Catch me if you can

Interaction studies of AB_5 toxins with PDI by small-angle scattering

Natalia Mojica Cortés

Master Thesis in Biochemistry

60 credits

Department of Chemistry
Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

06/2021

2021

Catch me if you can- Interaction studies of AB_5 toxins with PDI by small-angle scattering

Natalia Mojica Cortés

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Acknowledgements

First, I would like to express my gratitude to my main supervisor, Professor Ute Krengel. You gave me the opportunity to join your lab initially as a bachelor's student and later as a master's. Thanks for all the time you dedicated to my supervision, for your unconditional support and for promoting and organizing all the social gatherings that made my stay in Oslo and work in the lab very enjoyable. You have inspired me to continue in the path of structural biology and I simply would not be here without you. Thank you very much!

Thanks to Henrik, who trained me with protein expression, purification and SAXS. You have always been available to answer my questions and I truly appreciate that. Also, special thanks to Joel, who introduced me to the project and gave me feedback valuable feedback. I would also like to express my gratitude to Gabriele for all the nice scientific discussions, the insight you always shared and the patience when I asked so many questions. Thanks to Hedda for being always so positive and fun. Thanks to Clara for being a nice study partner and such a good baker. Takk til Abelone, Helene, Clara, Hedda og Anna for å snakke Norsk med meg. Det var alltid gøy! Thank you Tamjid for helping me with crystallization and for always being kind and inviting me for eating your delicious dumplings! Thanks to Anna for being an interested student and joining my project. I am also thankful to Dipankar, for providing insight and teaching me about Indian culture.

Thanks to Nikolina and Dario for allowing me to perform cloning experiments in their lab. Thanks to Marta Hammerstad for always being so nice whenever I needed to sonicate or prepare samples under inert atmosphere.

Thanks to Daniela Braun for the friendship and nice desserts! I always enjoyed our late conversations in the kitchen. Thanks to all my other friends: Masaki, Huan, Bisma, Valentine, Meg and (Colombian) Daniela. It was always fun to cook together and learn about your cultures.

Gracias a mis padres y mi hermano Nico. Ustedes siempre confiaron en mi y me dieron todo el apoyo emocional a lo largo de estos dos años en Noruega. Soy muy afortunada de tenerlos como mi familia. Sin su apoyo y amor no estaría aquí.

Finally, I would like to express my gratitude to Oliver for his unconditional support throughout the master's. Thanks for teaching me Matlab, helping me to overcome the stress and always listening to me. You are simply amazing!

Natalia,

May 2021

Abstract

Diarrheal diseases are one of the leading causes of death in the world. Almost half of the diarrhea cases each year are caused by the bacterial pathogens *Vibrio cholerae* and enterotoxigenic *Escherichia coli* (ETEC). After reaching the host intestine via contaminated food or water, these pathogens release two closely related toxins responsible for severe diarrhea: the cholera toxin (CT) and the heat-labile enterotoxin (LT), respectively. Both toxins have a similar AB_5 structure, characterized by one catalytically active A1 subunit (CTA1, LTA1), an A2 linker and a pentamer of cell-binding B subunits (CTB, LTB). To become active, the A1 subunit must be separated from the rest of the toxin. This is known to be mediated by the protein disulfide isomerase (PDI) in the endoplasmic reticulum of the host and has been shown to be a physical rather than an enzymatic process. Although the interaction between PDI and CT/LT has been investigated by various biophysical methods, the mechanism of toxin disassembly is not yet completely understood at the structural level. Better insight into this important step of the intoxication process could be exploited for development of potential therapeutics. The aim of this thesis was to gain structural insight into the interaction of CT with PDI by Small-Angle X-ray Scattering (SAXS). This technique was chosen as it is suitable to study highly dynamic systems, such as the CT/LT-PDI. Since the combination of SAXS with Small-Angle Neutron Scattering (SANS) would be useful for the characterization of the CT-PDI interaction, preliminary deuteration work for SANS was also pursued in this thesis.

This work focused initially on producing milli-gram amounts of pure proteins for SAXS and SANS. We developed a novel, simple, and cost-efficient protocol to produce CT and the porcine heat-labile enterotoxin (pLT) in *Vibrio natriegens.* The new protocol increased the yield of CT by at least 10-fold compared to expression in *E. coli*. PDI was successfully produced following an already established protocol that we slightly modified to improve protein purity. Moreover, a preliminary protocol for production of deuterated PDI was developed, nevertheless, further optimization is required as the yields were too low for SANS studies. Formation of a complex between CT and PDI was studied by size-exclusion chromatography (SEC) and SAXS under different temperatures, pH, reducing agents and specific treatments to inhibit PDI and potentially stabilize its interaction with CT, however, no interaction was observed in any of the conditions tested. Despite not obtaining a CT-PDI complex, this work made important progress in establishing efficient protocols for protein production, identified conditions were the CT-PDI interaction does not take place, and discussed the possible reasons for the lack of interaction, paving the way for future studies.

Abbreviations

Amino-terminus	N-terminus
Amp	Ampicillin
CAM	Chloramphenicol
CD	Circular Dichorism
Cryo-EM	Cryogenic electron microscopy
CT	Cholera toxin (classical)
CTA1	CT A-subunit fragment 1
CTA2	CT A-subunit fragment 2
CTB	CT B-subunit
Da	Dalton
d_{max}	Maximum diameter
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDC	(1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride
EDTA	Ethylenediamine-tetraacetate
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
FTIR	Isotope-edited Fourier Transform Infrared Spectroscopy
GAC	Galactose affinity chromatography
GSH	Glutathione
His-tag	Polyhistidine-tag (6X)
hLT	Human heat-labile enterotoxin
HRP	Horseradish peroxidase
IEX	Ion-exchange chromatography
IMAC	Immobilized metal affinity chroamtography
IPTG	Isopropyl β-d-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K_d	Dissociation constant
LB	Luria-Bertani medium
LB+v2 salts	Luria-Bertani medium supplemented with NaCl, KCl and MgCl2
mAU	Milli absorption unit
MES	2-(N-monopholino)ethanesulfonate
MQ-H2O	Milli Q-water
MS	Mass spectrometry
MST	Microscale thermophoresis
OD_{600}	Optical density measured at a wavelength of 600 nm
PBS	Phosphate-buffered saline
PDB	Protein Data Bank
PDI	Protein disulfide isomerase

pI	Isoelectric point
pLT	Porcine heat-labile enterotoxin
pLTA	pLT A-subunit fragment 1
pLTB	pLT B subunit
PMSF	Phenylmethylsulfonyl fluoride
Q3R	Quercetin-3-Rutinoside
rcf	Relative centrifugal force
R_g	Radius of Gyration
rpm	Revolutions per minute
SANS	Small-angle Neutron scattering
SAS	Small-angle scattering
SAXS	Small-angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SPR	Surface plasmon resonance
TCEP	Tris(2-carboxyethyl)phosphine
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
Vmax	*Vibrio natriegens*
ε	Mola extinction coefficient

Table of Contents

1 Introduction

1.1 Infectious diseases

Until the end of the twentieth century, infectious diseases were the main cause of premature death and disability[1]. Although the mortality levels have decreased considerably over the years due to increased vaccination and availability of antibiotics, infectious diseases continue to be a major global burden[2]. A good example is the ongoing coronavirus pandemic (COVID-19), a severe acute respiratory disease that has not only taken millions of lives worldwide but has had a serious economic and social impact[3]. Prior to COVID-19, infectious diseases were already among the leading causes of death, especially in low-income countries (Figure 1.)[4]. These include infections of the lower respiratory tract as well as diarrheal diseases (Figure 1.B).

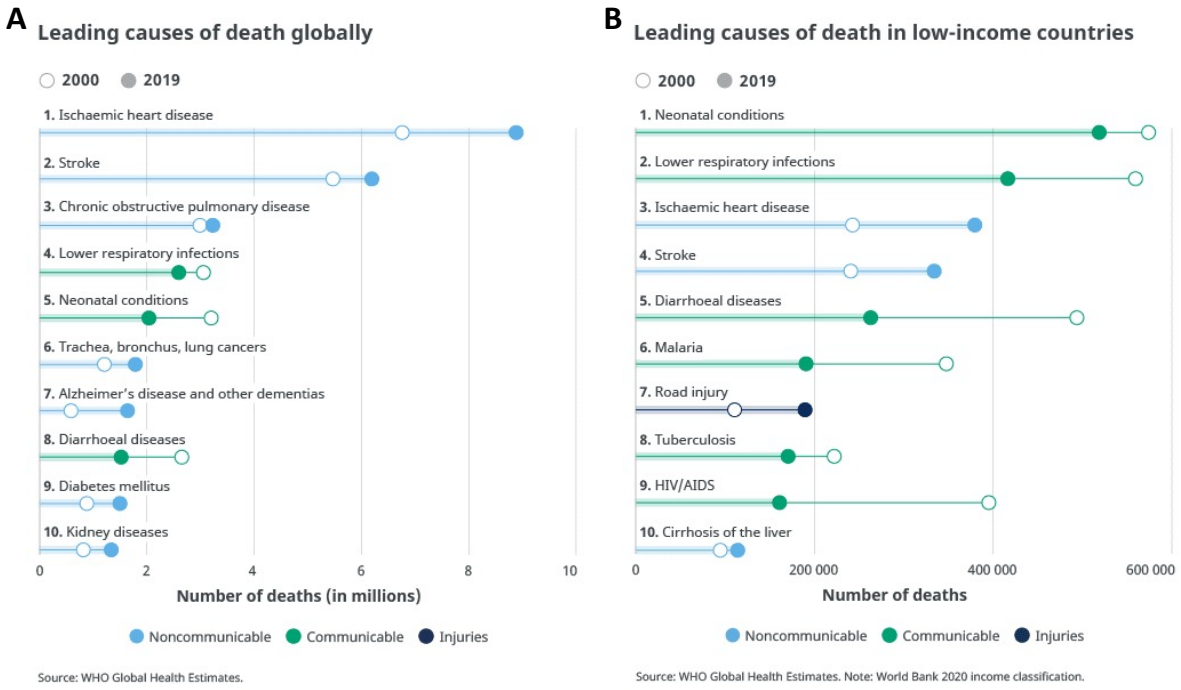


Figure 1. Top ten leading causes of death globally (a) and in low-income countries (b). World Health Organization, 2019. Communicable disease: Illness caused by an infectious agent. It can be passed from one person to another. Non-communicable disease: Disease that is non-transmissible between people.

1.1.1 Diarrheal diseases: cholera and traveler's diarrhea

Diarrheal diseases are the fifth leading cause of death in low-income countries and among the ten leading killers globally[4]. They are especially deadly among children under five years old, killing more than half a million children every year[5].

Diarrhea can be produced by infectious microorganisms including bacteria, viruses and parasites, which spread through contaminated food or drinking water or from person to

person due to poor sanitary conditions[5]. One of the most severe diarrheal diseases is cholera, an infection caused by the bacterium *Vibrio cholerae*. A similar, but less severe disease is caused by enterotoxigenic *Escherichia coli* (ETEC). These two infections are responsible for the majority of diarrhea episodes in the world[6].

Cholera

Cholera is a severe diarrheal disease produced by the Gram-negative bacterium *V. cholerae*. This bacterium is naturally present in aqueous environment, usually attached to crustaceans and plankton[7]. Infections occur by ingestion of water or food contaminated with *V. cholerae*[8]. It is estimated that between 1 to 4 million cases of cholera occur each year, leading to hundred thousand deaths[9]. The Yemen cholera outbreak in 2016 increased the number of diarrhea-related deaths to the point that it became the third most common cause of death worldwide that year[10]. If not treated quickly, cholera can lead to death in a matter of hours. Rehydration therapy is the most effective treatment, and in the most severe cases, antibiotics are also provided[8]. Cholera vaccines are available; however, they offer only short-term protection, and a booster dose is usually required[11].

V. cholerae can be classified by serogroup[1], serotype[2] (Ogawa, Inaba, Hikojima) and biotype[3](classical and El Tor). Out of more than 200 serogroups, only two of them have been associated with epidemics: O1 and O139[12,13]. The O1 serogroup contains the Classical and El Tor biotypes, which have been responsible for the majority of outbreaks[13]. One of the main differences between these biotypes is that they produce different variants of the cholera toxin (CT), the main virulence factor responsible for the severe diarrhea[14]. Although since 1960s El Tor strains became the major causative agent of cholera, a hybrid biotype took over in the beginning of the 2000s, which has characteristics of El Tor but produces a classical CT[15,16]. The O139 serogroup was discovered in 1992 as the cause of an epidemic in South East Asia, but only few cases have been reported after its initial outbreak[17].

ETEC: Traveler's diarrhea

ETEC, another type of Gram-negative bacterium, is also a major cause of diarrhea in young children in developing countries as well as travelers to these areas[18]. It is transmitted in the same way as *V. cholerae* and causes similar symptoms, although ETEC diarrhea is usually milder[19]. ETEC strains are characterized by their ability to produce heat-labile (LT) and/or heat-stable toxins (ST), which stimulate the massive secretion of fluids from epithelial cells that eventually lead to diarrhea[20]. While STs are small peptides stable up to 100 °C[21], LT is a protein that can be inactivated at 60 °C [22]. The high toxicity and low immunogenicity of ST

[1] Serogroup: group of bacteria that share a common cell surface antigen.

[2] Serotype: group of bacteria that exhibit same type and number of antigens.

[3] Biotype: organisms that share a specific variation in biological behavior.

makes them less attractive vaccine antigens[23,24]. In contrast, heat-labile toxins have drawn more interest as vaccine adjuvants due to their ability to enhance immune responses[20]. In fact, both non-toxic LT mutants as well as LTB subunits[4] have been used as adjuvants for several vaccine candidates[20,25,26]. It is worth mentioning that Dukoral®, one of the approved cholera vaccines, also offers cross-protection against diarrhea caused by ETEC[11,27]. This oral vaccine contains a mixture of CTB[5] and killed *V. cholerae* strains, but since LT cross-reacts with CT, protection against ETEC is also achieved[11].

As opposed to cholera, ETEC can infect both humans and animals[28]. LT toxins can be classified in two groups: LT-I and LT-II[20]. The former (referred here simply as LT) is a close homologue of the cholera toxin , sharing 82% sequence identity[28]. It exists as two variants: one that infects humans, known as hLT; and one that infects pigs, referred to as pLT[28]. The following section will explain in more detail some of the characteristics of these bacterial toxins.

1.2 Bacterial toxins

Pathogenic bacteria produce a wide range of biomolecules that enable them to invade and colonize the host, cause disease and escape the host's immune defenses[29]. These molecules are known as virulence factors and include adhesins, toxins, enzymes, iron-acquisition factors, etc[30]. Toxins play a fundamental role during infection because they can manipulate host cell functions and cause damage to tissues[31]. A large family of bacterial toxins comprises the AB-toxins. They are composed of two parts: the B component that binds to receptors on the surface of target cells, and the active A component that usually translocates to the host cytoplasm and carries out an enzymatic function on a specific cytosol target[32]. Some examples include diphtheria toxin[33], pertussis toxin[34], exotoxin A from *Pseudomonas aeruginosa*[35], the cholera toxin and the heat-labile enterotoxin[28].

1.2.1 Cholera toxin (CT)

The genes that code for the cholera toxin were acquired from a filamentous phage that infects *V. cholerae*. They are encoded as an operon (ctxAB), within the genome from the lysogenic bacteriophage CTXΦ [14]. The toxin subunits are translated from polycistronic mRNA and then translocated to the periplasm by the Sec system, where they are assembled and finally transported to the outer membrane by the type II secretion system (T2SS)[14,36]. The secretion signal lies on the B pentamer and has been suggested to be a structural motif, since no linear primary sequence has been identified[37].

The structure of CT was initially solved in 1991 by X-ray crystallography [38], however the model had poor geometry. The structure was later redetermined in 2004 by another group

[4] LTB: The cell-binding B subunit of the heat-labile enterotoxin. See section 1.2.2.
[5] CTB: The cell-binding B subunit of cholera toxin. See section 1.2.1.

and refined to a resolution of 1.9 Å[39].CT is an AB_5 toxin, consisting of a catalytically active A subunit (CTA; ~28 kDa) and five copies of the cell-binding B subunit (CTB; 11.6 kDa). The B subunits are arranged in a doughnut-shaped ring (Figure 2.A) and are held together by non-covalent interactions including hydrogen bonds and salt bridges. The A subunit is produced as a single polypeptide chain that is proteolytically cleaved by *V. cholerae* proteases between Arg_{192} and Ser_{193}[40]. This results in two peptides, the toxic A1 fragment (~22 kDa) and the A2 fragment (~6 kDa) which remain connected by a disulfide bridge[28]. CTA1 exhibits ADP-ribosyltransferase activity (in its active form) and adopts a loosely packed structure with two globular domains containing two β-sheets and several small helices (Figure 2.A). CTA2 forms a long helix stacked under CTA1 and a C-terminal tail inserted into the pore of the B-pentamer. It maintains several non-covalent interactions with CTB, and therefore acts as a linker between the CTA1 and B subunits. CTA2 ends in a short helix and a KDEL motif that protrudes from the B-pentamer and helps in the transport of the toxin to the endoplasmic reticulum (ER)[40].

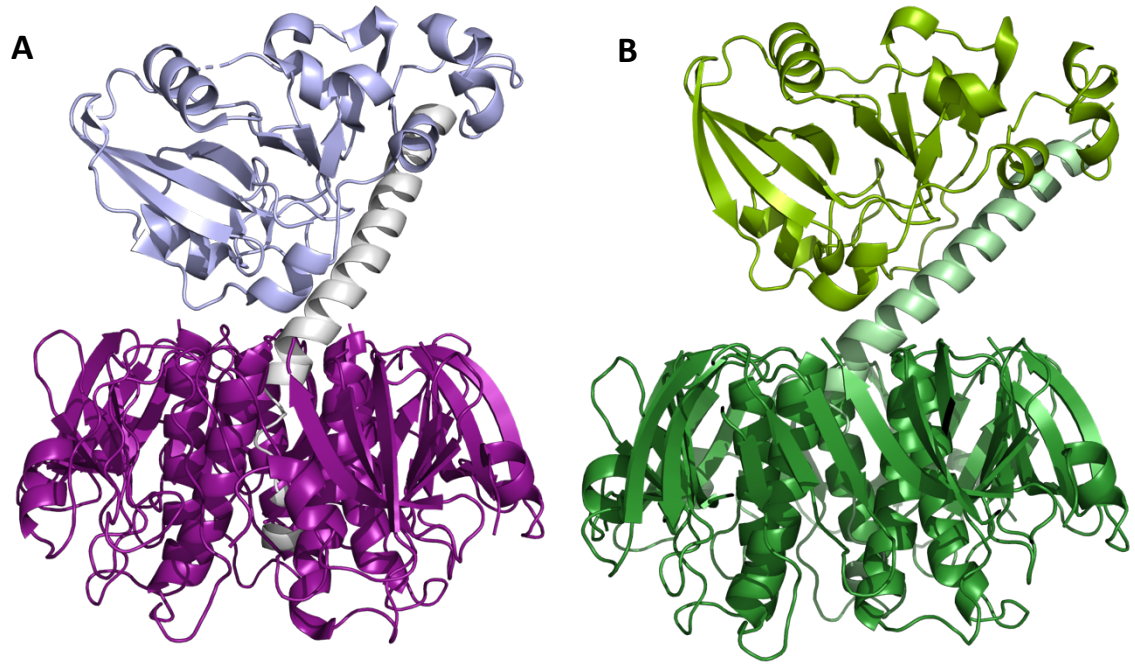


Figure 2. Crystal structures of cholera toxin (a; PDB:1S5E[41]) and heat-labile enterotoxin (b; PDB:1LTS[42]). CTA1: light blue, CTA2: gray, CTB: purple. LTA1: green, LTA2: light green, LTB: dark green. Structures were generated on PyMOL 2.4.0 (Schrodinger, Inc).

