDDKLYRAD  SRPPDEIKQS  GGLMPRGQSE  YFDRGTQMNI  NLYDHARGTQ  TGFVRHDDGY
61    VSTSISLRSA  HLVGQTILSG  HSTYYIYVIA  TAPNMFNVND  VLGAYSPHPD  EQEVSALGGI
121   PYSQIYGWYR  VHFGVLDEQL  HRNRGYRDRY  YSNLDIAPAA  DGYGLAGFPP  EHRAWREEPW
181   IHHAPPGCGN  SHHHHHH
```

Theoretical protein parameters

Mw: 22 358 Da

pI: 6.21

ε_{280}: 38 850 M^{-1} cm^{-1}

A_{280}: 0.575

CTA1

```
1     MNDDKLYRAD  SRPPDEIKQS  GGLMPRGQSE  YFDRGTQMNI  NLYDHARGTQ  TGFVRHDDGY
61    VSTSISLRSA  HLVGQTILSG  HSTYYIYVIA  TAPNMFNVND  VLGAYSPHPD  EQEVSALGGI
121   PYSQIYGWYR  VHFGVLDEQL  HRNRGYRDRY  YSNLDIAPAA  DGYGLAGFPP  EHRAWREEPW
181   IHHAPPGCGN  S
```

Theoretical protein parameters

Mw: 21 535 Da

pI: 5.88

ε_{280}: 38 850 M^{-1} cm^{-1}

A_{280}: 0.554

CdtB-His_6 Thrombin cleavage sites

```
1     MGSSHHHHHH  SSGLVPR|GSH  MASMTGGQQM  GR|GSNLSDFK  VATWNLQGSS  AVNESKWNIN
61    VRQLLSGEQG  ADILMVQEAG  SLPSSAVRTS  RVIQHGGTPI  EEYTWNLGTR  SRPNMVYIYY
121   SRLDVGANRV  NLAIVSRRQA  DEAFIVHSDS  SVLQSRPAVG  IRIGTDVFFT  VHALATGGSD
181   AVSLIRNIFT  TFNSSSSPPE  RRVYSWMVVG  DFNRAPANLE  VALRQEPAVS  ENTIIIAPTE
241   PTHRSGNILD  YAILHDAHLP  RREQARERIG  ASLMLNQLRS  QITSDHFPVS  FVRDR
```

Theoretical protein parameters

Mw: 32 500 Da

pI: 9.50

ε_{280}: 30 940 M^{-1} cm^{-1}

A_{280}: 1.050

CdtB (cleavage 1)

```
1     GSHMASMTGG  QQMGRGSNLS  DFKVATWNLQ  GSSAVNESKW  NINVRQLLSG  EQGADILMVQ
61    EAGSLPSSAV  RTSRVIQHGG  TPIEEYTWNL  GTRSRPNMVY  IYYSRLDVGA  NRVNLAIVSR
121   RQADEAFIVH  SDSSVLQSRP  AVGIRIGTDV  FFTVHALATG  GSDAVSLIRN  IFTTFNSSSS
181   PPERRVYSWM  VVGDFNRAPA  NLEVALRQEP  AVSENTIIIA  PTEPTHRSGN  ILDYAILHDA
241   HLPRREQARE  RIGASLMLNQ  LRSQITSDHF  PVSFVRDR
```

Theoretical protein parameters

Mw: 30 618 Da

pI: 9.15

ε_{280}: 30 940 M^{-1} cm^{-1}

A_{280}: 0.989

CdtB (cleavage 2)

```
1     GSNLSDFKVA  TWNLQGSSAV  NESKWNINVR  QLLSGEQGAD  ILMVQEAGSL  PSSAVRTSRV
61    IQHGGTPIEE  YTWNLGTRSR  PNMVYIYYSR  LDVGANRVNL  AIVSRRQADE  AFIVHSDSSV
121   LQSRPAVGIR  IGTDVFFTVH  ALATGGSDAV  SLIRNIFTTF  NSSSSPPERR  VYSWMVVGDF
181   NRAPANLEVA  LRQEPAVSEN  TIIIAPTEPT  HRSGNILDYA  ILHDAHLPRR  EQARERIGAS
241   LMLNQLRSQI  TSDHFPVSFV  RDR
```

Theoretical protein parameters

Mw: 29 101 Da

pI: 8.32

ε_{280}: 30 940 M^{-1} cm^{-1}

A_{280}: 0.940

References