1.2.2 Heat-labile enterotoxin (LT)

The heat-labile enterotoxin (LT) is a close homologue of CT produced by ETEC, sharing 82 % sequence identity[28]. In contrast to CT, LT is usually encoded in a plasmid[14]. After the individual subunits are assembled in the periplasm, they are either secreted by the T2SS or through outer membrane vesicles (OMVs) released by ETEC[20,43]. Although the crystal structure of the LT that infects humans (hLT) has not yet been obtained, the structure of the LT that infects pigs (pLT) was obtained in 1991 and was in fact the first structure known of an

AB_5 toxin (Figure 2.b) [42]. LT has a very similar three-dimensional structure to CT, however, the latter is more toxic to cells and results in more severe symptoms compared to LT intoxication[28]. The A subunit of LT (LTA) is also produced as a single polypeptide chain, but is cleaved into LTA1 and LTA2 by intestinal trypsin and not by the bacterial proteases[14]. One of the most important structural differences with CT lies on the A2 fragment, which forms a more acute angle with respect to the B-pentamer in LT (39° in LT compared to 50° in CT)[44]. It has been recently suggested that this difference could be a determinant factor during toxin activation by PDI, which could explain the difference in toxicities between LT and CT[44], however, structural data supporting this hypothesis are currently missing.

1.3 Cholera toxin and heat-labile enterotoxin intracellular activation

Upon secretion, both LT and CT attach *via* their B subunits to intestinal epithelial cells through binding to their primary receptor, the monosialoganglioside GM1[45]. After endocytosis and retrograde vesicular transport to the endoplasmic reticulum (ER), the CTA1/CTA2 disulfide bond is reduced and CTA1 dissociates from the rest of the toxin with assistance of the chaperone protein disulfide isomerase (PDI)[46,47] (Figure 3). This step is crucial in the intoxication process, since PDI-deficient cells are resistant to CT[48]. However, the molecular interactions mediating toxin disassembly are not well understood.

Once free, CTA1 unfolds naturally at the physiological temperature and is recognized by the ER-associated degradation (ERAD) pathway[49]. However, it escapes ubiquitin-mediated degradation mainly due to the few number of lysine residues in its structure[50]. With the help of the chaperones Hsp90 and Hsc70[51,52], CTA1 is extracted from the ER and refolded[52]. In the cytosol, CTA1 can then interact with a host ADP-ribosylation factor (ARF-6) to activate the G protein stimulatory α-subunit ($G_{s\alpha}$) through ADP-ribosylation[53,54]. The modified $G_{s\alpha}$ remains in its GTP-bound state, which keeps it constitutively active and in turn activates adenylyl cyclase[55,56], a protein that catalyzes the conversion of ATP to cAMP. cAMP activates the protein kinase A (PKA), an activator of the membrane chloride channel cystic fibrosis transmembrane conductance regulator (CFTR). Upon phosphorylation, CFTR secretes chloride ions to the intestinal lumen[57]. Increased levels of cAMP ultimately result in increased efflux of ions, disturbing the electrolytic balance across the epithelial membrane. The increased salt concentration converts the intestine into a hypertonic medium, forcing large amounts of water to move out of the cell. This results in the severe watery diarrhea characteristic of cholera. LT follows the same intoxication mechanism as CT[58,59].

1.4 The role of host chaperones in toxin activation

In the crowded cellular environment, proteins are at risk of forming aberrant structures that can become toxic. To surpass this challenge, cells produce proteins known as chaperones that help other proteins to fold[60]. Interestingly, bacteria have found ways to exploit the host chaperones' abilities to activate their toxins and facilitate their translocation. This is the case for the cholera and heat-labile enterotoxin, as briefly mentioned in the previous section, but it has also been observed for toxins from other organisms including ricin[61], diphtheria[62], shiga toxin[63].

Despite the key role of human chaperones in the activation of CT and LT, very little is known about how these toxins interact with the chaperones. Of particular interest is the cholera toxin disassembly by PDI, which is essential for intoxication since PDI deficient cells are resistant to CT[48]. The following two sections will provide more details about PDI and its interactions with the cholera toxin.

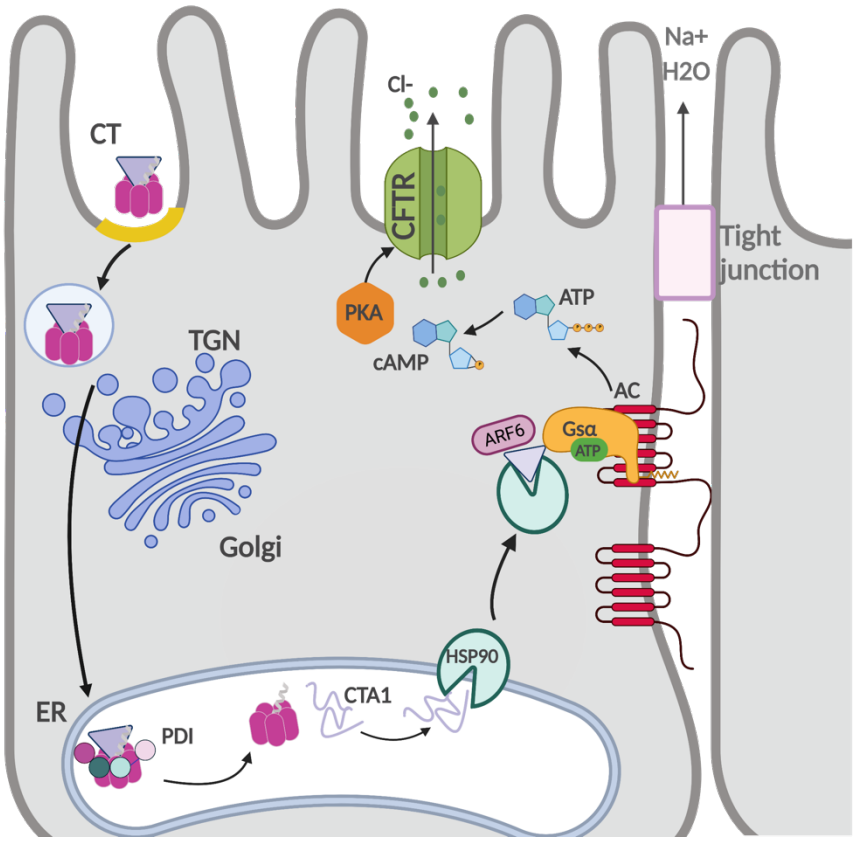


Figure 3. Intracellular activation and intoxication mechanism of the Cholera toxin. After binding to GM1 on epithelial cells, CT is endocytosed and retro-translocated to the ER, where the CTA1 subunit is separated from the rest of the toxin by PDI. On its own, CTA1 rapidly unfolds and is transported back into the cytosol by chaperones like HSP90, which also help to refold the toxin. CTA1 further interacts with ARF6 and activates $G_{s\alpha}$, resulting in increased levels of cAMP and ultimately large efflux of ions. This leads to a massive loss of water that cause the severe diarrhea typical of cholera. Illustration created with BioRender.com.

1.4.1 Protein disulfide isomerase (PDI)

PDI is a protein localized mainly in the ER, that helps other proteins to fold by catalyzing the formation and rearrangement of disulfide bonds[64]. PDI forms a U-shaped structure with four thioredoxin-like domains organized in an abb'xa' topology (Figure 4.A)[65]. The a and a' domains are located at either end of the U and contain -CGHC active motifs that oxidize thiols and reduce and isomerize disulfide bonds. The b and b' domains are substrate-binding domains located at the base of the U and are not catalytically active. b' and a' are connected by the x-linker, which confers remarkable flexibility to PDI. The C-terminal part of the protein has a short flexible tail with a KDEL motif that helps in the localization of PDI to the ER[66].

PDI cycles between a "closed conformation" when the protein is reduced, and an "open conformation" when it is oxidized [67] (Figure 4.A and 4.B). In the oxidized form of PDI, all domains are in the same plane. In contrast, reduced PDI has domains a, b and b' in the same plane, while domain a' is twisted 45° out of the plane, resulting in a smaller distance between the two active sites in domains a and a' (Figure 4.B). This redox-dependent conformation affects PDI ability to bind to substrates[68], as will be discussed later for CT disassembly.

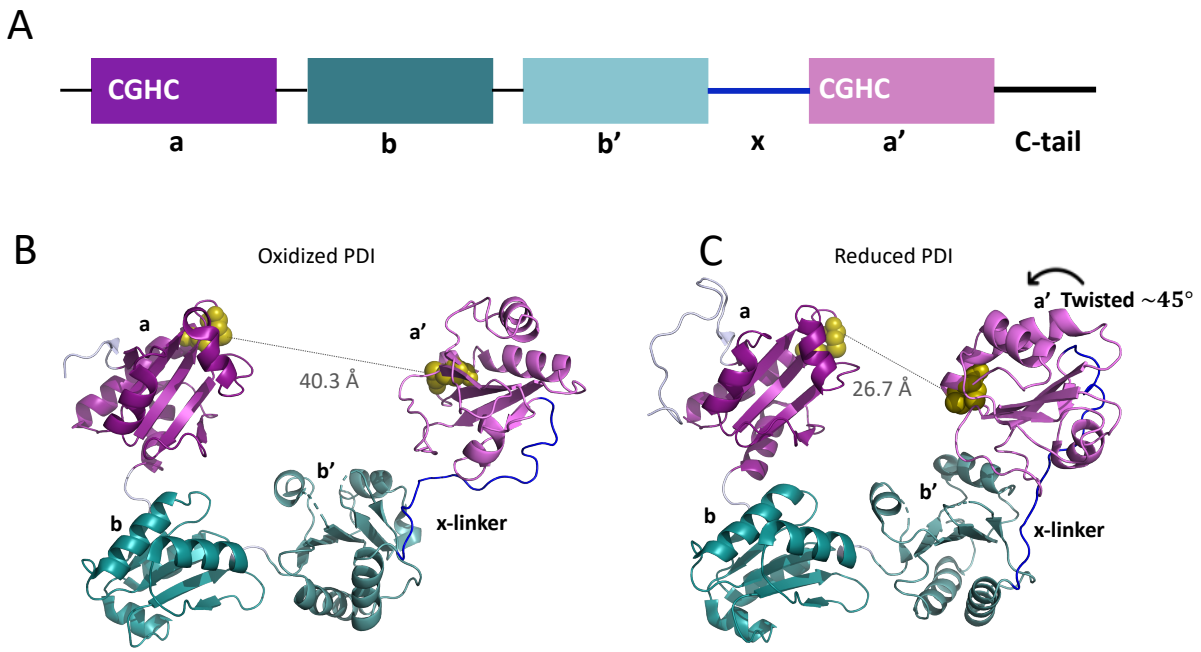


Figure 4. Domain organization (A) and crystal structures of oxidized (B; PDB: 4EL1) and reduced (C; PDB: 4EKZ) PDI. Individual domains are shown in different colors and the cysteine residues are displayed as yellow spheres. The distance between the active sites is indicated.

Apart from its oxidorectucase activity, PDI can also act as a chaperone by binding to proteins and helping them to fold by themselves[69,70]. Recently, PDI itself was found to be able to unfold and refold upon interaction with specific substrates, acting as a "disaggregase"[71]. This activity has been associated with the reversal of α-synuclein fibrils aggregation, an amyloid-forming protein involved in Alzheimer and Parkinson's disease[71]. It has been suggested that

the cholera toxin disassembly by PDI follows the same- but yet poorly understood- disaggregase mechanism[71].

The ability of PDI to displace the A1 subunit of the cholera toxin from the rest of the toxin is unique, since other PDI family members like ERp57 and ERp72 are not able to disassemble CT[72]. Even though PDI can reduce the disulfide bond that connects CTA1 to CTA2[73], this can also happen in the redox environment of the ER and does not lead to disassembly of the toxin, as the subunits stay together through several non-covalent interactions[74]. The main role of PDI in this process is the physical displacement of CTA1. It was initially thought that PDI-mediated holotoxin disassembly by unfolding the CTA1 subunit[75]. However, it has been demonstrated that CTA1 unfolds naturally at the physiological temperature of 37 °C[76]. Interestingly, PDI partially unfolds upon binding to CT[72]. It has been hypothesized that the resulting increase in PDI's hydrodynamic radius helps the chaperone to act as a "wedge" to push away the CTA1 subunit from CTA2/CTB (Figure 5)[72]. The oxidoreductase activity of PDI does not seem to be required for holotoxin disassembly, as PDI treated with bacitracin, a redox inhibitor, was still able to disassemble CT[72]. In contrast, PDI locked in a folded conformation by treatment with an intramolecular cross-linker was able to bind, but not to disassemble CT[72]. The next section will present a more detailed explanation of the experiments that support the physical model of PDI-driven CT disassembly, which is central for this thesis.

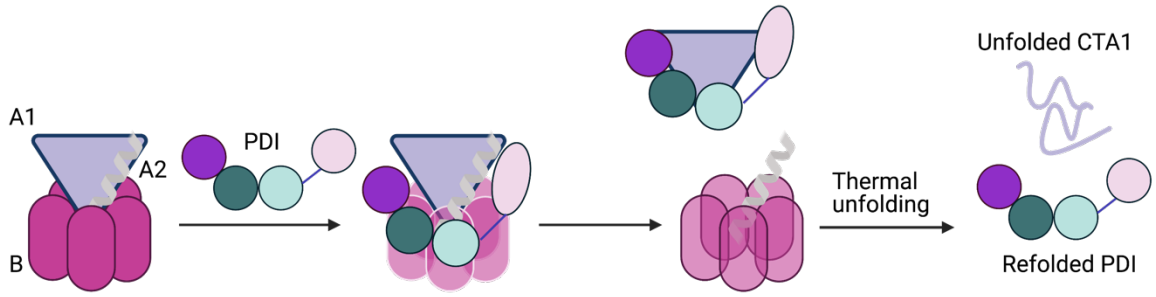


Figure 5. Wedge model of PDI-driven CT disassembly. PDI partially unfolds upon binding to the CT holotoxin. The increased size of PDI allows it to act as a wedge and push CTA1 away from the rest of the toxin[76]. Free CTA1 unfolds naturally at the physiological temperature and is released by PDI. After dissociating from CTA1, PDI regains its folded conformation[72]. Created with BioRender.com

1.5 Biophysical studies providing the foundations of CT and PDI interaction

The interaction between CT and PDI has been demonstrated by Surface Plasmon Resonance (SPR) as well as enzyme-linked immunosorbent assay (ELISA)[48,76]. These two techniques immobilize the toxin on a solid surface and detect changes in the signal after the addition of PDI and protein-specific antibodies (a theoretical explanation of how these techniques work will be provided in Section 1.6.4). Both experiments have shown that only reduced PDI is able

to bind and disassemble CT, and that a molecular excess of the chaperone over the toxin is required for efficient disassembly at neutral pH[48,76].

The SPR experiments in particular demonstrate that PDI-driven CT disassembly can occur at both pH 7.0 and 6.5 as well as at different temperatures (37 ℃, 10 ℃ and 4 ℃)[48] **(Figure 6)**. Similar SPR experiments were used to investigate the interaction between PDI and CTA1 alone and showed that this interaction at pH 7.0 is temperature-dependent: it is stronger at 10 ℃, weaker at 25 ℃ and no interaction was detected at 37 ℃ [72](Figure 7). Since the isolated CTA1 subunit is an unstable protein with a disordered structure at 37 ℃ [48,77,78], the fact that no interaction was detected at the physiological temperature indicates that PDI does not interact with unfolded CTA1. Unfolding of this subunit only occurs when it is free (not associated to CTA2/CTB)[48,77], which explains why the interaction between the CT holotoxin and PDI occurs at 37 ℃.

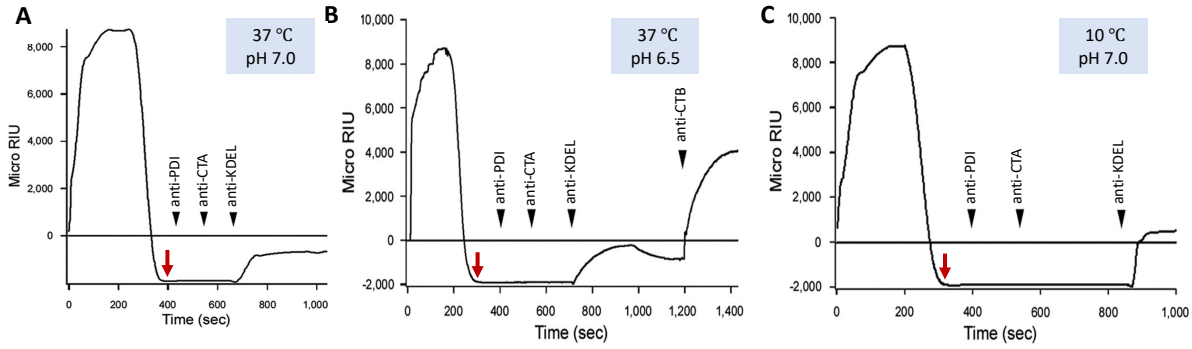


Figure 6. SPR experiments showing CT disassembly by reduced PDI at different temperatures and pH values. A-C In the SPR experiment CT is appended to a GM1-coated sensor slide and records this as the baseline signal. Since the signal is sensitive to the mass of the molecules bound to the sensor, binding of PDI to the holotoxin is observed as an increase in the signal, which then drops below the baseline level after PDI is no longer perfused (indicated by the red arrow). Sequential addition of antibodies is demarked by the black arrows. If only PDI was lost, the signal would have returned to the baseline level. Instead, it drops to a negative value indicating that also a portion of the holotoxin had been released from the sensor slide. The addition of the anti-CTA antibody did not lead to a signal increase, indicating that CTA1 is absent from the slide. This would have been the result of PDI displacing CTA1 from the holotoxin. Figure adapted from Taylor *et al* [48].

Although SPR has the advantage of monitoring CT disassembly in real-time, it is not possible to determine the relative concentrations of CT and PDI in the assay because the amount of CT bound to the sensor surface cannot be calculated[76]. However, with an ELISA-based experiment, the Teter group (University of Central Florida, UCF) showed that at least a ~2.8-fold molar excess of PDI is required to disassemble 50 % of CT[76]. The SPR experiment used an even higher excess of PDI in the perfusion buffer and disassembly appears to be more efficient than with the ELISA assay. This could possibly be due to the flow of PDI in the perfusion buffer, which could provide an additional “push” for displacing CTA1. Disassembly efficiency is also dependent on the temperature: it is faster at 37 ℃ than at 20℃ [76] and very slow at 4 ℃ [48].

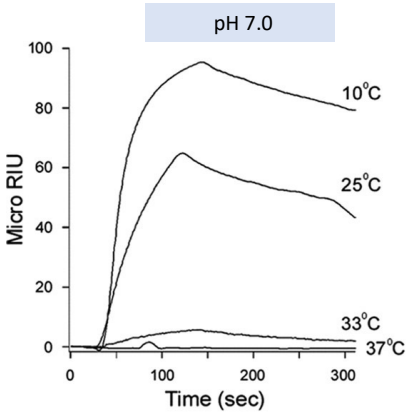


Figure 7. SPR experiment showing the interaction between PDI and CTA1 at different temperatures when the pH is 7.0. CTA1 is immobilized to an anti-CTA-coated SPR sensor. Figure adapted from Taylor *et al* [48].

The initial foundations for understanding the molecular mechanism of PDI-driven CT disassembly have been provided by Isotope-edited Fourier Transform Infrared Spectroscopy (FTIR) and Circular Dichroism (CD)[6]. Both experiments are carried out in solution and provide information about changes in the secondary structure of PDI upon interaction with CT or only CTA1[72]. The CD spectra of reduced PDI and CTA1 at 10 ℃ and 1:1 molar ratio indicated that the interaction between CTA1 and PDI resulted in conformational changes of either one of the proteins or both. FTIR resolved this ambiguity. In this technique, one of the proteins must be labeled with ^{13}C. The conformation of the protein is not altered by the labeling, however, the increased mass of the ^{13}C isotope causes a downshift of the amide I band that allows to distinguish between the FTIR spectrum of the ^{13}C-labeled protein from the spectrum of the same unlabeled protein[79,80]. The recorded FTIR spectrum is used to manually model the contribution of secondary structure elements. These experiments show that reduced PDI partially unfolds in the presence of an equimolar ratio of CTA1 at 10℃ [72]. In contrast, PDI did not unfold at 37 ℃, which confirmed the previous observation that PDI can only bind to folded CTA1[48,72].

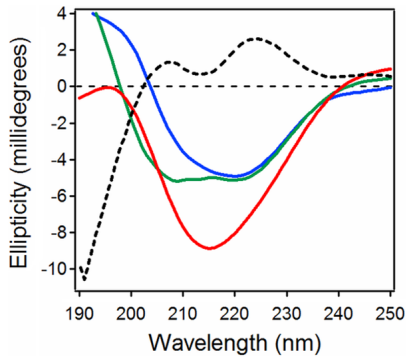


Figure 8. Far-UV CD spectra of CTA1 (blue), PDI (green) and a mixture of both proteins at a 1:1 molar ratio (red). The black dotted line corresponds to the subtraction of individual CTA1 and PDI spectra from the spectrum of the protein mixture. Figure taken from Taylor *et al*. [72]

[6] For a theoretical explanation on how these techniques work see Barth[160] and Ismail *et al*[161] .

The secondary structure of PDI was also monitored by FTIR in the presence of the CT holotoxin[72]. Interestingly, it was observed that PDI had started to unfold already after 1 minute of mixing the proteins at 1:1 molar ratio and 37 °C, suggesting that the interaction under these conditions is a very fast process, contrary to the results obtained by ELISA using an immobilized toxin. Within 25 minutes of exposure to CT, the chaperone exhibited an increase in its secondary structure, which was interpreted as PDI starting to refold after displacing CTA1 from the toxin[72].

Additional SPR and FTIR experiments using either an intra-crosslinked or inhibited PDI showed that the chaperone can bind to CT, but is unable to disassemble it[72,81]. With both treatments, PDI assumes a more rigid conformation[72,82] that blocks its ability to unfold and thus disassemble CT.

These studies have shed light onto the conditions required for the interaction between PDI and CT or CTA1 (e.g., PDI redox state, temperature, pH) as well as the changes in the PDI secondary structure when the toxin is present (summarized in Table 1). However, many questions regarding the disassembly mechanism remain unsolved: which domain(s) of PDI bind to CT? What is the role of PDI redox state on binding and disassembly? Which domain (s) are directly responsible for CTA1 displacement? Which domains of PDI unfold? The Teter group at the University of Central Florida (UCF) has recently carried out preliminary ELISA and SPR experiments with PDI deletion constructs (personal communication), which have provided insight into the role of individual PDI domains on CT binding and disassembly. These results show that the ab domains alone can disassemble CT, which occurs only when the a domain is reduced but not when it is oxidized. In contrast, the bb'x domains are sufficient for binding but cannot disassemble CT. Interestingly, the a' domain unfolds upon contact with CTA1 but is not required for CTA1 binding or CT disassembly.

Table 1. Summary of conditions that have been explored in previous interaction studies involving CT, CTA1 and PDI. Experiments that used CTA1 are highlighted in blue; experiments with CT in white.

Observation		**Conditions**				**Method**
Binding	CT Disassembly	Temperature	pH	Molar ratio CT:PDI	Specific treatments	
		37 ℃	7.0	PDI excess	None	SPR[48,72], ELISA[76]
X	X	37 ℃, 10 ℃, 4 ℃ 30 ℃ 37 ℃	7.0 7.4 6.5	PDI excess	GSH	SPR[72,76]
		37 ℃, 20 ℃	7.4	1: 2.8	GSH/DTT	ELISA[76]
X		10 ℃ 37 ℃	7.0 6.5	1:1	GSH	FTIR[72]
X	X	37 ℃	7.0	1:1	GSH	FTIR[72]
X		37 ℃ 37 ℃ 10 ℃	7.0 6.5 7.0	1:1	GSH+EDC	FTIR[72]
X		37 ℃	7.0	PDI excess	GSH+Q3R	SPR[81]
		20 ℃	7.0	1:1	GSH+Q3R	FTIR[81]
X		37 ℃	7.0	PDI excess	GSH+Q3R	SPR[81]

1.6 Theory background

1.6.1 Protein X-ray Crystallography

X-ray crystallography is the traditional technique to determine three-dimensional structures of molecules, and has almost no size restriction. It relies on the interaction of X-rays (electromagnetic waves in the range of angstroms) with electrons, which diffract the X-ray beam[83]. Since this interaction is weak, millions of copies of the molecule of interest organized within a crystal are used to magnify the signal.

Crystals are obtained from a solution of highly concentrated pure protein (in general >95% pure on SDS-PAGE) that is placed under conditions that significantly reduce its solubility. If this happens too fast, precipitation occurs, but if it is done slowly under the right conditions, crystals will grow[84]. This is achieved by mixing the protein with a crystallization solution, which often contains precipitants like salts and polymers that reduce protein solubility. The most common method for protein crystallization is vapor diffusion, which exists in two formats: sitting-drop and hanging-drop[85]. Briefly, a drop containing a mixture of the protein sample and crystallization solution is placed in a sealed well containing a liquid reservoir of the crystallization solution. The drop can be placed either on a plateau above the reservoir

(sitting-drop) or on a surface inverted above the well (hanging-drop). Since the drop contains a lower concentration of the crystallization solution than the reservoir, water evaporates and transfers from the drop to the reservoir until equilibrium is reached. As water leaves, the protein and precipitant concentration in the drop increases, which may lead to formation of crystals[86]. Crystallization depends on many variables, including the temperature, pH, buffer, protein concentration, type of precipitant, presence of additives, etc. Although automation has made it possible to screen hundreds of conditions at a time, crystallization is not a predictable and controllable task yet, and it continuous to be a trial and error process[86].

Crystallization of proteins bound to small molecules is also possible. This is usually carried out by two methods: soaking and co-crystallization[87]. In the former, an existing protein crystal is soaked with a high concentration of ligand solution. The ligand can diffuse into the crystal lattice due to the solvent channels and bind to the protein if the binding sites are accessible (e.g., not part of crystal contacts). An alternative approach is co-crystallization, where the protein and ligand solution are mixed and then crystallization screens are set-up[87].

Once crystals are obtained, they are exposed to a monochromatic X-ray beam either in home X-ray sources or at synchrotrons. The latter uses high intensity X-rays that have the advantage of reducing exposure times and provide high signal to noise ratios, making synchrotrons the preferred X-ray source for protein crystallographers[84]. When X-rays strike the protein crystal, they are scattered by the electrons giving rise to a diffraction pattern. The amplitudes of the waves are obtained from the intensities of the diffracted X-rays; however, the information of the phases is lost, that is, the origin of each wave with respect to each other. Without the phases, the position of the atoms in the structure cannot be determined. This is known as the phase problem in X-ray crystallography, and albeit challenging, can be solved by different methods[88]. One of them is molecular replacement, which uses the phases of another molecule whose structure is known and resembles the protein of interest. If no similar structure is known, the phases can be obtained experimentally by several methods that usually involve introducing heavy atoms in the protein structure[7][89]. Once the phases are found, they are summed in all directions to generate an electron density distribution of the molecules. A chemical model can be fitted into this density, providing a first model of the 3D structure of the protein. This model is improved by moving the position of the atoms such that they fit the experimental data as good as possible in a process known as refinement (Figure 9)[83]. This is done in several rounds (iterative process) until a good fit is obtained.

[7] Nowadays, a common approach is to replace methionine with selenomethionine in the protein. Selenium does not perturb the structure of the protein, but it is an effective anomalous scatter. Anomalous scattering occurs when the wavelength of the incident X-rays is close to the transition between different electron shells in the atom. At this point, the scattering changes rapidly and results in a change of both the amplitude and the phase[77], which can be used to derive the position of the atoms in the molecule.

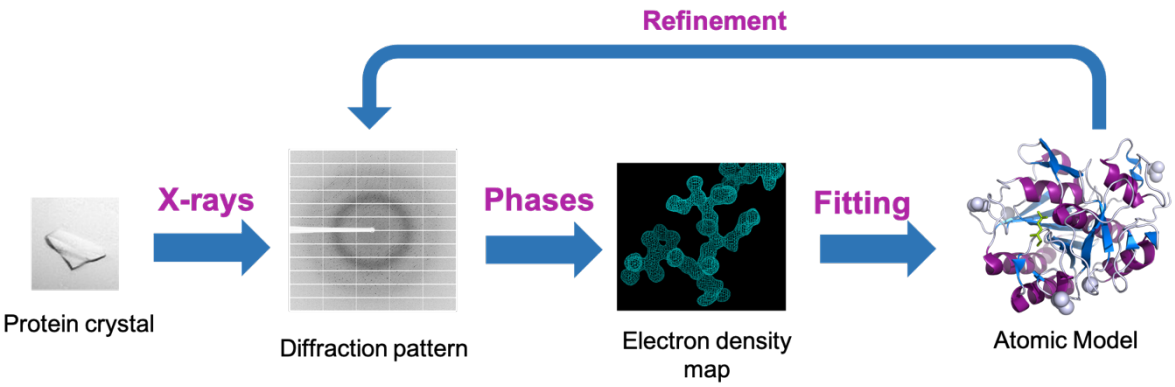


Figure 9. X-ray crystallography workflow. When an X-ray beam strikes a protein crystal, X-rays are scattered and detected as a diffraction pattern. This pattern only contains information of the amplitudes of the waves but not the phases. Once the phases are determined (either experimentally or computationally), they are combined to generate an electron density map of the protein in the crystal. An atomic model is fitted into the electron density, and after several rounds of refinement, a 3D structure of the protein is obtained. Images of the protein crystal and diffraction pattern kindly provided by Henrik Vinther Sørensen, University of Oslo.

The main bottleneck of this technique is obtaining protein crystals, since large amounts of highly pure protein are required and there is no guarantee that the protein will crystallize, e.g., if the protein is too flexible. In addition, X-ray diffraction depends on the crystal architecture, so imperfections in the crystal result in poor diffraction. Nevertheless, this technique usually provides the highest resolution structures and remains the gold standard in structural biology.

1.6.2 Single particle cryogenic-electron microscopy (Cryo-EM)

Cryo-EM is an imaging technique used for investigating three-dimensional structures of frozen samples near their native state[90]. This technique comes in two main flavors: cryo-electron tomography (cryo-ET), to visualize whole cells and organelles, and single particle cryo–electron microscopy (SP cryo-EM) to visualize macromolecules[91]. Unlike X-ray crystallography, single particle cryo-EM does not require crystals, and only very small amounts of protein are needed (microgram range). The number of structures determined by this technique have increased exponentially in the recent years, and notable improvements in both instrumentation and imagining processing have allowed to reach atomic resolution, like in the case of Apoferritin[92,93].

Cryo-EM samples are prepared by placing a few microliters of the pure protein on a support grid made of a conductive material (e.g., Cu or Au) and coated with a carbon film. Grids are blotted to remove excess liquid and then vitrified by plunging them into liquid ethane[90]. This freezes the sample fast enough to prevent water molecules from forming ice crystals, preserving the specimen in a close to native state that protects it from radiation damage[94]. The grid is placed in a transmission-electron microscope (TEM), where an electron beam is

focused at the sample by electromagnetic lenses. When the electrons interact with the atoms of the sample, they are scattered and detected, forming a two-dimensional image. Hundreds to thousands of short movies are recorded, which after correction of specimen movements, result in images. These images are sorted according to their orientations and then combined to reconstruct a three-dimensional structure of the protein[95]. In contrast to X-ray crystallography, the phases of the scattered waves are not lost, and a structure of the protein can be directly determined after a thorough computational analysis.

The radiation sensitivity and low signal-to-noise ratio of macromolecules in vitreous ice, makes the visualization of specimens <200 kDa difficult[96]; however, size limits are continuously challenged and several structures of macromolecules below 200 kDa have been determined[97–99], the smallest being the SAM-IV riboswitch (39 kDa)[100].

In many cases, the quality of the sample is assessed prior to cryo-EM by a technique known as negative stain electron microscopy. After adsorbing the sample on an EM grid, an electron dense compound is added (e.g., uranyl acetate or formate) to stain the background, but not the particles, which appear as lighter areas in the image compared to the surrounding solvent[101]. This technique allows to visualize the distribution of the particles (is the sample too crowded or can individual particles be clearly identified?), determine homogeneity/heterogeneity of the sample (are the proteins in different states or forming larger structures?), and evaluate formation of complexes or large assemblies. In this way, the samples can be optimized before starting cryo-EM experiments.

Besides the fact that no crystals are required, the increase in resolution and decrease in sample size limits have made this technique very attractive for studying 3D structures of a wide range of proteins.

1.6.3 Small-Angle X-ray and Neutron Scattering (SAXS/SANS)

Small angle X-ray and neutron scattering are techniques that offer low resolution structures of molecules in solution (nanometer scale)[102]. They are especially useful for studying multi-domain and flexible proteins which are difficult to investigate by other methods like X-ray crystallography[103]. In a typical experiment, a sample is placed in a capillary and exposed to a beam of X-rays (SAXS) or neutrons (SANS) (Figure 10). The incident radiation is elastically scattered by the sample and measured at low angles (0.1 to 5 deg). The scattering is analyzed and provides information on the size, shape and oligomeric state of the molecules[102]. It is also possible to generate a 3D representation of the molecule or “envelope”. This can be achieved either by *ab initio* methods, where no previous information of the structure of the molecule is known, or by modeling the X-ray, NMR or EM structure if it has been solved for the macromolecule[104].

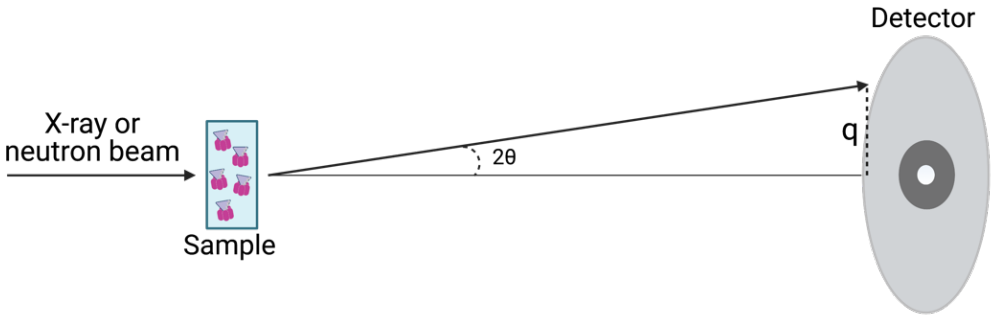


Figure 10. Schematic representation of a typical small angle scattering experiment. An X-ray or neutron beam is shot at a sample, and the scattered radiation is detected at low angles. Created in Biorender.com

In this thesis, only a brief explanation of how SAS works is presented. An advanced explanation of the physics behind these techniques can be found in Feigin and Svergun[105] and Svergun *et al* [106].

The scattering intensity, I, depends on the scattering angle (2θ) and is related to the structure of the molecules in solution by the following expression:

$$I\,(q) = S\,(q)\sum_{i}^{n}[(\Delta\rho_i V_i)^2 P_i(q)] \quad (1)$$

Where q is the scattering vector and relates to the scattering angle through λ (wavelength of incident radiation) as shown in (2) :

$$q = \frac{4\pi sin\theta}{\lambda} \quad (2)$$

The relationship in (1) indicates that the intensity of scattered radiation is the sum of the scattering from each particle in the sample, i, within the illuminated sample volume. The scattering intensity depends on several variables, including the volume squared of each particle V_i^2; the contrast square $\Delta{\rho_i}^2$, which is the difference in the scattering density of the macromolecules and the solvent; the structure factor $S\,(q)$, which gives information regarding the interactions between particles; and the form factor $P_i(q)$, that contains information about the size and shape of the particles. Due to the random orientation of molecules in solution, the resulting scattering is isotropic and is radially averaged[107]. This results in a one-dimensional scattering curve that is usually plotted as $I\,(q)\ vs\ q$, which gives information about the intramolecular distances of the protein but not their spatial orientation.

As all atoms scatter radiation, all the species present in the sample contribute to the measured scattering. To obtain accurate structural information of individual macromolecules

or complexes, the sample must be homogeneous and not affected by interparticle interactions ($S\ (q) = 1$). Under these conditions, (1) can be simplified to

$$I(q) = N(\Delta\rho V)^2 P(q) \quad (3)$$

Where N is the number of homogeneous particles in the sample.

The success of SAXS and SANS experiments relies on the production of monodisperse and highly pure samples. Since the scattering intensity increases with the squared volume of the particles as shown in (1), if large impurities or aggregates are present –even at low concentrations– they can lead to alterations on the scattering curves[103,107], making the calculation of 3D spatial models very difficult.

Another important factor for an accurate SAS analysis is the correction of the scattering from background contributions. The buffer used in the sample must match exactly the one used to measure the background scattering, so that the subtraction of both measurements gives rise to the scattering of the macromolecule[107].

The overall size of the macromolecules is derived from the Guinier analysis, which provides the Radius of gyration (R_g), a measure of the intramolecular distances with respect to the center of the electron density of the molecule[108]. R_g is obtained from the slope of the "Guinier region[8]" of the plot $n(I)\ vs\ q^2$, where $ln(I) = \ln\big(I(0)\big) - \frac{1}{3}R_g^2 q^2$. The Y-intercept, $I(0)$, is the scattering intensity at *q = 0* and provides an estimation of the molecular mass. Furthermore, the hydrated volume of the macromolecule can be determined by Porod's law and the maximal intramolecular distance (D_{max}) can be estimated by an inverse Fourier transformation[9] of the intensity, known as the pair distance distribution function $P\ (r)$ [109]. This function represents the probability distribution of distances r between any two atoms in the macromolecule, resulting in a curve with peaks where the most probable distances in the molecule are[110]. The overall particle shape can be estimated from this curve: globular molecules display a symmetric bell-shaped plot, while unfolded particles display an extended line[111] (Figure 11.A). Another presentation of the SAXS data is the Kratky plot, $I(q)q^2\ vs\ q$, which gives information about the flexibility and degree of unfolding of the protein[112]. Molecules with different shapes display characteristic curve shapes in the Kratky plot as shown in Figure 11.B.

[8] The Guinier region corresponds to the low-q portion of the data.

[9] A Fourier transformation is a mathematical tool that allows to convert signals from the time domain to the frequency domain. An inverse Fourier function does the opposite.

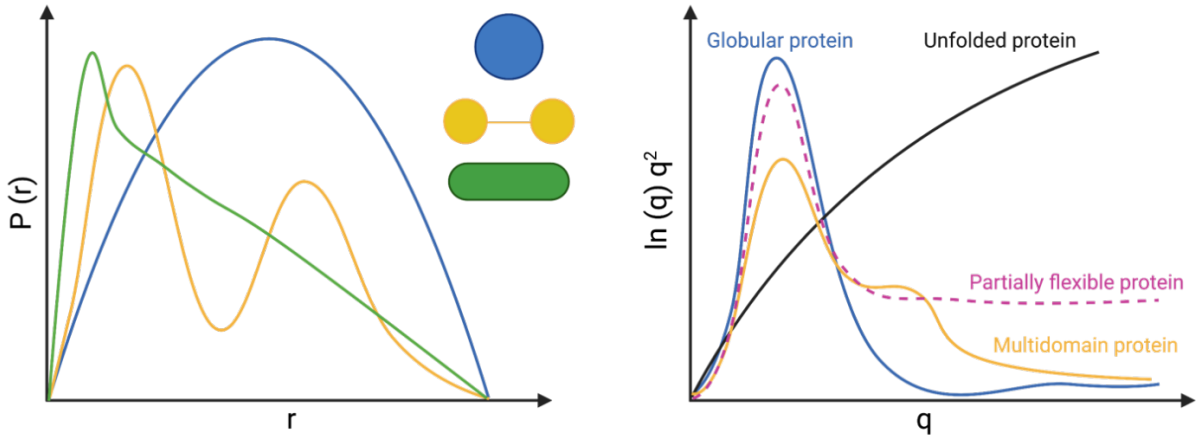


Figure 11. Typical pair distance distribution plots (left) and Kratky plots (right) for proteins. Shown are the P(r) curves for a sphere (blue), dumbbell (yellow) and rod (green). Typical Kratky plots for a globular protein (blue), unfolded protein (black), partially flexible protein (pink dashed-line) and multidomain protein (yellow). Image generated in Biorender.com.