[1] U. Püntener *et al.* (2012) "Long-term impact of systemic bacterial infection on the cerebral vasculature and microglia", *J Neuroinflammation*, vol. 9, pp. 146-146

[2] L. Leibovici (2013) "Long-term consequences of severe infections", *Clin Microbiol Infect*, vol. 19, no. 6, pp. 510-512

[3] B. Albiger *et al.* (2007) "Role of the innate immune system in host defence against bacterial infections: focus on the Toll-like receptors", *J Intern Med*, vol. 261, no. 6, pp. 511-528

[4] J. E. Galán (2005) "Bacterial toxins and the immune system: show me the in vivo targets", *J Exp Med*, vol. 201, no. 3, pp. 321-323

[5] S. Leekha *et al.* (2011) "General principles of antimicrobial therapy", *Mayo Clin Proc*, vol. 86, no. 2, pp. 156-167

[6] M. A. Kohanski *et al.* (2010) "How antibiotics kill bacteria: from targets to networks", *Nat Rev Microbiol*, vol. 8, no. 6, pp. 423-435

[7] R. Laxminarayan *et al.* (2013) "Antibiotic resistance – the need for global solutions", *Lancet Infect Dis*, vol. 13, no. 12, pp. 1057-1098

[8] A. K. Sharma *et al.* (2017) "Bacterial virulence factors: secreted for survival", *Indian J Microbiol*, vol. 57, no. 1, pp. 1-10

[9] J. Pizarro-Cerdá and P. Cossart (2006) "Bacterial adhesion and entry into host cells", *Cell*, vol. 124, no. 4, pp. 715-727

[10] S. J. Krebs and R. K. Taylor (2011) "Protection and attachment of Vibrio cholerae mediated by the toxin-coregulated pilus in the infant mouse model", *J Bacteriol*, vol. 193, no. 19, pp. 5260-5270

[11] E. Wong *et al.* (2012) "The Vibrio cholerae colonization factor GbpA possesses a modular structure that governs binding to different host surfaces", *PLoS Pathog*, vol. 8, no. 1, p. e1002373

[12] T. J. Kirn *et al.* (2005) "A colonization factor links Vibrio cholerae environmental survival and human infection", *Nature*, vol. 438, no. 7069, pp. 863-866

[13] R. Bhowmick *et al.* (2008) "Intestinal adherence of Vibrio cholerae involves a coordinated interaction between colonization factor GbpA and mucin", *Infect Immun*, vol. 76, no. 11, pp. 4968-4977

[14] B. B. Finlay and P. Cossart (1997) "Exploitation of mammalian host cell functions by bacterial pathogens", *Science*, vol. 276, no. 5313, pp. 722-724

[15] T. W. Mak and M. E. Saunders (2006), "Immunity to pathogens," in *The Immune Response*, Burlington: Academic Press, pp. 642-670.

[16] J.-M. Cavaillon (2017) "Exotoxins and endotoxins: inducers of inflammatory cytokines", *Toxicon*, vol. 149, pp. 45-53

[17] X. Wang and P. J. Quinn (2010) "Endotoxins: lipopolysaccharides of gram-negative bacteria", *Subcell Biochem*, vol. 53, pp. 3-25

[18] O. Odumosu *et al.* (2010) "AB Toxins: a paradigm switch from deadly to desirable", *Toxins*, vol. 2, no. 7, pp. 1612-1645

[19] V. Luigi *et al.* (2010) "Environmental reservoirs of Vibrio cholerae and their role in cholera", *Environ Microbiol Rep*, vol. 2, no. 1, pp. 27-33

[20] S. Almagro-Moreno and R. K. Taylor (2013) "Cholera: environmental reservoirs and impact on disease transmission", *Microbiol Spectr*, vol. 1, no. 2, pp. 1-10

[21] World Health Organization (August 2017) "Cholera vaccines: WHO position paper" in "Weekly epidemiological record", vol. 92

[22] C. J. O'Neal *et al.* (2004) "Crystal structures of an intrinsically active cholera toxin mutant yield insight into the toxin activation mechanism", *Biochemistry*, vol. 43, no. 13, pp. 3772-3782
[23] I. Majoul *et al.* (1997) "Reduction of protein disulfide bonds in an oxidizing environment: the disulfide bridge of cholera toxin A-subunit is reduced in the endoplasmic reticulum", *FEBS Lett*, vol. 401, no. 2-3, pp. 104-108
[24] P. A. Orlandi (1997) "Protein-disulfide isomerase-mediated reduction of the A subunit of cholera toxin in a human intestinal cell line", *J Biol Chem*, vol. 272, no. 7, pp. 4591-4599
[25] A. H. Pande *et al.* (2007) "Conformational instability of the cholera toxin A1 polypeptide", *J Mol Biol*, vol. 374, no. 4, pp. 1114-1128
[26] M. Taylor *et al.* (2011) "Protein-disulfide isomerase displaces the cholera toxin A1 subunit from the holotoxin without unfolding the A1 subunit", *J Biol Chem*, vol. 286, no. 25, pp. 22090-22100
[27] M. L. Torgersen *et al.* (2001) "Internalization of cholera toxin by different endocytic mechanisms", *J Cell Sci*, vol. 114, no. 20, pp. 3737-3747
[28] G. H. Hansen *et al.* (2005) "Cholera toxin entry into pig enterocytes occurs via a lipid raft- and clathrin-dependent mechanism", *Biochemistry*, vol. 44, no. 3, pp. 873-882
[29] Y. Feng *et al.* (2004) "Retrograde transport of cholera toxin from the plasma membrane to the endoplasmic reticulum requires the trans-Golgi network but not the Golgi apparatus in Exo2-treated cells", *EMBO Rep*, vol. 5, no. 6, pp. 596-601
[30] K. Chaudhuri and S. N. Chatterjee (2008), "Cholera toxin (CT): Entry and retrograde trafficking into the epithelial cell," in *Cholera toxins*, pp. 185-197.
[31] K. Sandvig *et al.* (2013) "Retrograde transport of protein toxins through the Golgi apparatus", *Histochem Cell Biol*, vol. 140, no. 3, pp. 317-326
[32] C. Rodighiero *et al.* (2002) "Role of ubiquitination in retro-translocation of cholera toxin and escape of cytosolic degradation", *EMBO Rep*, vol. 3, no. 12, pp. 1222-1227
[33] M. Taylor *et al.* (2010) "Hsp90 is required for transfer of the cholera toxin A1 subunit from the endoplasmic reticulum to the cytosol", *J Biol Chem*, vol. 285, no. 41, pp. 31261-31267
[34] H. Burress *et al.* (2014) "Co- and post-translocation roles for HSP90 in cholera intoxication", *J Biol Chem*, vol. 289, no. 48, pp. 33644-33654
[35] C. J. O'Neal *et al.* (2005) "Structural basis for the activation of cholera toxin by human ARF6-GTP", *Science*, vol. 309, no. 5737, pp. 1093-1096