While X-rays are scattered by electrons, neutrons are mainly scattered by the atomic nuclei[103]. This fundamental difference between SAXS and SANS makes the combination of these two techniques especially useful for studying multi-component samples like protein-protein complexes. Neutron scattering depends on the isotopic composition of a macromolecule and the solvent and can be either positive or negative. Particularly interesting is the fact that hydrogen and its isotope deuterium have dramatically different scattering lengths[113], and this difference can be exploited to enable contrast variation. By varying the ratio of H_2O and D_2O of the solvent, it is possible to find conditions where one component of the complex has the same scattering as that of the solvent. Under these conditions, the "matched" component becomes invisible to SANS, allowing the visualization of only the other component in the complex[110] (Figure 12). For most proteins, the matching point is between 40-45% D_2O, while for DNA/RNA it is 60-70%[113]. In the case of complexes consisting of two proteins, one of them must be partially deuterated to obtain different match points, otherwise both will be matched simultaneously, or they will contribute equally to the scattering. Deuterated proteins are obtained through production in cell cultures grown in D_2O-based media with a deuterated carbon source (e.g., glycerol or glucose)[114]. Efficient protocols for the production of deuterated proteins have been developed for bacteria (*Escherichia coli*) and yeast (*Pichia pastoris*)[114].

SANS is significantly more demanding than SAXS, as it requires higher amounts of protein (200-500 μL of the sample at 5-10 mg/mL for a single measurement)[10] due to the relatively low flux of neutron sources[103]. In addition, the large incoherent scattering of hydrogen makes measurements in H_2O-based solvents problematic, as it contributes to a significant

[10] The amount of protein depends on the size (the larger the protein, the more it scatters) as well as the neutron source. Decent data can be obtained with concentrations as low as 1.5-2 mg/mL for a 50 kDa protein (PhD thesis of Henrik Vinther Sørensen, UiO).

amount of noise in the SANS data. Nevertheless, this is a powerful technique that can give valuable structural information.

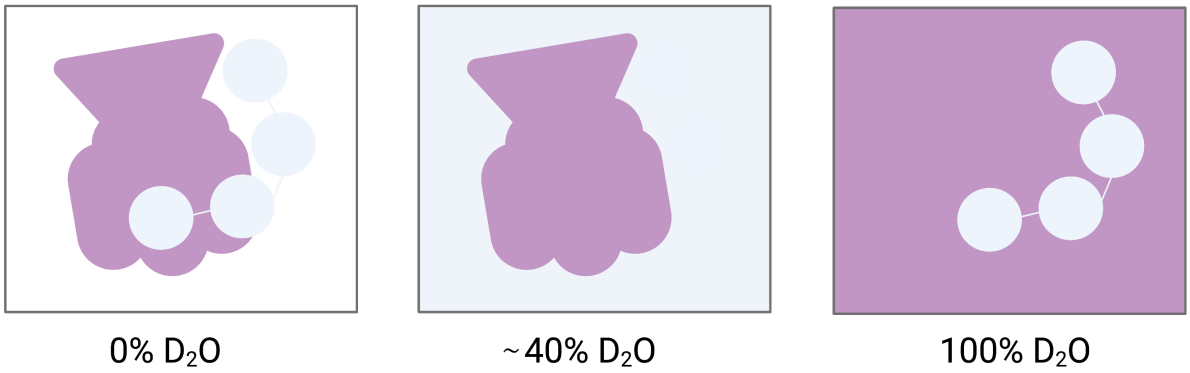


Figure 12. Contrast-matching for a protein-protein complex. At 0% D_2O, both proteins in the complex are visible. At the match-point of the hydrogenated protein component (e.g., 40% D_2O), only the deuterated protein is visible, while at 100 % D_2O only the hydrogenated component is visible.

1.6.4 Methods for studying protein-protein interactions

Surface Plasmon Resonance (SPR)

SPR is one of the most widely used techniques for monitoring binding between two or more molecules in real time[115]. SPR allows to determine both the binding affinity and association/dissociation kinetics of an interaction, based on changes in the refractive index of the medium when molecules bind to the surface of a sensor slide[116]. The slide is made of gold or silver and is placed at the interface of two dielectric (poor conductive) media with different refractive indices. When polarized light is focused on the surface, it interacts with surface plasmons – excited free electrons oscillating along the surface, which resonate with the electric field of incident light[117]. The oscillations are highly sensitive to changes in the surrounding, such as adsorption of molecules.

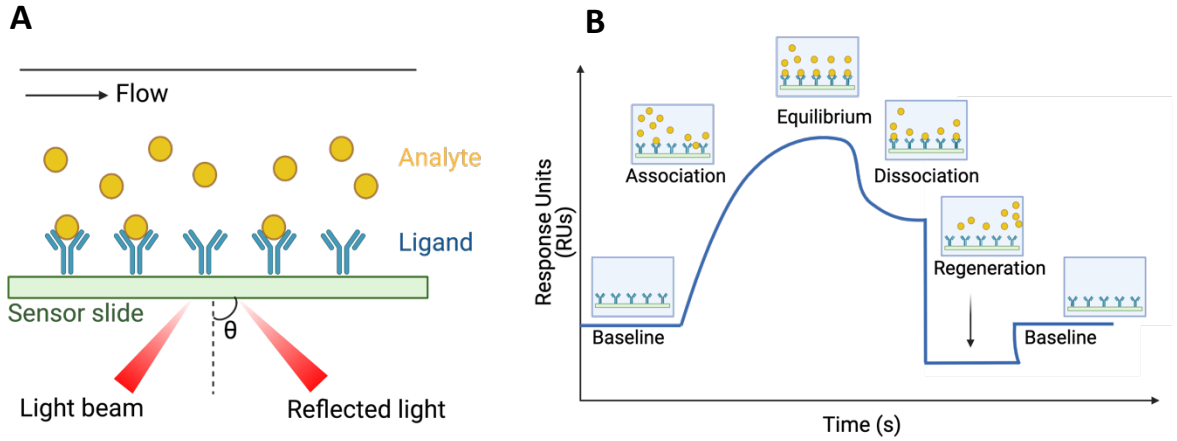


Figure 13. Experimental SPR set-up (A) and typical sensorgram (B). When polarized light is shot on a metal film, the reflected light is detected and analyzed. Image generated with Biorender.com.

In a typical SPR experiment, a molecule (called the ligand) is immobilized on a metal surface, followed by injection of the binding partner, called the analyte (Figure 13.A). As the refractive index is proportional to the mass on the sensor slide, if a binding event occurs, the refractive index increases. This is observed as a change in the resonance angle of refracted light, which is converted to resonance units (RU). The change in RU signal over time generates a curve known as sensorgram (Figure 13.B), where different stages of the binding event can be evaluated. A baseline signal is initially recorded, which corresponds to the RUs when the ligand is bound to the surface. Then, the analyte is added and an increase in the RU signal is observed if it binds to the ligand (association phase). After a certain time, the number of association and dissociation events are the same and equilibrium is reached. Injection of the analyte is then stopped, causing the analyte molecules to dissociate which leads to a decrease in RUs. Complete removal of the analyte occurs during the regeneration step until the baseline is reached again. The curve can be used to calculate the binding affinity (K_d).

Micro-scale thermophoresis (MST)

MST is another useful technique for studying interactions between biomolecules[118]. It measures changes in the mobility of fluorescent molecules under microscopic temperature gradients, known as thermophoresis[119]. When the buffer is kept constant, the thermophoresis only depends on changes in the charge, size and hydration shell of the molecules[118,120]. If an interaction between two molecules occurs, at least one of these properties changes, resulting in a different thermophoretic movement compared to that of the unbound molecule[121].

In MST, a fluorescent protein (either labeled or intrinsically fluorescent) is titrated with a non-fluorescent molecule and loaded into a glass capillary. A temperature gradient is induced by an infrared laser, and the thermophoretic movement is measured as the fluorescence distribution inside the capillary. The change in thermophoresis of the fluorescent molecule can be used to calculate an equilibrium binding constant[120]. Compared to other methods, MST has the advantage that no immobilization of the biomolecules is required and that very small amounts of protein are needed (few microliters of samples in nM range)[120].

Isothermal titration calorimetry (ITC)

ITC is the standard method for determining thermodynamic parameters of biomolecular interactions in solution by measuring the heat that is absorbed or released during a binding event[122]. As no immobilization or labeling of the binding partners is required, this method provides affinity measurements in the native state. ITC uses a calorimeter containing a reference cell filled with solvent, a sample cell with the analyte and an injection syringe[123]. The ligand is titrated into the sample cell, and if binding occurs, the temperature in the cell

changes. Since the calorimeter must keep the reference and sample cell at the same temperature, it applies or reduces heat to return the cells to the same temperature. From each ligand injection, a change in the heat of reaction (ΔH) is obtained, which can be plotted and integrated against the molar ratio of the injection to create a titration curve[123]. This curve directly provides the binding affinity, enthalpy, entropy and stoichiometry of the binding reaction[122].

Enzyme-linked immunosorbent Assay (ELISA)

ELISA is a technique used to detect and quantify the presence of a protein in a sample using specific antibodies. There are many ways to perform an ELISA, but all of them share the following steps[124]. First, an antigen is immobilized on the surface of a microtiter plate (coating). Then, the remaining available sites on the surface are coated with a protein that does not disturb the assay, such as bovine serine albumin (BSA) (blocking). In some cases, a primary antibody covalently liked to an enzyme is added, which reacts upon addition of a substrate, leading to absorption or a fluorescence signal. This signal is proportional to the amount of binding antibody and can be quantified. If the first antibody is not linked to an enzyme, a second antibody with affinity for the primary antibody is added[125]. This is called the secondary antibody and is conjugated to an enzyme that gives rise to a signal. Typical enzymes used are Horseradish peroxidase (HRP) and alkaline phosphatase (AP)[125]. A crucial step in ELISA experiments is washing, which is done thoroughly after the addition of the antibodies to remove unbound proteins.

In addition to the methods described above, protein-protein interactions can also be studied by other techniques including fluorescence resonance energy transfer (FRET), bio-layer interferometry (BLI) and size-exclusion chromatography (SEC). For a theoretical background on these techniques, the reader can refer to reference [126].

2 Aim of the thesis

The mechanism of cholera toxin disassembly by PDI is not well understood at the structural level. The interaction between these proteins is highly dynamic and requires an integrative structural biology approach. Although a model for toxin disassembly has been proposed[76], the structural details regarding the CT binding sites as well as the specific domains of PDI involved in binding and disassembly remain unknown. Our collaborators at UCF have identified different conditions under which the interaction between CT with PDI takes place, which could serve as a starting point for obtaining a stable protein complex. **The aim of this thesis was to gain structural insight into the interaction of CT with PDI by Small-Angle X-ray Scattering (SAXS).** As high amounts of protein are required for this technique, the initial focus of this project was to optimize the expression yields. The combination of SAXS with Small-Angle Neutron Scattering (SANS) could provide additional information about the disassembly mechanism, and therefore preparatory work for SANS studies was also pursued in this thesis. PDI also mediates disassembly of other CT homologues including pLT, which was used as an alternative target in this thesis.

This thesis had three main objectives:

1. Production of hydrogenated and deuterated protein targets: Establish and optimize protocols for the expression and purification of the cholera toxin, pLT and both hydrogenated and deuterated PDI.
2. Structural characterization of CT and PDI by SAXS and crystallization of CT.
3. Interaction studies: Investigate conditions that lead to a stable interaction between PDI and CT by SEC and SAXS.

3 Methods

This chapter describes the experimental procedures used and developed in this thesis. A list of all the materials and equipment used can be found in the Appendix Section A.

3.1 Cloning and transformation of CT, pLT and PDI

CT is encoded by a pARCT5 vector generously provided by Professor Randall Holmes. It contains an L-arabinose-inducible operon with periplasmic signal sequences[11] derived from the LT-IIB gene, which directs secretion of both CTA and CTB subunits to the periplasmic space more efficiently than the native CT secretion signals[127].

The pLT gene was synthesized by Genscript, inserted into a pUC57 plasmid and subcloned[12] in this thesis into a pARCT5* vector, where the NcoI site in the chloramphenicol gene had been previously removed by a silent mutation using primers NcoI_Fwd and NcoI_Rev (Appendix Section D)[13]. The DNA coding for CTA and CTB in pARCT5* was replaced by pLTA and pLTB DNA, keeping the LT-IIB signal sequence.

Full-length human PDI with an N-terminal His_6-tag is encoded in a pOLR130 plasmid kindly provided by Professor Ken Teter (UCF). The plasmid harbors an ampicillin resistance gene.

3.1.1 Cloning of pLT

The CT gene in pARCT5* and pLT gene in pUC57 are flanked by two *NcoI* restriction sites (Figure 15). Both vectors were digested with the restriction enzyme NcoI-HF® (New England Biolabs, NEB) following the protocol described below. Using the same restriction enzyme for cutting both plasmids creates complementary sticky ends that help to ligate the two DNA pieces. To separate the insert (pLT) and the destination vector (pARCT5*) from the rest of DNA fragments generated, the digestion reactions were run on an agarose gel. The two relevant bands were cut out and purified with the QIAquick Gel Extraction Kit (QIAGEN). The destination vector was dephosphorylated using shrimp alkaline phosphatase rSAP (NEB) and then ligated with the insert. The ligation mixture was used to transform DH5α *E. coli* competent cells as described in Section 3.1.2. The orientation of the insert was verified by a restriction enzyme digest and positive clones were sent for sequencing.

[11] Amino acid sequence that directs proteins to the bacterial periplasmic space (region between inner cytoplasmic membrane and outer bacterial membrane).
[12] Technique to move a DNA sequence from a parent vector to a destination vector.
[13] The silent mutation was introduced by Joel Benjamin Heim, Department of Chemistry, UiO.

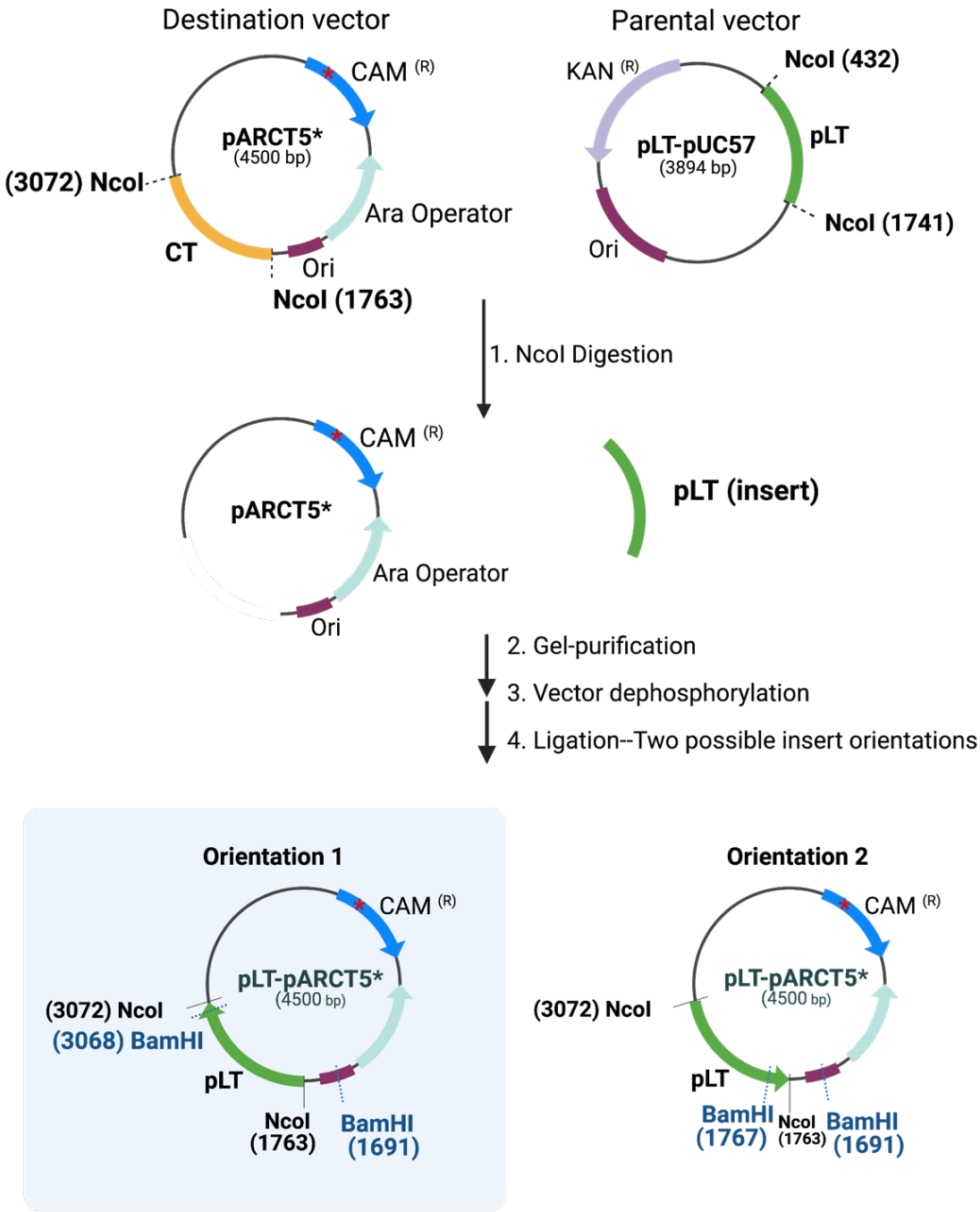


Figure 15. pLT subcloning workflow. The CT and pLT genes were excised from their respective parental vectors by digestion with the NcoI restriction enzyme. After gel-purifying the pLT insert and the pARCT5* vector, the latter was dephosphorylated by treatment with shrimp alkaline phosphatase, and then ligated with the insert. Correct orientation of the insert (plasmid within light blue square) was verified by a diagnostic restriction digest with the enzyme BamHI, followed by sequencing. NcoI and BamHI restriction sites indicated by the dashed line. Relevant plasmid elements are highlighted in different colors. Blue: chloramphenicol resistance marker (CAM $^{(R)}$). Light purple: kanamycin resistance marker (KAN $^{(R)}$). Light green: arabinose operator. Dark purple: origin of replication (Ori). The red star in the CAM region represents the silent mutation introduced to remove an NcoI site (leaving only two NcoI restriction sites in the vector). Figure generated with Biorender.com.

Restriction enzyme ligation

Two parallel digestion reactions were set up. Briefly, 1 μg of pARCT5* or pLT-pUC57 was mixed with 5 μL of 10X rCutsmart™ buffer (NEB) (1:10 dilution), 1 μL of NcoI-HF (20 units, NEB) and milli-Q water to a final volume of 50 μL. The reactions were incubated for 1 hour at 37 ℃ in a bead bath (VWR) followed by heat-inactivation at 65 ℃ for 15 minutes.

Agarose gel electrophoresis and gel purification

Restriction-enzyme digests were separated according to their size by agarose gel electrophoresis. This method applies an electric field to move the negatively charged DNA molecules through the pores of an agarose gel. As smaller molecules move faster than larger molecules, it is possible to estimate the approximate size of DNA fragments by comparing their position in the gel with that of a DNA ladder (mixture of DNA fragments with known lengths).

Standard protocol: A 1% agarose solution in TAE buffer was heated in a microwave until complete dissolution and then allowed to cool to 50-60 ℃, before addition of ethidium bromide[14] to a final concentration of 10 μg/mL. The solution was poured into a gel tray with an 8-well comb and allowed to solidify at room temperature for 20-30 minutes. After the comb was removed, the gel was placed in an electrophoresis tank (HE33 Mini submarine electrophoresis unit, Hoefer) and covered with TAE buffer. Samples were prepared by mixing the DNA with 6X DNA Loading Dye (Thermo Scientific™) and then loaded into the gel wells. Perfect DNA 1 kbp ladder (Novagen) was also loaded in one of the wells. The tank was connected to a power supply (PowerPack HC, Bio-Rad) to apply an electric field of 90 V for 60 minutes. The bands were visualized under ultraviolet-light (UV Transilluminator 2000, BioRad) and gel pieces containing relevant DNA fragments were cut out with a scalpel and further purified using the QIAquick Gel Extraction Kit (QIAGEN) according to the manufacturer's protocol.

Vector dephosphorylation

To avoid recircularization of the pARCT5* vector, the 5'-end phosphates were removed by treatment with rSAP (NEB). In a PCR tube, 0.1 pmol of gel-purified pARCT5* was mixed with 2 μL of 10X rCutSmart™ buffer (NEB) and 1 μL of rSAP. The reaction was incubated at 37 ℃ for 30 minutes in a bead bath and then stopped by heat-inactivation at 65 ℃ for 5 minutes.

[14] Ethidium bromide is a DNA intercalating agent that fluoresces under UV light. It is used to visualize DNA bands in a gel. This compound is mutagenic and must be handled with care.