[36] M. Field (2003) "Intestinal ion transport and the pathophysiology of diarrhea", *J Clin Invest*, vol. 111, no. 7, pp. 931-943
[37] L. Whitesell and S. L. Lindquist (2005) "HSP90 and the chaperoning of cancer", *Nat Rev Cancer*, vol. 5, no. 10, pp. 761-772
[38] J. Wang *et al.* (2013) "High expression of heat shock protein 90 is associated with tumor aggressiveness and poor prognosis in patients with advanced gastric cancer", *PLoS One*, vol. 8, no. 4, p. e62876
[39] X. Liu *et al.* (2016) "HSP90 inhibits apoptosis and promotes growth by regulating HIF-1α abundance in hepatocellular carcinoma", *Int J Mol Med*, vol. 37, no. 3, pp. 825-835
[40] S. Yamada *et al.* (2003) "A hydrophobic segment within the C-terminal domain is essential for both client-binding and dimer formation of the HSP90-family molecular chaperone", *Eur J Biochem*, vol. 270, no. 1, pp. 146-154
[41] P. Hawle *et al.* (2006) "The middle domain of Hsp90 acts as a discriminator between different types of client proteins", *Molecular and Cellular Biology*, vol. 26, no. 22, pp. 8385-8395

[42] B. T. Scroggins *et al.* (2007) "An acetylation site in the middle domain of Hsp90 regulates chaperone function", *Molecular cell*, vol. 25, no. 1, pp. 151-159
[43] M. Jahn *et al.* (2014) "The charged linker of the molecular chaperone Hsp90 modulates domain contacts and biological function", *Proc Natl Acad Sci USA*, vol. 111, no. 50, pp. 17881-17886
[44] A. Dixit and G. M Verkhivker (2012) "Probing molecular mechanisms of the Hsp90 chaperone: Biophysical modeling identifies key regulators of functional dynamics". p. e37605.
[45] B. Chen *et al.* (2005) "The HSP90 family of genes in the human genome: insights into their divergence and evolution", *Genomics*, vol. 86, no. 6, pp. 627-637
[46] M. M. U. Ali *et al.* (2006) "Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex", *Nature*, Article vol. 440, no. 7087, pp. 1013-1017
[47] Y. Minami *et al.* (1994) "The carboxy-terminal region of mammalian HSP90 is required for its dimerization and function in vivo", *Mol Cell Biol*, vol. 14, no. 2, pp. 1459-1464
[48] A. Dixit and G. M. Verkhivker (2012) "Probing molecular mechanisms of the Hsp90 chaperone: biophysical modeling identifies key regulators of functional dynamics", *PLoS One*, vol. 7, no. 5, p. e37605
[49] J. Fontana *et al.* (2002) "Domain mapping studies reveal that the M domain of hsp90 serves as a molecular scaffold to regulate Akt-dependent phosphorylation of endothelial nitric oxide synthase and NO release", *Circ Res*, vol. 90, no. 8, pp. 866-873
[50] J. Cheung-Flynn *et al.* (2003) "C-terminal sequences outside the tetratricopeptide repeat domain of FKBP51 and FKBP52 cause differential binding to Hsp90", *J Biol Chem*, vol. 278, no. 19, pp. 17388-17394
[51] R. L. Matts *et al.* (2011) "Elucidation of the Hsp90 C-terminal inhibitor binding site", *ACS Chem Biol*, vol. 6, no. 8, pp. 800-807
[52] A. J. Ramsey *et al.* (2009) "C-terminal sequences of hsp70 and hsp90 as non-specific anchors for TPR proteins", *Biochem J*, vol. 423, no. 3, pp. 411-419
[53] O. Hainzl *et al.* (2009) "The charged linker region is an important regulator of Hsp90 function", *J Biol Chem*, vol. 284, no. 34, pp. 22559-22567
[54] L. Guerra *et al.* (2011) "The biology of the cytolethal distending toxins", *Toxins*, vol. 3, no. 3, pp. 172-190
[55] W. M. Johnson and H. Lior (1987) "Response of Chinese hamster ovary cells to a cytolethal distending toxin (CDT) of Escherichia coli and possible misinterpretation as heat-labile (LT) enterotoxin", *FEMS Microbiol*, vol. 43, no. 1, pp. 19-23
[56] W. M. Johnson and H. Lior (1987) "Production of Shiga toxin and a cytolethal distending toxin (CLDT) by serogroups of Shigella spp", *FEMS Microbiol*, vol. 48, no. 1-2, pp. 235-238
[57] W. M. Johnson and H. Lior (1988) "A new heat-labile cytolethal distending toxin (CLDT) produced by Campylobacter spp", *Microb Pathog*, vol. 4, no. 2, pp. 115-126
[58] M. Ohara and M. Sugai (2005) "Cytolethal distending toxin and its implication in periodontal diseases", *J Periodontal Res*, vol. 47, no. 1, pp. 18-24
[59] M. D. Scuron *et al.* (2016) "The cytolethal distending toxin contributes to microbial virulence and disease pathogenesis by acting as a tri-perditious toxin", *Front Cell Infect Microbiol*, vol. 6, pp. 9-12
[60] T. Faïs *et al.* (2016) "Impact of CDT toxin on human diseases", *Toxins*, vol. 8, no. 7, pp. 5-7
[61] C. A. Elwell and L. A. Dreyfus (2000) "DNase I homologous residues in CdtB are critical for cytolethal distending toxin-mediated cell cycle arrest", *Mol Microbiol*, vol. 37, no. 4, pp. 952-963

[62] M. Lara-Tejero and J. E. Galan (2000) "A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein", *Science*, vol. 290, no. 5490, pp. 354-357

[63] D. Nešić *et al.* (2004) "Assembly and function of a bacterial genotoxin", *Nature*, vol. 429, pp. 429-433

[64] T. Frisan *et al.* (2003) "The Haemophilus ducreyi cytolethal distending toxin induces DNA double-strand breaks and promotes ATM-dependent activation of RhoA", *Cell Microbiol*, vol. 5, no. 10, pp. 695-707

[65] E. Bezine *et al.* (2014) "The cytolethal distending toxin effects on mammalian cells: a DNA damage Perspective", *Cells*, vol. 3, no. 2, pp. 592-615

[66] S. Y. Peres *et al.* (1997) "A new cytolethal distending toxin (CDT) from Escherichia coli producing CNF2 blocks HeLa cell division in G2/M phase", *Mol Microbiol*, vol. 24, no. 5, pp. 1095-1107

[67] C. A. Whitehouse *et al.* (1998) "Campylobacter jejuni cytolethal distending toxin causes a G2-phase cell cycle block", *Infect Immun*, vol. 66, no. 5, pp. 1934-40

[68] M. Ohguchi *et al.* (1998) "Actinobacillus actinomycetemcomitans toxin induces both cell cycle arrest in the G2/M phase and apoptosis", *Infect Immun*, vol. 66, no. 12, pp. 5980-5987

[69] R. N. Jinadasa *et al.* (2011) "Cytolethal distending toxin: a conserved bacterial genotoxin that blocks cell cycle progression, leading to apoptosis of a broad range of mammalian cell lineages", *Microbiology*, vol. 157, no. Pt 7, pp. 1851-1875

[70] L. A. McSweeney and L. A. Dreyfus (2005) "Carbohydrate-binding specificity of the Escherichia coli cytolethal distending toxin CdtA-II and CdtC-II subunits", *Infect Immun*, vol. 73, no. 4, pp. 2051-2060

[71] K. Boesze-Battaglia *et al.* (2006) "Cholesterol-rich membrane microdomains mediate cell cycle arrest induced by Actinobacillus actinomycetemcomitans cytolethal-distending toxin", *Cell Microbiol*, vol. 8, no. 5, pp. 823-836