Ligation

The sticky-ended pARCT5* vector (dephosphorylated) and pLT insert were mixed at 1:2 and 1:3 molar ratios with 2 μL of 10X T4 DNA ligase buffer (NEB), 1 μL of T4 DNA ligase (NEB, added last) and milli-Q H_2O to a final volume of 20 μL. In addition, two control ligation reactions with the cut vector alone, either phosphorylated or non-phosphorylated (rSAP-treated), were set up. The reactions were incubated at 20 ℃ for 10 minutes in a bead bath and then heat-inactivated at 65 ℃ for 10 minutes. The amount of destination vector used in each reaction was approximately 50 ng, and the amount of insert was calculated according to (4):

$$ng\ insert = \frac{ng\ vector\ x\ kb\ insert}{kb\ vector} x \frac{molar\ ratio\ insert}{molar\ ratio\ vector} \quad (4)$$

The ligation mixtures were used to transform NEB DH5α *E. coli* competent cells according to the protocol described in section 3.1.2.

Diagnostic restriction digestion

Since the vector and insert were digested with a single restriction enzyme, the sticky ends are the same and the insert can be ligated in two different orientations into the vector backbone. To verify the insert orientation, plasmids isolated from several colonies were digested with BamHI restriction enzyme, which cuts the vector in two fragments with different sizes for each orientation (Figure 15). The digestion reactions were set up by mixing 200 ng of plasmid with 2 μL of 10X FastDigest (FD) green buffer (ThermoFisher Scientific), 1 μL of Fast digest BamHI enzyme (ThermoFisher Scientific) and milli-Q water to a final volume of 20 μL. The reactions were incubated at 37 ℃ for 10 minutes in a TC-3000 PCR thermocycler (Techne), followed by heat-inactivation at 80℃ for 5 minutes. The bands were visualized under UV-light (UV Transilluminator 2000, BioRad), and the plasmids that appeared to have the correct insert orientation (downstream of the promoter) were sent for sequencing.

3.1.2 Transformation

At the beginning of this project, OverExpress C43 cells (Sigma) and BL21 (DE3) pLysS cells (Novagen) were already transformed with the vectors coding for CT and PDI, respectively.

NEB DH5α and Turbo *E. coli* competent cells were transformed with pARCT5 and pLT-pARCT5 vectors. These vectors were also used to transform Vmax™ Express Chemically Competent Cells (SGI-DNA) for protein expression. The standard transformation protocols for these strains are described below.

E. coli transformation

Standard protocol: 1-5 μL of a solution containing 50-100 ng of plasmid DNA was added to a tube with 50 μL of DH5α or Turbo chemically competent *E. coli* cells. After 15 minutes of incubation on ice, the tube was gently flicked 4-5 times and then incubated for 15 more minutes. The cells were then heat-shocked at 42 ℃ for either 30 seconds (DH5α cells) or 45 seconds (Turbo cells) and immediately returned to ice for 5 minutes. Room temperature SOC medium (NEB) was added to the mixture at a 10:1 volume ratio, and the cells were incubated at 37 ℃ for 1 hour with shaking (180 rpm; Multitron Standard incubator-shaker, InforsHT) to allow them to recover and start expressing the antibiotic resistance marker. Two dilutions of the transformation reaction were spread onto pre-warmed LB-agar plates containing 25 mg/L chloramphenicol and incubated overnight at 37 ℃ (Kelvitron T incubator, Heraeus). The next day, single colonies were taken and used to inoculate 5 mL of LB-medium containing 25 mg/L chloramphenicol (CAM). The cultures were grown overnight at 37 ℃ with shaking (180 rpm, InforsHT) and used the following day for isolating DNA and making glycerol stocks.

Vibrio natriegens (Vmax) transformation

100-200 ng of plasmid DNA was added to one vial of Vmax™ Express Chemically Competent Cells (SGI-DNA). The tube was gently mixed and incubated on ice for 1 hour. The cells were then heat-shocked at 42 ℃ for 30 seconds and immediately returned to ice for 2-5 minutes. After incubation, the mixture was transferred to a pre-warmed 14 mL Falcon® tube containing 950 μL of Vmax™ Chemicompetent Cell Recovery Medium and incubated at 30 ℃ for 2 hours in a shaker (200 rpm, InforsHT). Then, two dilutions of the transformation reaction were spread onto LB-agar plates supplemented with 12.5 mg/L chloramphenicol and incubated overnight at 30 ℃ (Kelvitron T incubator, Heraeus). The next day, single colonies were taken to inoculate LB+ v2 salts medium[15] (25 mg/L chloramphenicol). The culture was grown overnight at 30 ℃ and used the following day for making glycerol stocks.

Glycerol stocks

Standard protocol: An aliquot of an overnight culture was diluted 1:1 with 50% glycerol (sterile-filtered) into a sterile 2-mL cryogenic vial. The tube was gently mixed and then stored at -80 ℃.

[15] Regular LB-Miller medium supplemented with additional salts including NaCl, KCl and $MgCl_2$. For a detailed composition of this medium, please see Appendix Section B.

Plasmid isolation and quantification

Standard Protocol: Plasmid DNA was isolated from overnight cultures of *E. coli* strains using the NucleoSpin Plasmid Kit (Machery-Natel), following the protocol for low-copy plasmids described by the manufacturer. The plasmid yield was quantified by measuring the absorbance at 260 nm with a NanoPhotometer nanodrop (Implen).

3.1.3 DNA Sequencing

Plasmids positive to the diagnostic digestion were sent for sequencing of the *pLT* gene region using the primers listed in the Appendix Section D (Eurofins Genomics). 10 μL samples containing 0.5-1 μg of plasmid DNA and 25 μM of primer were sent for sequencing to Eurofins Genomics. Sequences were visualized and aligned in the software MEGA 7.0.

3.2 Protein production

***Safety/Risks associated with toxin work**

E. coli strains for production of CT are classified as biosafety level 2. However, intoxication would require ingestion of mg-amounts of the toxin. For this reason, toxin production was performed in a BSL-1 laboratory, taking additional safety measures such as biological risk signs installed at shakers, centrifuges, and autoclave while they were in use, and vaccination of the researcher involved in the project (safety clearance for this work has been obtained from Helsedirektoratet, including flexible use of facilities).

3.2.1 Expression of CT in E. coli

50 mL Terrific Broth (TB) medium (Appendix Section B) containing 25 μg/mL CAM were inoculated with OverExpress C43 cells (Sigma) transformed with the CT-encoding pARCT5 plasmid. The culture was incubated overnight at 30 ℃ with shaking (150 rpm, Multitron Standard Incubator Shaker). The next day, 10 mL of pre-culture were transferred to Erlenmeyer baffled flasks containing 500 mL TB medium (1:50 dilution) supplemented with 25 μg/mL CAM and grown until the optical density at 600 nm (OD_{600}) reached approximately 2.0. Protein expression was induced with 0.2 % L-arabinose (Sigma) at 37 ℃, 150 rpm for 3 hours. Cells were harvested by centrifugation for 20 minutes at 6000 rcf (relative centrifugal force), 4 ℃ (Avanti J-26 XP centrifuge, Beckman- Coulter; J-Lite JLA-8.1000 fixed-angle rotor, Beckman-Coulter). To isolate CT from the periplasmic space, cell pellets were resuspended in 1/40th volume of TALON A buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl) supplemented with 1 tablet of cOmplete™ protease inhibitor cocktail per 50 mL lysis buffer (Roche), 3 μL of benzonase (Sigma) and 1 mg/mL polymyxin B sulfate[16] (Sigma). This solution

[16] PolymyxinB can disrupt the outer membrane of cells and has been used for releasing periplasmic proteins[162].

was incubated for 15 minutes at 37 ℃ with shaking followed by centrifugation at 8000 rcf, 4 ℃ for 20 min. The supernatant (periplasmic fraction) was filtered through a 0,2 μm filter (PES membrane, VWR) and used immediately for further purification steps (Section 3.3.1).

3.2.2. Expression of CT and pLT in Vmax

Vmax cells transformed with plasmids encoding CT or pLT were grown overnight at 30 ℃ and 180 rpm in LB+v2 salts medium containing 25 μg/mL CAM. The following day, the cultures were diluted 1:100 in 500 mL of LB+v2 salts medium (25 μg/mL CAM) and grown until OD_{600} reached ~0.8 before induction with 0.2% or 0.02% L-arabinose (Sigma) at 30 ℃, 140 rpm for 14-16 hours. Since CT is naturally secreted into the medium by Vmax, the cells were removed by 2-3 rounds of centrifugation at 15,000 rcf and 4 ℃ for 30 minutes. The supernatant was supplemented with 1 % NaN_3 to avoid bacterial growth, then filtered through a 0.22 μm filter (PES membrane, VWR) and used as soon as possible for further purification steps (Section 3.3.1).

3.2.3. Expression of human PDI in *E. coli*

Expression of hydrogenated PDI

LB medium containing 100 μg/mL ampicillin (Amp) and 20 μg/mL CAM was inoculated with BL21 pLysS cells transformed with the pOLR130 plasmid encoding human PDI. Cultures were grown overnight at 30 ℃ on shakers set to 140 rpm. The next day, the pre-culture was diluted 1:100 into flasks containing 1 L LB medium supplemented with 100 μg/mL ampicillin and grown at 30 ℃, 140 rpm. When OD_{600} reached 0.7-1.0, protein expression was induced overnight with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG; VWR) at 20 ℃, 140 rpm. Cells were harvested by centrifugation at 5000 rcf for 20 minutes and stored at -80 ℃ until further use. To isolate PDI from the cell cytosol, protein pellets were defrosted overnight in the fridge and then resuspended in His A buffer (50 mM Tris-HCl pH 7.5, 200 mM NaCl, 20 mM Imidazole; Appendix Section B) supplemented with 1 tablet of protease inhibitor cocktail (per 50 mL lysis bufferRoche), 200 μg/mL lysozyme (Sigma), 3 μL benzonase (Sigma) and incubated on ice for 20 minutes with stirring. For reduced PDI, 5 mM TCEP was also included in the lysis buffer. The solution was then sonicated at 40 % amplitude with 4 rounds of 20 second pulses, 40 seconds break, followed by centrifugation at 11,000 rcf, 4 ℃ for 40 minutes (Eppendorf 5810R bench-top refrigerated centrifuge; Eppendorf F-34-6-38 fixed angle rotor). The supernatant (cell lysate) was filtered through a 0.22 μm filter (PES membrane, VWR) and used immediately for purification, as described in Section 3.3.2.

PDI expression test in M9+ minimal medium

SANS studies involving a protein-protein complex require labeling of one of the components with deuterium. For the CT-PDI complex, we decided to work establish a deuteration protocol for PDI using minimal medium. This medium contains the minimum components that support bacterial growth and is used to incorporate specific isotopes in proteins. Since deuterated compounds are expensive, conditions for optimal expression of PDI were first tested in M9+ minimal medium without deuterium (Appendix Section B).

2.5 mL of LB medium containing 100 μg/mL ampicillin and 20 μg/mL CAM were inoculated with BL21 plysS cells and grown for 8 hours at 37 ℃ on a shaker set to 130 rpm. (InforsHT) These cultures were transferred to 25 mL of M9+ medium (100 μg/mL ampicillin, 20 μg/mL CAM) in 250 mL baffled flasks and grown for 15 hours on shakers at 37 ℃, 130 rpm. The next day, overnight cultures were transferred to 225 mL of M9+ media supplemented with 100 μg/mL ampicillin, and grown at 37 ℃, 130 rpm. Protein expression was induced with different concentrations of IPTG (0.25-1.0 mM) and at different OD_{600} values (0.7-5.0) for 20 hours at 25 ℃. The conditions tested are summarized in Table 2. To evaluate protein expression, cell samples before and after IPTG induction were taken and analyzed by SDS-PAGE as described in Section 3.3.3.

Table 2. Summary of expression conditions tested for overnight expression of PDI in H_2O-based M9+ medium

Condition	IPTG Concentration(mM)	OD_{600} at point of induction
1	0.25	0.7
2		1.2
3		1.6
4		5.0
5	0.5	0.7
6		1.2
7		1.6
8		5.0
9	1.0	0.7
10		1.2
11		1.6
12		5.0

Expression of perdeuterated PDI

To express perdeuterated PDI, BL21 pLysS cells were grown in D_2O-based M9+ minimal medium containing deuterated glycerol (Appendix Section B) following the protocol described for the minimal media expression test, except for one additional step: Cells were

initially grown in 1 mL of H_2O-based LB medium for 3 hours and then diluted 1:10 in 2.5 mL of D_2O-based LB medium, followed by incubation for another 5 hours. Protein expression was induced at $OD_{600} \approx 0.7$ with 0.5 mM IPTG (best condition found). By mistake, the temperature was not changed to 25 ℃ during expression but kept at 37 ℃. Deuterated PDI was extracted from the cell cytosol in the same way as it was described above for hydrogenated PDI.

3.3. Protein purification

All protein purification steps were carried out at 4-10 ℃ unless stated otherwise. Size-exclusion chromatography (SEC) was performed in an Äkta purifier or an Äkta pure protein purification system (GE Healthcare). All protein samples were filtered prior to SEC (Spin-X® centrifuge tube filters; Sigma)

3.3.1 CT and pLT purification

Affinity chromatography was carried out using a Miniplus 3 Peristaltic pump (Gilson) at room temperature[17]. Ion-immobilized affinity chromatography (IMAC) was used as the capture step for CT, since the toxin has two histidine amino acid residues that confer natural weak affinity for Co^{2+} or Ni^{2+} ions[128]. As pLT does not contain such residues (Appendix Section F), D-galactose affinity chromatography was used as capture step for this protein.

Talon affinity chromatography

CT was purified from the periplasmic fraction (*E. coli* expression) or the medium supernatant (Vmax expression) by loading these solutions into a HiTrap TALON® crude 5 mL column (Cytiva) previously equilibrated with Talon A. The column was washed with 15 column volumes (CV) of Talon A before eluting CT with 10 CV of Talon B (50 mM sodium phosphate pH 8,0, 300 mM NaCl, 50 mM imidazole). The protein was then concentrated as described in section 3.3.4 and stored at 4 ℃ until further use.

Galactose affinity chromatography

This method was used as main capture step for pLT and as an alternative for the purification of CT from the medium supernatant. While pLT purification was carried out by gravity flow, CT purification was performed in a peristaltic pump using an *in-house* packed column.

Standard protocol: Filtered supernatants were applied onto 3-4 mL Pierce™ D-Galactose Agarose resin (Thermo Scientific™) equilibrated with Gal A (50 mM sodium phosphate pH 7.4, 200 mM NaCl; Appendix Section B). The column was washed with 15 CV of Gal A, followed by

[17] Purification of the toxins was initially performed at 4 ℃, however, to avoid precipitation of LB+v2 salts medium components, the purification was carried out at room temperature.

protein elution with 10 CV of Gal B (50 mM sodium phosphate pH 7.4, 200 mM NaCl, 300 mM galactose; Appendix Section B). Proteins were concentrated to 2-8 mg/mL by ultrafiltration and stored at 4 °C.

Cation-exchange Chromatography (for CT produced in *E. coli*)

CT eluted by Talon affinity chromatography was dialyzed into 50 mM Tris-HCl pH 8 and loaded onto a HiTrap™SP (GE Healthcare) previously equilibrated with binding buffer. Elution was carried out with a linear gradient of 1 M NaCl. Fractions collected were analyzed by SDS-PAGE and further purified by SEC.

Size exclusion chromatography (SEC)

Standard Protocol: Concentrated proteins eluted by affinity chromatography were filtered with Spin-X® centrifuge tube filters (Sigma). 500 μL of filtered samples were loaded into a Superdex 200 Hiload 16/60 GL (CT produced in *E. coli*) or Superdex 200 increase 10/300 GL column (GE Healthcare) previously equilibrated with Phosphate-Buffered Saline (PBS) pH 7.4. Toxins were eluted with PBS buffer at a flow rate of 0.5 mL/min. Fractions containing pure protein were pooled, concentrated to 2-8 mg/mL by ultrafiltration and stored at 4 °C.

3.3.2 PDI purification

The same purification protocols were used for PDI and deuterated PDI. For the affinity and anion exchange chromatography steps, samples were loaded using a peristaltic pump (Miniplus 3 Peristaltic pump, Gilson). Protein elution was performed with an ÄKTA pure protein purification system (GE Healthcare). To keep PDI in a reduced state throughout the purification steps, a reducing agent was added to all the buffers and the protein was immediately frozen until use. Oxidized PDI was produced in the same way as reduced PDI, except that no reducing agent was included in the buffers.

His-tag affinity chromatography

Standard protocol: Cell lysates were loaded onto a 5 mL Histrap™ HP column (GE Healthcare) previously equilibrated with His A buffer supplemented with 5 mM TCEP (Sigma)[18]. After washing the column with 15 CV of His A buffer, elution was performed in three steps with HisA buffer containing 50 mM, 150 mM and 500 mM imidazole (Appendix Section B). PDI-containing fractions were pooled and dialyzed by two IEX A buffer (20 mM Bis-Tris, 5 mM DTT, pH 6.0) exchanges at 4 °C.

Anion-exchange chromatography

[18] DTT was initially used as reducing agent, but later changed to TCEP as it is more stable over time.

Optimized protocol[19]: Dialyzed fractions from the protein capture step were loaded into a 5 mL Hitrap™ Q HP column (GE Healthcare) previously equilibrated with IEX A buffer. After washing the column with 15 CV of IEX A, the protein was eluted with IEX B (20 mM Bis-tris pH 6.0, 1 M NaCl, 5 mM DTT) by the step-gradient shown in Figure 16. PDI-containing fractions were pooled, concentrated to 5-15 mg/mL, snap-frozen in liquid nitrogen and stored at -80 ℃ until further use.

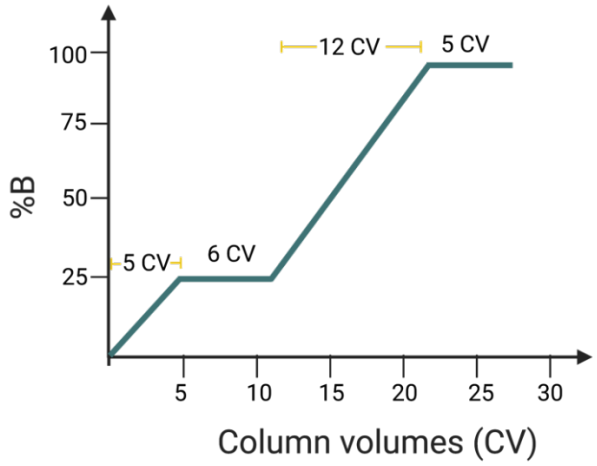


Figure 16. Step-gradient applied during PDI purification by anion-exchange chromatography. Figure generated in Biorender.com.

Size exclusion chromatography

Concentrated protein samples were loaded onto a Superdex 200 increase 10/300 GL column (GE Healthcare) equilibrated with either 20 mM Bis-Tris, 200 mM NaCl, 20 mM TCEP, pH 6.5 or PBS pH 7.4 containing 20 mM TCEP. Protein was eluted with a 0.5 mL/min flowrate and the PDI-containing fractions were pooled, concentrated to 5-10 mg/mL and stored at -20 ℃.

3.3.3 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Standard protocol: Samples were prepared by mixing an aliquot of protein with 4X Bolt LDS Sample Buffer (Invitrogen), 100 mM DTT (Applichem) and milliQ-H_2O to a final volume of 20 μL for 15 or 17-well gels, or 30 μL for 10-well gels. The samples were heated for 5 minutes at 95 ℃ and then loaded into NuPAGE™ Bis-Tris 4-12% polyacrylamide gels (Invitrogen, Appendix Section A) pre-assembled into a Mini Gel Electrophoresis Tank (Appendix Section A). SeeBlue Plus 2 ladder was also included in one of the wells. The tank was connected to a Power Supply-EP602 to apply an electric field of 200 V for 25-30 min. Gels were stained overnight with Coomassie staining solution and distained with milli-Q H_2O until the desired background was obtained.

[19] The initial protocol applied a linear gradient of 0-100% elution buffer over 20 CV.