[72] K. Mise *et al.* (2005) "Involvement of ganglioside GM3 in G(2)/M cell cycle arrest of human monocytic cells induced by Actinobacillus actinomycetemcomitans cytolethal distending toxin", *Infect Immun*, vol. 73, no. 8, pp. 4846-4852

[73] K. Boesze-Battaglia *et al.* (2015) "The Aggregatibacter actinomycetemcomitans cytolethal distending toxin active subunit CdtB contains a cholesterol recognition sequence required for toxin binding and subunit internalization", *Infect Immun*, vol. 83, no. 10, pp. 4042-4055

[74] K. Deng *et al.* (2001) "Investigation of the interaction among the components of the cytolethal distending toxin of Haemophilus ducreyi", *Biochem Biophys Res Commun*, vol. 285, no. 3, pp. 609-15

[75] A. Frisk *et al.* (2001) "The role of different protein components from the Haemophilus ducreyi cytolethal distending toxin in the generation of cell toxicity", *Microb Pathog*, vol. 30, no. 6, pp. 313-324

[76] M. Damek-Poprawa *et al.* (2012) "Localization of Aggregatibacter actinomycetemcomitans cytolethal distending toxin subunits during intoxication of live cells", *Infect Immun*, vol. 80, no. 8, pp. 2761-70

[77] S. D. Dixon *et al.* (2015) "Distinct roles for CdtA and CdtC during intoxication by cytolethal distending toxins", *PLoS One*, vol. 10, no. 11, p. e0143977

[78] A. Gargi *et al.* (2013) "Cellular interactions of the cytolethal distending toxins from Escherichia coli and Haemophilus ducreyi", *J Biol Chem*, vol. 288, no. 11, pp. 7492-7505

[79] A. Eshraghi *et al.* (2014) "Cytolethal distending toxins require components of the ER-associated degradation pathway for host cell entry", *PLoS Pathog*, vol. 10, no. 7, p. e1004295
[80] S. Nishikubo *et al.* (2003) "An N-terminal segment of the active component of the bacterial genotoxin cytolethal distending toxin B (CDTB) directs CDTB into the nucleus", *J Biol Chem*, vol. 278, no. 50, pp. 50671-81
[81] X. Mao and J. M. DiRienzo (2002) "Functional studies of the recombinant subunits of a cytolethal distending holotoxin", *Cell Microbiol*, vol. 4, no. 4, pp. 245-255
[82] J. A. Bornhorst and J. J. Falke (2000) "Purification of proteins using polyhistidine affinity tags (16) ", *Methods Enzymol*, vol. 326, pp. 245-254
[83] M. Gallwitz *et al.* (2012) "The extended cleavage specificity of human thrombin", *PLoS One*, vol. 7, no. 2, p. e31756
[84] T. Sun *et al.* (2004) "A study of the separation principle in size exclusion chromatography", *Macromolecules*, vol. 37, no. 11, pp. 4304-4312
[85] H. Hedlund (2006) "Desalting and buffer exchange of proteins using size-exclusion chromatography", *CSH Protoc*, vol. 2006, no. 1,
[86] B. J. Smith (1984) "SDS polyacrylamide gel electrophoresis of proteins", *Methods Mol Biol*, vol. 1, pp. 41-55
[87] K. Teter *et al.* (2006) "The cholera toxin A13 subdomain is essential for interaction with ADP-ribosylation factor 6 and full toxic activity but is not required for translocation from the endoplasmic reticulum to the cytosol", *Infect Immun*, vol. 74, no. 4, pp. 2259-2267
[88] J. S. Hontz *et al.* (2006) "Crystallization of Escherichia coli CdtB, the biologically active subunit of cytolethal distending toxin", *Acta Crystallogr Sect F Struct Biol Cryst Commun*, vol. 62, no. 3, pp. 192-195
[89] X. Wu *et al.* (2009) "A novel method for high-level production of TEV protease by superfolder GFP tag", *J Biomed Biotechnol*, vol. 2009, pp. 1-8