For samples containing whole cells, the OD_{600} was measured and normalized to a value of 2.0 according to (5). The calculated aliquot of cells was centrifuged for 5 minutes at 12,000 rcf. The supernatant was discarded, and the cell pellet was resuspended in 30 μL of a solution containing 4X Bolt LDS Sample Buffer pre-mixed with 8 M urea at a 1:4 volume ratio and supplemented with 100 mM DTT.

$$Cell\ extract\ aliquot\ for\ OD_{600}\ normalization = \frac{30\ \mu L\ x\ 2.0}{OD_{600} sample} \qquad (5)$$

3.3.4 Protein concentration and quantification

Standard protocol: Proteins were concentrated by ultrafiltration (4 ℃, 3500 *xg*) using Amicon Ultra Centrifugal Filter Units tubes 10,000 MWCO (Merck) and Vivaspin500 tubes 5,000 MWCO (Sartorious). The protein concentration was determined by measuring absorbance at 280 nm with a spectrophotometer (Ultrospec 2100 pro, Amersham Biosciences), using the molar extinction coefficient (ε) of the protein calculated in the ProtParam Expasy website (https://web.expasy.org/protparam/; Appendix Section F).

3.3.5 Dialysis

Standard protocol: Proteins were dialyzed at 4 ℃ using a 10K MWCO SnakeSkin dialysis Tubing (Thermo Scientific) or Pure-A-Lyzer™ Midi 3.5K MWCO (Sigma) with magnetic stirring against the desired buffer. After 2 hours, the dialysis buffer was changed, and samples were dialyzed overnight.

3.4 PDI intra-crosslinking with EDC

In order to lock PDI in a rigid conformation, the protein was treated with EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride; Thermo Scientific™), a "zero-length" intramolecular cross-linker[72].

Standard protocol: PDI was dialyzed by two buffer exchanges against 50 mM MES, 150 mM NaCl, pH 6.5. 4 or 6 mg/mL of PDI was incubated with a 10-fold excess of EDC for 30 min at 20 ℃ in an Eppendorf Thermomixer® comfort (Appendix Section A) set to 300 rpm. To remove excess EDC and separate PDI from oligomers, the mixture was filtered and loaded onto a Superdex 200 increase column (GE Healthcare) equilibrated with PBS, pH 7.4. SEC fractions were analyzed by SDS-PAGE and those containing monomeric PDI were pooled, concentrated to 2 - 3 mg/mL, and stored short-term at 4 ℃.

3.5 CT cleavage with trypsin

Recombinant CT holotoxin produced in *E. coli* has an intact CTA polypeptide that can be cleaved *in vitro* into CTA1 and CTA2 fragments by limited trypsin digestion as described by Jobling *et al*[129].

Standard protocol: 4 mg of toxin were incubated with 1/10th of trypsin (Sigma; Appendix Section A) for 30 minutes at 30 ℃, followed by addition of 1 mM phenylmethylsulfonyl fluoride (PMSF) to stop the reaction. The samples were filtered and purified by SEC to remove trypsin. The integrity of the toxin was analyzed by SDS-PAGE.

3.6 CT/PDI complex formation test by SEC

Complex formation between CT and PDI was evaluated by SEC under three different conditions (Table 3).

Proteins were incubated separately with 20 mM TCEP or DTT for 30 minutes at room temperature and then placed on ice for at least ten minutes. After mixing the proteins at a 1:1 molar ratio, the sample was quickly loaded onto a Superdex 200 increase column (GE Healthcare) equilibrated with the corresponding buffer (Table 3). SEC fractions were analyzed by SDS-PAGE.

Table 3. Conditions evaluated by SEC for the formation of a CT/PDI complex

CT/PDI mixture	Buffer
Nicked CT + PDI **Un-nicked CT + PDI**	20 mM Bis-Tris, 200 mM NaCl, 20 mM TCEP, pH 6.5
Nicked CT + EDC-treated PDI	PBS, pH 7.4, 20 mM DTT

3.7 Small-angle X-ray scattering (SAXS)

3.7.1 In-house data collection

Preliminary SAXS measurements of the individual proteins were performed in a Bruker NanoStar SAXS instrument at the University of Oslo. Data was collected at a wavelength of 1.54 Å. Protein samples and matching buffers were measured for 2 hours each. Samples were treated with TCEP or GSH for 30 min at room temperature right before the measurements. CT (un-nicked) was measured at 2.4 mg/mL in 20 mM Bis-Tris, 200 mM NaCl, 20 mM TCEP, pH 6.5 and at 2.1 mg/mL in 1X PBS, 20 mM GSH, pH 7.4. PDI was measured at 1.4 mg/mL in the same Bis-Tris buffer. Measurements in Bis-Tris buffer were performed at 10 ℃, whereas those in PBS buffer were done at 37 ℃. Intensities were collected as a function of the scattering vector q in the range 0.009 to 0.3 Å^{-1}. Data processing was performed with the

SUPERSAXS program package[130]. The scattering from the protein-free buffer was used for background subtraction.

3.7.2 Synchrotron data collection

Protein samples were measured at the beamline BM29 at the European Radiation Synchrotron Facility (ESRF) in Grenoble, France[20]. Data were collected at a wavelength of 0.992 Å and scattering intensities were recorded as a function of the scattering vector, within the q-range 0.00449-0.5187 $Å^{-1}$. Measurements were performed at 10 ℃ or at 37 ℃ in the following buffers:

1) 20 mM Bis-Tris, 200 mM NaCl, pH 6.5
2) 1X PBS, 20 mM GSH, pH 7.4
3) 1X PBS, 2 mM GSH, pH 7.4

Samples measured under reducing conditions were treated with TCEP or GSH for 30 min at room temperature right before the measurements. The scattering of individual proteins in oxidizing conditions (no TCEP or GSH treatment) was also measured.

Unnicked CT was measured at three different concentrations (0.5-2 mg/mL) in the three different buffers outlined above. Nicked toxin was measured at two concentrations (2 mg/mL and 0.5 mg/mL) in only Buffer 1.

PDI was measured at different concentrations three different concentrations (0.33- 2 mg/mL). EDC-treated PDI and Q3R-treated PDI[21] were measured in Buffers 2 and 3, respectively. Q3R-treated PDI samples were prepared *in situ* by mixing the protein with 0.5 mM Q3R, followed by incubation at room temperature for at least 10 minutes before the measurements.

The CT-PDI samples were also prepared *in situ* by mixing the two proteins at a 1:1 molar ratio, before incubation at 10 ℃ or 37℃. A detailed list of all the conditions tested can be found in Section J of the Appendix.

3.7.3 SAXS data analysis

Guinier and Porod analyses were done using PRIMUS[131] from the ATSAS software package[132]. *Ab initio* models of the individual proteins were generated with DAMIFF[133], averaged with DAMAVER [134] and refined with DAMMIN[135]. The crystal structures of CT and

[20] All samples were fast-frozen in liquid nitrogen and sent on dry ice.

[21] In addition to the buffer components described in the main text, samples containing Q3R have 2% DMSO to solubilize the inhibitor.

PDI were fitted to the SAXS data using Pepsi-SAXS[136]. The resulting atomistic models were superimposed with the *ab initio* model using SUPCOMB[137].

The molecular mass of the proteins was calculated from the Guinier analysis according to (6)

$$MM = \frac{N_A \; I\,(0)}{c\Delta\rho_M^2} \qquad (6)$$

Where MM is the molecular mass, c is the concentration, N_A is Avogadro's constant and ρ_M is the scattering contrast per mass, assumed to be $2.17\text{x}10^{10}$ cm/g (similar value among proteins).

To evaluate if the experimental scattering plots of the CT-PDI mixtures might indicate formation of a complex, the scattering of the individual proteins (at the same concentrations as in the mixture) were added, plotted as a function of q and compared to the experimental CT-PDI plots.

In addition, to get insight into how the theoretical scattering of a possible CT-PDI complex may look, models of PDI docked to CT were calculated using the web servers ClusPro[138] and HDOCK[139]. Models where PDI was bound to the bottom of the CT pentamer were filtered out, as this orientation may not facilitate toxin disassembly. The theoretical scattering of the remaining models was calculated with Crysol[132] (ATSAS package), and compared with the experimental scattering curves. Data plots were generated with MATLAB® (R2021a) and figures of structural models with PyMOL.

3.8 CT crystallization

CT was crystallized by sitting-drop vapor diffusion using a crystallization robot (Oryx 4, Douglas Instruments). A 48-well customized screen based on conditions where the protein has been previously crystalized[41] was designed (Appendix Section G). Reagents were dispensed using a Formulator® Liquid Handler (FORMULATRIX). 2 μL drops were set up by mixing CT (5.76 mg/mL) in buffer G (50 mM Tris, 200 mM NaCl, 1 mM EDTA, 3 mM NaN_3, pH 7.4) with the reservoir solution at a 1:1 volume ratio. The crystallization plate was stored at 20 °C in a Formulatrix RI-182 0056 crystal hotel. Drop images were taken with the RockImager software (Formulatrix). Small crystals were visible after one week.

One of the crystals from condition B4 was fished, transferred to 20 % glycerol (cryo-protectant) in the corresponding mother liquor, mounted in a loop and flash-cooled with liquid nitrogen. The crystal was kept in a Dewar (Taylor-Wharton CX-100) for future data collection.

4. Results and discussion

4.1 Toxin cloning and production

4.1.1 CT production in *E. coli*

The first part of this thesis focused on producing substantial amounts of pure CT for SAS experiments. At the beginning of the project, a protocol to express CT as a periplasmic protein in Overexpress™ C43 *E. coli* cells had already been established (PhD thesis of Joel Bejamin Heim, UiO). This strain was chosen as it had proven to be more effective in expressing toxic proteins than other *E. coli* strains[140]. Cells were grown until OD_{600} reached 2.0, before addition of L-arabinose to induce protein expression. This is a relatively high OD_{600} value for induction, however, due to the toxic nature of CT, higher cell densities were needed to obtain significant amounts of protein. After induction, two clear bands at around 28 kDa (CTA) and 11 kDa (CTB) were observed by SDS-PAGE analysis (Figure 17.A), confirming expression of the toxin. CT was then purified from the *E. coli* periplasmic space by Talon affinity chromatography. The protein eluted upon addition of 50 mM imidazole. Although most of the contaminants were removed (Figure 17.B), the CT holotoxin co-purifies with free CTB pentamer (referred in this section as simply CTB unless stated otherwise). This is difficult to observe by SDS-PAGE alone, however, CTB contamination has been reported previously by Jobling *et al.*[129] and Krengel's lab (UiO). CTB excess may be due to degradation of several copies of the unstable CTA subunit before they assemble with the CTB subunits in the periplasm or incomplete assembly.

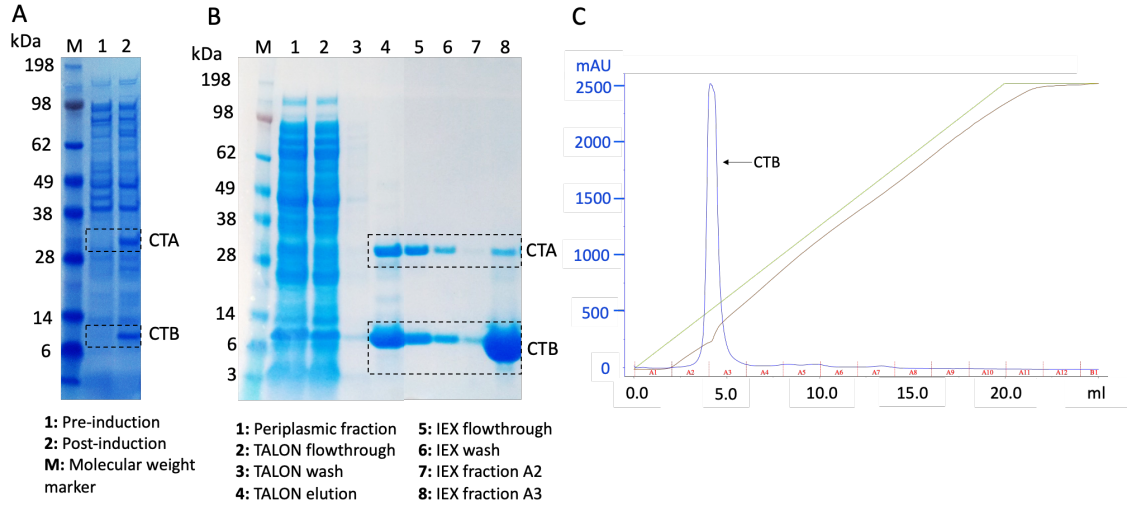


Figure 17. Production of CT in *E. coli* and purification. (A) SDS-PAGE analysis of cell samples before and after induction of CT expression. Under denaturing conditions, CT runs in SDS-PAGE as two bands: one at around 28 kDa corresponding to the CTA subunit (CTA1+CTA2) and the other one at approximately 11 kDa, corresponding to individual CTB subunits. (B) SDS-PAGE of protein samples collected during purification of the toxin by Talon affinity chromatography and IEX. (C) IEX chromatogram. At pH 8.0, most of the cholera holotoxin does not bind to the IEX column and elutes in the flow-through and wash, allowing separation from free-CTB pentamers. Elution was monitored

by absorbance at 280 nm (blue line). Green line: Linear gradient of buffer B concentration. Brown line: Conductivity. Red: Collected fractions.

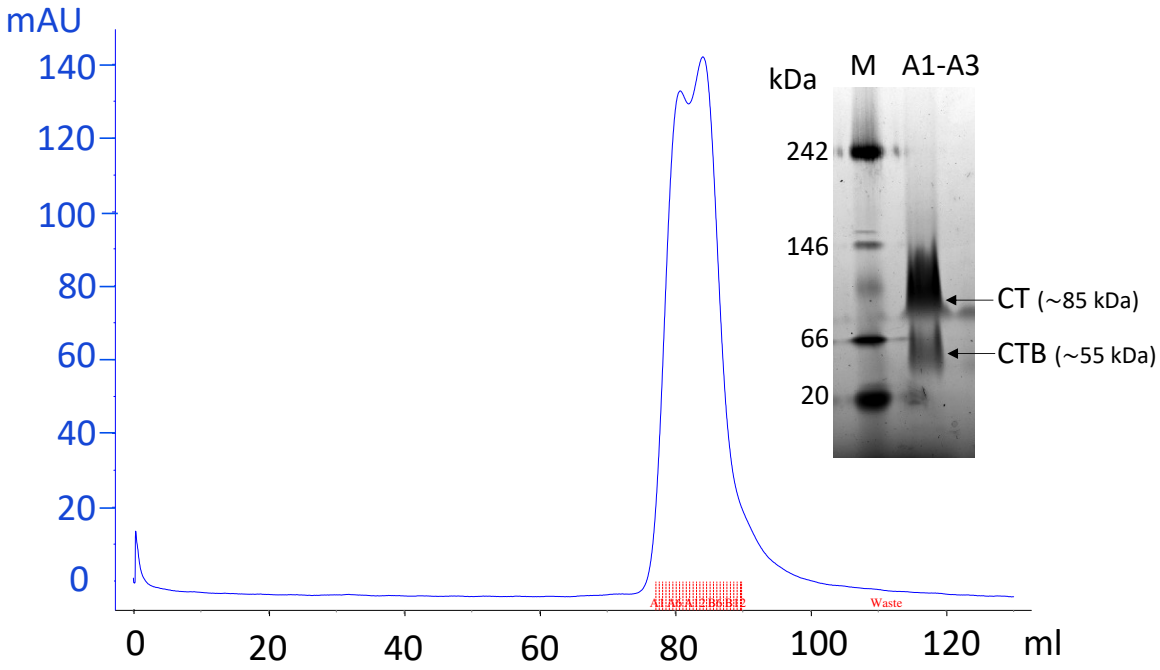


Figure 18. SEC purification of CT from IEX flow-through and wash. (A) SEC chromatogram and (B) native PAGE analysis of fractions from first SEC peak. Contamination of the B-pentamer in CT fractions is still visible.

To remove contamination by CTB, the toxin was further purified by cation exchange chromatography followed by SEC. Based on the difference in the theoretical isoelectric points (pI) of the CT holotoxin (pI=7.24) and the CTB (pI=8.48), a buffer at pH= 8.0 was chosen for the IEX binding and elution steps. The CTB-pentamer will carry a positive charge as its pI is above the pH and is expected to bind the cation exchange column. In contrast, the CT holotoxin will carry a negative charge and thus elutes in the flow-through. Although most of the CTB pentamers were removed by IEX (Figure 17.B and 17.C) when the CT-containing IEX flow-through and wash were loaded onto a SEC column, two overlapping peaks were obtained (Figure 18.A), suggesting that some contamination by CTB remained.. The left peak corresponds to the cholera holotoxin since it has a larger hydrodynamic radius. Therefore, fractions A1-A3 from the first half of this peak were pooled and ran on a native PAGE to evaluate the purity of the holotoxin. This is a better alternative than SDS- PAGE, as the proteins keep their oligomeric states. Although the gel bands appear smeared, two bands with matching molecular weights for CT and CTB-pentamer are clearly visible (Figure 18.B), confirming that free pentamers were not removed completely. Despite obtaining ~6 mg of protein per L of growth medium after the Talon step, the final holotoxin yield was no more than 0.2 mg per L of medium. Considering that protein purity is very important for SAS experiments, high amounts of holotoxin had to be discarded because they still contained CTB.

Several trials for optimizing the CT purification protocol were performed, but none of them significantly improved separation from free CTB pentamers. First, we replaced the cation exchange chromatography step by anion exchange, where at pH 8.0 the holotoxin is expected to bind to the column while CTB pentamers should not bind. Nevertheless, both proteins eluted in the flow-through and wash. There are several possible explanations: an uneven distribution of the surface charges on CT which may prevent the protein from binding to the resin; a defective column; using a pH in the buffer that is only few fractions of a unit above from the pI of CT[22] (if the theoretical value is very different from the experimental one). It is recommended to run isoelectric focusing (IEF) electrophoresis to determine the experimental pI of CT and CTB and hence choose an appropriate pH for the purification. Also, fresh columns or columns with other anion exchanger resins could be used.

The other two optimizing attempts were carried out for the SEC step. Those included applying a slower flow rate during the runs, which may allow for better separation of the components, and using a Superdex 200 Increase column (GE Healthcare), which has smaller resin beads and therefore confers higher resolution than the Superdex 200 16/60 Hiload column initially used[141]. However, these changes did not allow separation of overlapping CT and CTB peaks.

Although there is still room for further optimization of the purification protocol, we decided to try a different approach: expressing the protein in another bacterial host. This is discussed in the following section.

4.1.2 CT production in Vmax improved protein yield

As the main aim of this thesis was to investigate the interaction between CT and PDI by SAXS and SANS, mg amounts of both proteins were required. Production of CT in *E. coli* resulted in low protein yields, which were not high enough to support the experiments without significant scale-up. For this reason, we decided to try expression of the toxin in another non-pathogenic Gram-negative bacterium: *Vibrio natriegens* (Vmax). This is the fastest-growing known organism to date, with a growth rate that is twice as fast as *E. coli* and has recently been used for the expression of several proteins with significantly higher yields compared to production in *E. coli*[142,143]. Vmax is part of the *vibrio* genus and is therefore more closely related to *V. cholerae* than *E. coli*. One trait that both vibrios share is the type II secretion system (T2SS), a protein apparatus that enables secretion of proteins to the extracellular milieu. CT is naturally secreted by the *V. cholerae* T2SS, and we thus hypothesized that CT could also be secreted into the medium by the Vmax T2SS. In this thesis, a protocol for the expression of CT in Vmax was established, resulting in at least a 10-fold increase in protein yield.

[22] It is suggested that the pH of the buffers for IEX should be 0.5-1 units above or below the pI of the protein.

Since the pARCT5 plasmid contains several elements compatible with Vmax, including the p15A origin of replication[142,144] and the araBAD promoter, we transformed competent cells with the same plasmid. As a halophilic bacterium, Vmax requires the presence of higher concentration of salts in the medium for appropriate growth. Cells were therefore grown in LB-medium supplemented with additional salts (Appendix Section B) as recommended by the manufacturer[142]. For the initial expression test, cells were grown until OD_{600} reached ~0.8 before induction with 0.2% w/v L-arabinose overnight. Expression of the protein was not visible in the pre- and post-induction cell samples, nor in the supernatant (Figure 19.A). This was not surprising, as the protein was expected to be secreted to the growth medium, and it may be too diluted in the medium to be clearly seen by SDS-PAGE. To check if the protein was produced, 500 mL of supernatant were directly loaded onto a Talon column and the elution fraction was analyzed by SDS-PAGE. Two clear bands matching the sizes of CTA and CTB were observed (Figure 19.A), confirming successful expression of the toxin. CT appears already very pure after this step, nevertheless, the protein was further purified by SEC to remove contaminants that may not be visible from the gel. Interestingly, a single sharp peak was obtained (Figure 19.B), suggesting that there is no contamination by CTB. Since both the holotoxin and CTB-pentamers can be transported with equal efficiency to the extracellular medium by the T2SS[145], the lack of CTB excess may be due to a better transcriptional control in Vmax of the *ctxAB* operon or to less degradation of CTA, as the expression in Vmax is carried out at 30 °C compared to 37 °C in *E. coli.*

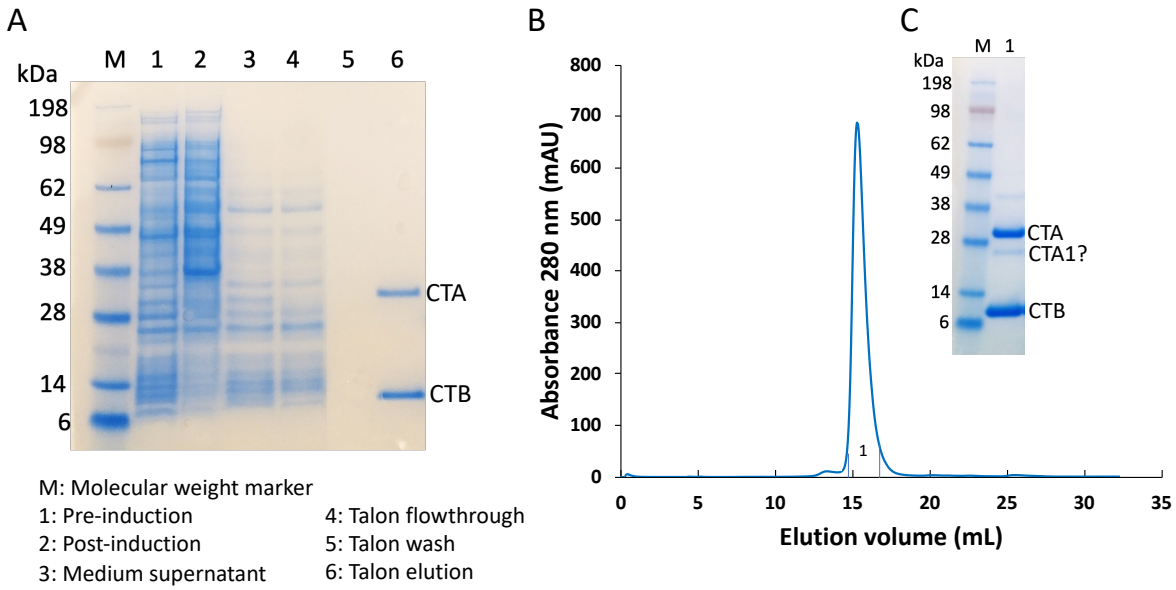


Figure 19. Production of CT in Vmax and purification by Talon affinity chromatography followed by SEC. (A) SDS-PAGE analysis of protein samples obtained throughout CT expression and purification by Talon. (B) Size-exclusion chromatogram of CT from Talon Affinity Chromatography and (C) SDS-PAGE of CT from SEC peak (lane 1). "M" denotes molecular mass marker.

Another aspect to discuss is the two additional faint bands observed in the SEC fractions (Figure 19. C). The one between 28 and 17 kDa may correspond to CTA1, whose size is 22 kDa,

while the band above 38 kDa may be a contaminant. As CT was expressed in a *vibrio* species, we hypothesized that CTA could be partially cleaved into CTA1 and CTA2 by Vmax proteases, in a similar way as it happens to CT when it is produced by *V. cholerae*. Both bands were analyzed by the Proteomics Core Facility at Oslo University Hospital and fragments of CTA1 and CTA2 were identified in both bands (Appendix Section I). Such results were puzzling, as the gel was run under reducing and denaturing conditions that separate the toxin into its individual subunits, which cannot run as higher molecular weight species. During LC-MS analysis, carry-over effects could happen if the measured species have a strong affinity for the column used and remain bound even after washing steps. This could be an explanation for why the top unknown band was also identified as CTA. Nevertheless, we do not have a logical explanation for the band around 22 kDa. We recommend repeating the MS analysis of these two bands to confirm if partial cleavage occurred.

The yield of the first CT production in Vmax was 12 mg of protein per 500 mL of growth medium (24 mg/L), however, when the protocol was reproduced several times, yields varied significantly, between 2 mg to 10 mg of protein per 500 mL of medium (4-20 mg/L). Initially, we hypothesized that the variations could be due to saturation of the column with cell debris that was not completely removed by centrifugation, therefore the column was stringently washed and regenerated after each purification. As an alternative to Talon affinity chromatography, we also used Galactose affinity chromatography (GAC) to capture CT from the supernatant. We were able to recover 25% more protein than with Talon Chromatography (3.8 mg vs 3 mg per 500 mL medium) and from this point onwards, GAC became the preferred method to capture CT. Alternatively, the protein could be precipitated with ammonium sulfate, which would help to separate the protein from cell debris and reduce the sample volume loaded on the column.

Although column regeneration and GAC slightly improved the yield, we still experienced yield variations between different rounds of expression. Such inconsistences appeared to be related to the expression rather than the purification itself. Similar observations have also been reported for expression of other toxic proteins and may be attributed to formation of heterogeneous populations of cells with different expression levels. This protocol can be further improved by investigating the effects of different concentrations of L-arabinose in the expression as well as different induction temperatures. In addition, protein production in Vmax has been improved by using codon-optimized sequences[146], therefore, codon-optimization of the CT gene could potentially improve the protein yield.

Although improvements can still be made, we established a new protocol that yields sufficient amounts of pure protein for structural studies.

4.1.3 pLT cloning and production

pLT was subcloned into the pARCT5* vector by restriction-enzyme based cloning. Double cuts generated by digestion with the enzyme NcoI. After gel-purifying the bands of the pARCT5 backbone and pLT insert, the vector was dephosphorylated to remove 5'-end phosphates and prevent its recircularization. The ligation was carried out using a 1:2 and 1:3 vector: insert ratios and colonies were obtained from both ligation reactions. Since both sticky-ends are the same, the insert can be ligated into the vector backbone in two different orientations. Therefore, we verified the insert orientation in plasmids isolated from several colonies by a diagnostic restriction enzyme digestion with BamHI. This enzyme cuts the vector in two fragments, whose size depends on the orientation of the insert as shown by the scheme in Figure 20.B. Six different colonies appeared to have the right insert orientation, as two bands of the expected fragment sizes were observed in the gel (Figure 20.A). To verify the exact sequence of the insert and its orientation, plasmids isolated from three positive colonies were sent for sequencing of the insert region and only one of them had the correct pLT sequence (Appendix Section E). The other two turned out to be the original pARCT5 plasmid, which might have been carried over as a contaminant when the bands were cut out from the NcoI restriction gel. One positive colony was enough to transform cells and proceed to express the toxin.

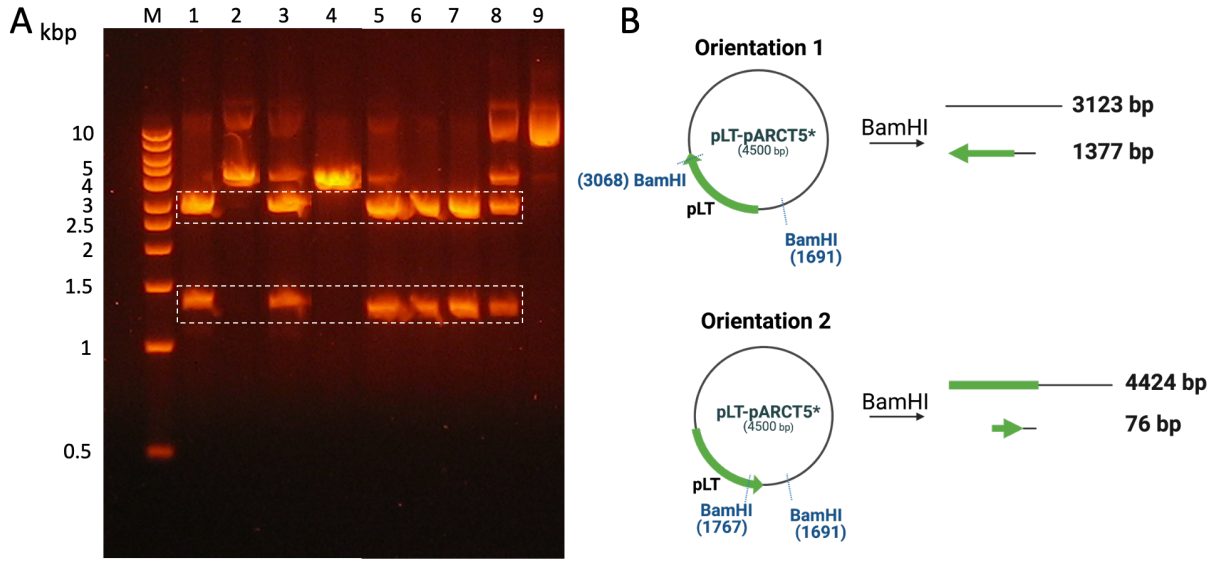


Figure 20. Evaluation of pLT insert orientation by digestion with BamHI. (A) Agarose gel electrophoresis containing the BamHI digestion products of plasmid isolated from 8 different colonies. (B) Representation of two possible DNA fragments obtained after digestion for each plasmid orientation. The dashed white rectangle highlights the DNA bands that might correspond to the fragments obtained for orientation 1 (downstream of the promoter). M=DNA molecular weight marker. Lane 1-8: Plasmid isolated from 8 different colonies. Lane 9: Uncut plasmid isolated from one of the colonies.

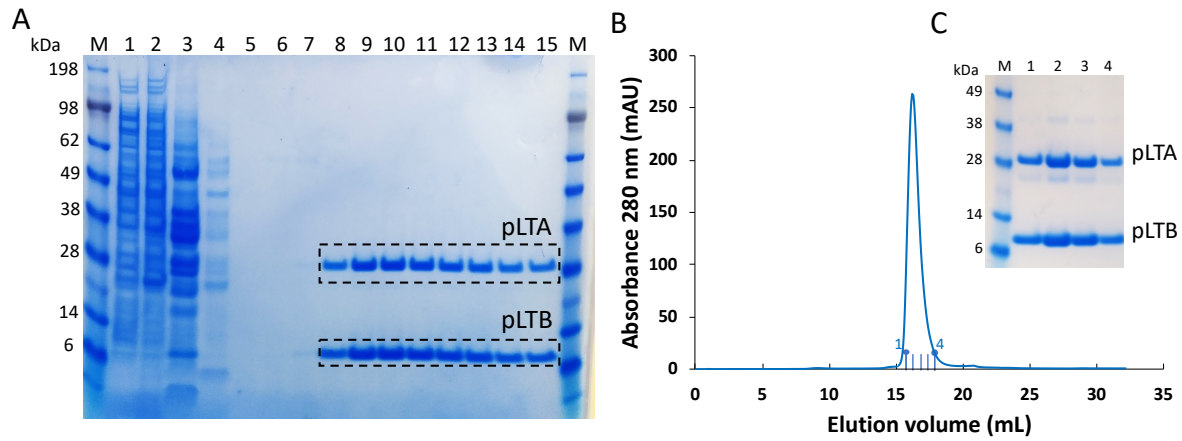


Figure 21. Expression of pLT in Vmax and purification. (A) SDS-PAGE analysis of protein samples obtained throughout the expression of pLT and purification by galactose affinity chromatography. Lane 1: Pre-induction sample. Lane 2: Post-induction sample. Lane 3=Periplasmic extract. Lane 4: Galactose affinity chromatography (GAC) flow-through. Lane 5: GAC wash. Lane 7-15: GAC elution fractions. (B) Size-exclusion chromatogram of pLT from pooled galactose fractions and (C) SDS-PAGE analysis of fractions 1-4 from SEC.

Due to the close sequence homology between pLT and CT, the same Vmax protocol developed for expression of CT was applied to pLT, except that only a 250 mL main culture was grown. The protein was successfully produced in the first trial, as seen by the two gel bands at around 28 kDa and 11 kDa (Figure 21.A). Mass spectrometry analysis of these two gel bands performed by the group of Bernd Thiede (University of Oslo) confirmed that the protein corresponded to pLT (Appendix Section I). The protein yield was 1.8 mg per 250 mL of growth medium (7.2 mg/L of medium), which is comparable to the yield obtained for CT. pLT was further purified by SEC and a single peak was obtained (Figure 21.B). The protein appears very pure by SDS-PAGE, although two other faint bands are present, which could be further investigated by MS analysis.

4.2 CT crystallization

In order to validate the correct folding, purity and homogeneity of the protein obtained by the new Vmax expression protocol, we decided to crystallize CT. We designed a 48-well customized screen based on the two conditions where the protein has been previously crystallized [41] (Appendix Section G). The variables that changed across the custom screen were the concentration of PEG 3350, $Mg(OAc)_2$ and sodium citrate dihydrate dibasic. To reproduce the same conditions, CT was dialyzed to buffer G (50 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA, 3 mM NaN_3) and concentrated to 5.7 mg/mL.

Initially, many small droplets in several conditions were obtained, indicating phase separation. Small crystals were formed after 7-14 days of incubation at 20 °C in conditions B4 and C1(Figure 22). These conditions contain the same PEG concentration (24 %) and similar Mg $(OAc)_2$ concentrations (100 mM and 125 mM, respectively), which can serve as a starting

point for setting up a fine-tuning optimization screen. These crystals are thin and do not have clear sharp edges, which could indicate that they are protein and not salt crystals. There are several methods to confirm this, including crashing the crystals (if they break easily, they are likely protein crystals), redissolving the crystals in their mother liquor and running SDS-PAGE (if the characteristic CT bands are observed, the crystals contain protein) and the final test: obtaining the X-ray diffraction pattern of the crystal (protein crystals have closely spaced diffraction spots whereas salts have more separated spots)[147]. One of the crystals from condition B4 (pointed out by the arrow in Figure 22.B) was fished, flash-cooled and stored in a dewar for future data collection at ESRF. Obtaining the diffraction pattern will confirm that protein crystals were obtained. This would be a good indication that the protein produced with the new Vmax protocol is highly pure and homogeneous, which is necessary to form a well-organized crystal lattice.

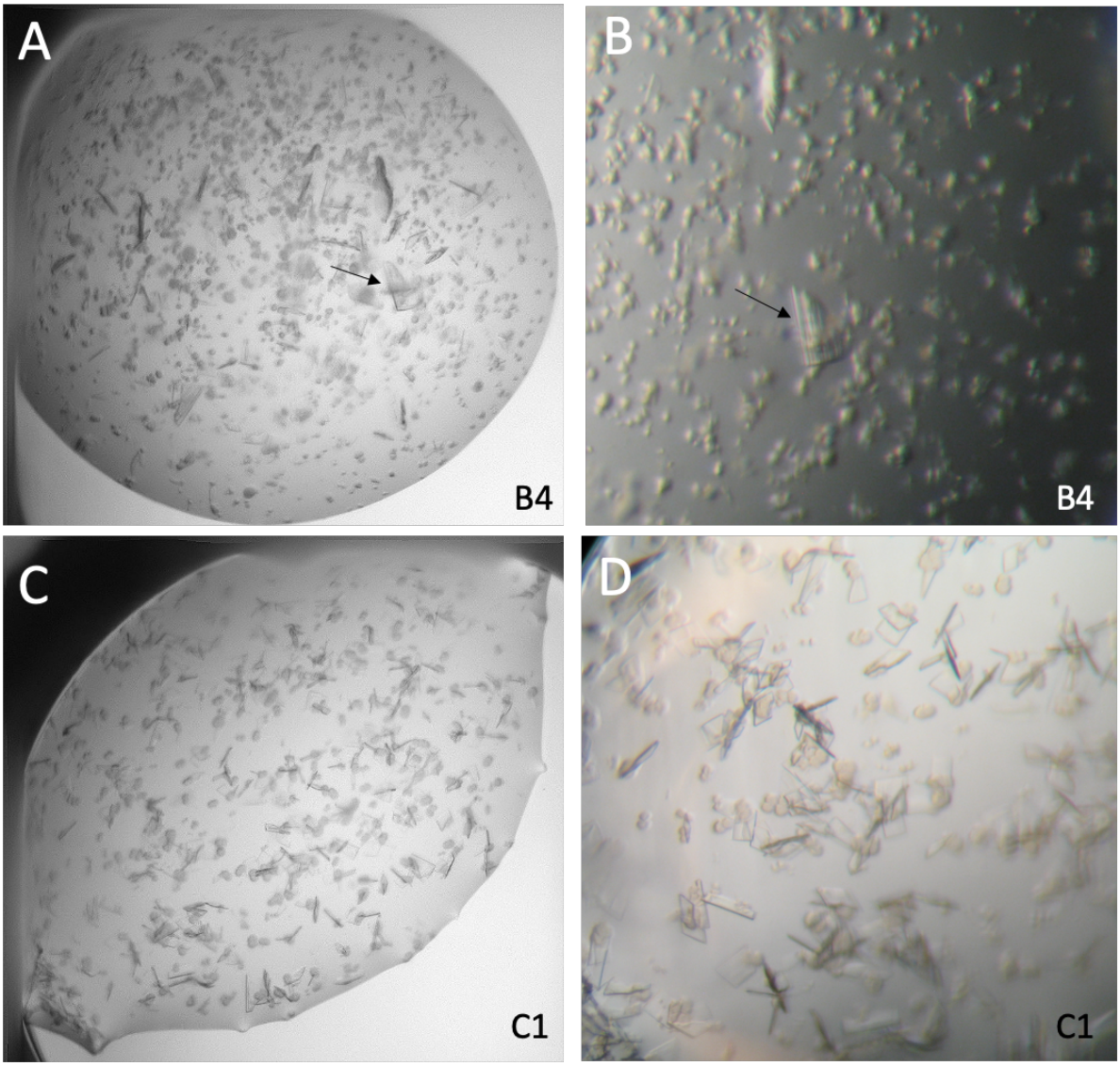


Figure 22. Crystals formed in well B4 (A, B) and well C1 (C, D) in the CT customized screen. The black arrow points at the crystal fished for future data collection.

Nevertheless, the plate-like aspect of the crystals suggests that there may be imperfections in crystal packing as the lattice does not extend easily in depth, which can affect the quality of X-ray diffraction. In addition to setting-up a fine-tuning screen, the small crystals could be used as seeds for growing larger crystals.

4.3 PDI expression and purification

4.3.1 Production of hydrogenated PDI

At the beginning of the project, a protocol to produce PDI in pLysS cells was already established by Joel Benjamin Heim (Department of Chemistry, UiO), however, some of the purification steps were modified in this thesis to improve protein purity. Human PDI carrying an N-terminal His-tag was produced in BL21 pLysS cells as described in the Methods section. SDS-PAGE analysis of the cytosolic extract showed a very strong band at around 56 kDa, which corresponds to the molecular weight of PDI and was later confirmed by MS analysis. The protein was purified from the cytosolic material by IMAC. Effective elution of PDI was achieved with 150 mM imidazole (Figure 23), however, due to the high amount of protein produced and loaded onto the IMAC column, a significant amount eluted in the IMAC wash, suggesting that the binding capacity of the column was exceeded. It is recommended to measure the protein concentration in the cell lysate prior to loading extracts in the column. Preferably, the cytoplasmic extract should be divided and purified by 2 or 3 batches of IMAC. The amount of protein after this step was around 15 mg per L of expression medium, which is a very efficient yield. PDI still contained several impurities that were further removed by anion-exchange chromatography followed by SEC.

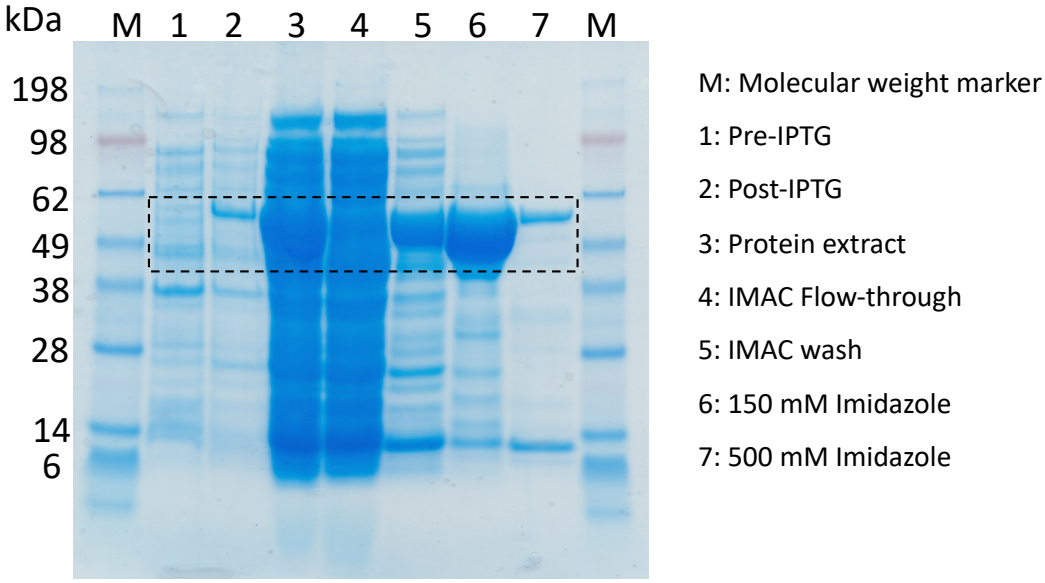


Figure 23. PDI expression in *E. coli* and purification by IMAC. SDS-PAGE analysis of cell samples obtained throughout purification of PDI by IMAC.

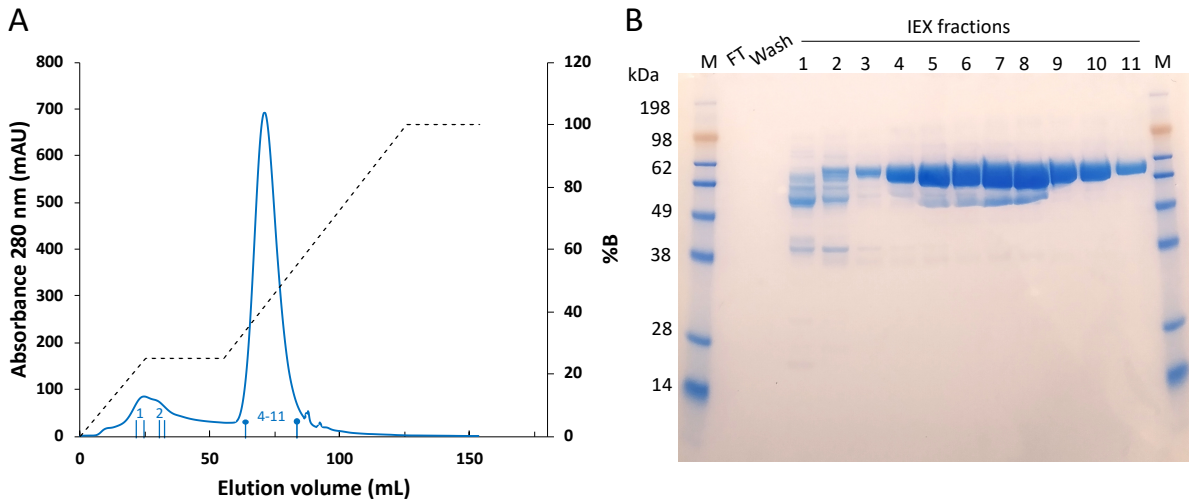


Figure 24. PDI purification by IEX. (A) IEX chromatogram of pooled PDI samples from IMAC purification. (B) SDS-PAGE analysis of IEX protein fractions.

At first, the protein was eluted with a linear gradient of NaCl from the IEX column, however, several contaminants from a small overlapping peak eluted together with PDI (Figure S7. Appendix Section H). By introducing a plateau of 25% NaCl over 6 CV, the separation of this peak was significantly improved (Figure 24.A) and PDI eluted almost free of contaminants as seen by SDS-PAGE (Figure 24.B). PDI-containing fractions were pooled and further purified by SEC, resulting in two peaks whose relative height varied between different preparations of the protein (Figure 24.A and 24.C). These peaks correspond to two different species of PDI, likely dimers (early peak) and monomers (later peak), which have been previously reported by SEC analysis[23][148,149]. SDS-PAGE confirmed that the main component of both peaks was PDI (Figure 24.B and 24.D). Even though two other faint high molecular weight bands were also observed in fractions from the first peak, they were identified as PDI by MS analysis (Appendix Section I). Oligomers are not seen in SDS-PAGE under reducing conditions unless they are covalently linked. The fact that we still identified higher molecular weight bands could be explained if the amount of reducing agent was not sufficient to cleave the disulfide bonds that may be forming between PDI molecules, giving rise to oligomeric species. Although PDI dimerization was initially attributed to formation of intermolecular disulfide bonds between the cysteines of its thioredoxin domains (a and a')[150], more recent data suggested that dimerization is mainly mediated by hydrophobic interactions between b' domains[151,152]. Yu *et al.*[149] reported observing high molecular weight bands by SDS-PAGE without reducing agent, thus we do not rule out that intermolecular disulfide bonds may also contribute to dimerization.

[23] The cited studies referred to these peaks as dimers and tetramers instead of monomers and dimers. This is because PDI appears as a larger protein on SEC due to its elongated shape and it was for a long time believed that the protein was a dimer. However, it has been confirmed to exist mainly a monomer by analytical ultracentrifugation and SAXS[159].

Another aspect to discuss is the peak profiles observed between different preparations. When PDI was purified from fresh IEX fractions, the dimer peak was smaller. However, when it was purified from IEX fractions stored for a prolonged period at 4 ℃ or -20 ℃, the dimer peak was larger (Figure 24.A and 24.C). This was observed for reduced PDI, but we did not keep track of oxidized PDI. It thus appeared that protein storage influenced the proportion of monomeric and dimeric forms of PDI, which has also been previously observed during PDI preparations by Morjana *et al.*[153]. Nevertheless, studies have shown that dimerization can be reversed at 37 ℃[149], suggesting that these dimers are held together by relatively weak forces, e.g. hydrogen bonds and hydrophobic and ionic interactions, which can be disrupted when the temperature is increased. The presence of reducing agents can also have an effect, as disulfide formation may have a role on PDI dimerization. It would be advisable to store the protein at -20 ℃ long-term or at 4 ℃ for short-term storage. If the monomeric form is required for experiments, prior exposure to 37 ℃ may reverse possible dimerization.

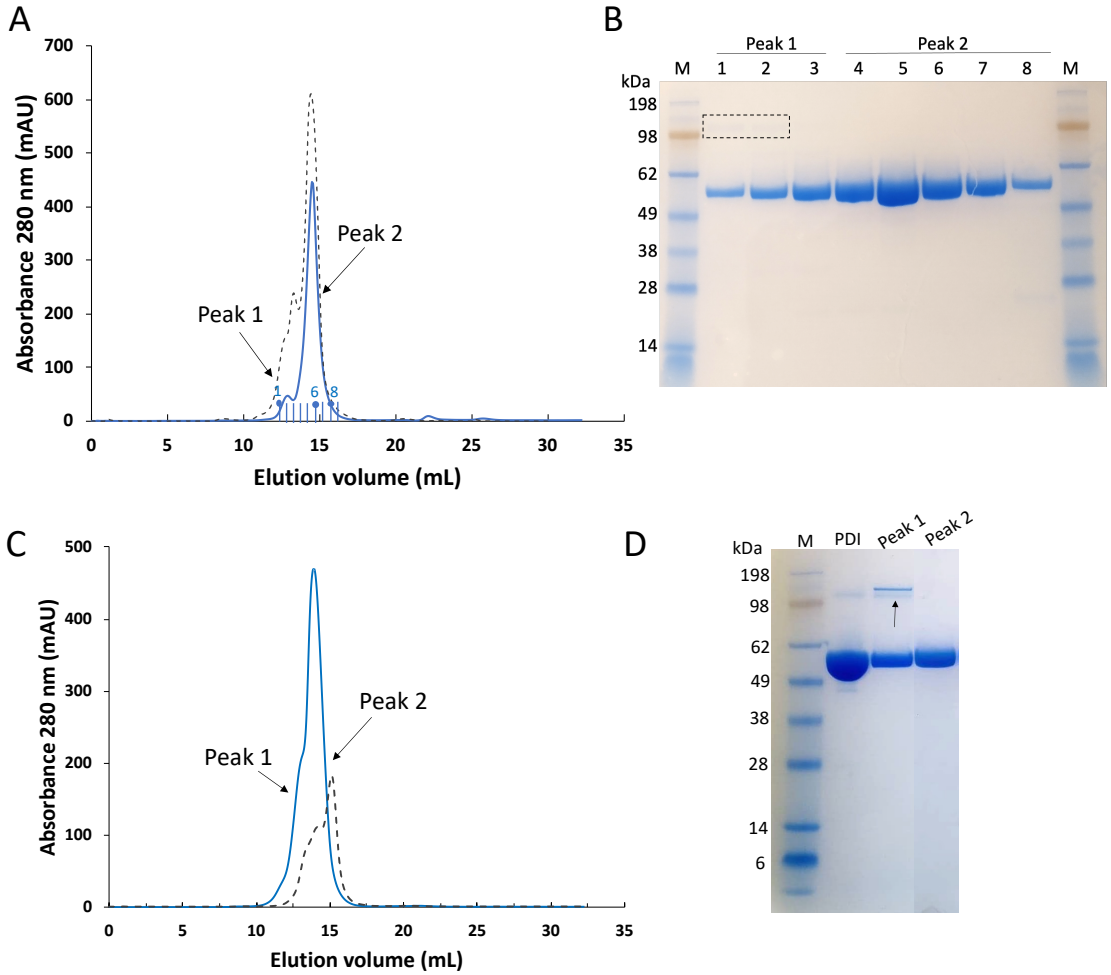


Figure 24. SEC purification of reduced PDI and oxidized PDI. Size-exclusion chromatograms of reduced (A) and oxidized (C) PDI. Dashed line: elution profile obtained from a sample that had been stored for longer periods of time. SDS-PAGE analysis of reduced PDI (B) and oxidized PDI (D) SEC samples.

4.3.2 Conditions tested for expression of PDI in M9+ medium

To study the interaction between CT and PDI by SANS, one of the proteins must be deuterated. Since the expression yields of PDI were significantly higher than CT, we decided to develop a protocol for deuteration of PDI. Due to the high cost of deuterated components, conditions for optimal expression of PDI were first tested in M9+ minimal medium without deuterium (Appendix Section B). These experiments were based on a protocol recently developed for production of deuterated proteins in M9+ medium in the PhD thesis of Henrik Vinther Sørensen (UiO), which was adapted from the protocol established by Cai *et al*.[154]. The M9+ medium used deuterated glycerol, which is a cheaper carbon source than deuterated glucose used by Cai *et al*[154].

PDI expression was induced at 3 different IPTG values (0.25, 0.5 and 1 mM) and at 4 different cell densities (OD_{600} = 0.7, 1.2, 1.6 and 5.0) for a total of 12 conditions (Figure 25). Expression levels were evaluated by comparing PDI band intensities of pre- and post- induction samples on a SDS gel. More protein was obtained when higher amounts of IPTG (0.5 mM or 1 mM) were used at lower cell densities (OD_{600} = 0.7 or 1.2), in contrast to Sørensen[155] and Cai *et al*.[154], who found optimal induction points at OD_{600} of 3 and 10, respectively. It is possible that PDI starts to degrade over time and therefore earlier induction may be more efficient for expression. Although it is normal that cells grow very slowly in M9+ medium, our main cultures of pLysS cells only reached an OD_{600} value of 5 after 13 hours of growth, where they appeared to have reached a plateau (Figure 26). For this reason, induction at higher OD_{600} values was not tested. The possible reasons for this slow growth and how it can be improved are discussed in the following section.

Based on the band intensities of PDI in this preliminary expression test, the highest amount of protein was obtained when induction was performed at OD_{600} 1.2 with 1 mM IPTG (Figure 25). Therefore, this condition was chosen to produce deuterated PDI.

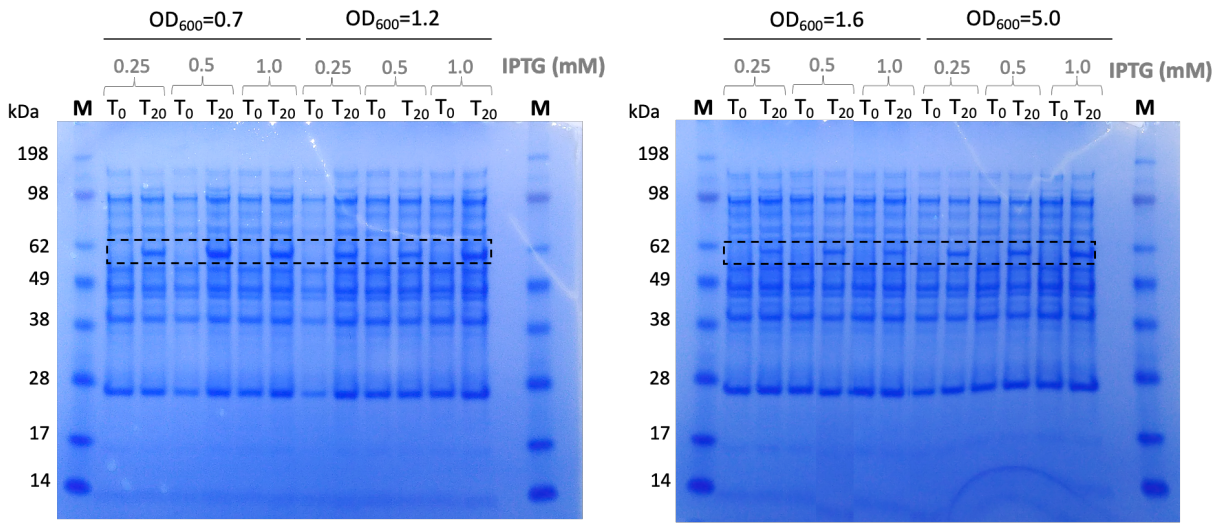


Figure 25. SDS-PAGE analysis of PDI expression in M9+ medium. Cell pellets before (T_0) and after (T_{20}) induction of PDI expression. The concentrations of IPTG tested are presented in grey. The dashed rectangle highlights PDI bands.

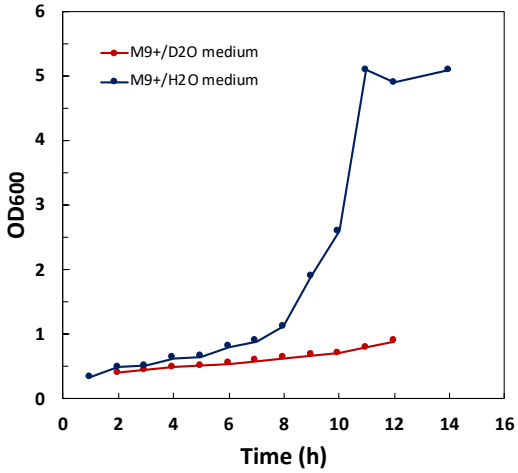


Figure 26. Growth curves of pLysS cells grown in H_2O-based M9+ medium during PDI expression test (blue line) and cells grown in D_2O-based M9+ medium during production of deuterated PDI (red line).

4.2.3 Production of deuterated PDI and preliminary troubleshooting

Deuterated PDI was produced according to the best condition found in the M9+ expression test and purified following the same protocol as for hydrogenated PDI. Growing cells in deuterated medium requires adaptation to M9+ medium and D_2O. This was carried out via a 3-step protocol: first cells were grown in LB/H_2O medium, then transferred to LB/D_2O and finally grown in M9+/D_2O medium. This has the advantage that cells do not have to adapt to the M9+ medium and D_2O simultaneously.

The elution profile and retention volume from SEC of deuterated PDI was similar to PDI expressed in LB medium, and both appeared equally pure (Figure 27.B and 27.C). Nevertheless, only 0.7 mg of protein per L of medium was obtained, which is approximately 20-fold lower compared to the yield of hydrogenated PDI. Such amount of protein would barely be enough for a single SANS experiment, where 0.5 mg -1 mg of protein is usually required for each measurement[103].

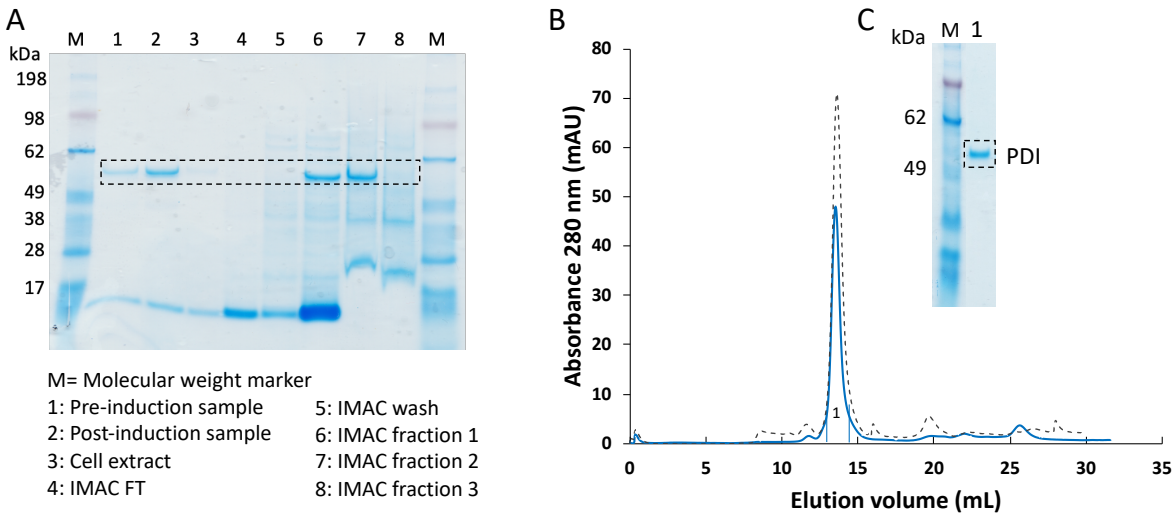


Figure 27. Production of deuterated PDI in M9+ medium. (A) SDS-PAGE analysis of pre- and post-induction samples and (B) Size-exclusion chromatogram of deuterated PDI from IEX step.

One of the main reasons for this low yield can be attributed to the temperature used during overnight protein expression, which was kept at 37 ℃ by mistake and not changed to 25 ℃. Protein production is faster when cells are grown at 37 ℃, however, if it is done overnight, high amounts of protein are rapidly accumulated which could lead to protein aggregation. In addition, the amount of nutrients available in the medium are also used up more rapidly as the cellular processes occur faster, eventually leading to cell death. The yield of deuterated PDI could therefore be improved by using a lower temperature during protein expression.

To obtain sufficient amounts of deuterated protein for SANS experiments in a cost-effective manner, this protocol must be first optimized in H_2O-based M9+ medium, as cell growth is mainly affected by the change from a nutrient-rich LB media to the limited-nutrient M9+ media[154]. Even though the growth in M9+ medium is slow, our cells grew slower than other strains using a similar protocol. The cell densities of overnight pre-cultures in hydrogenated and deuterated M9+ media were 4.0 and 2.7, respectively, whereas cell densities of 10 have been reported for BL21 (DE3) cells[154,155]. Growing the overnight pre-culture for a longer period before inoculating the main cultures may help to accelerate cell growth, as there will be a higher cell density when starting larger cultures. It is also possible that the use of two antibiotics (ampicillin and chloramphenicol) imposed a higher burden to the cells that hindered faster growth. In principle, only ampicillin is necessary, as it is the selection marker for the PDI-encoding plasmid, and chloramphenicol was only included for double selection. Therefore, we carried out a cell growth test where only ampicillin was included in the medium and the overnight pre-culture was grown for 16 hours. We reached an OD_{600} of 6.6 in this culture, which was then used to inoculate a larger culture whose growth was monitored for 16 hours (Figure 28). In contrast to our initial expression test, cells continued growing after 13 hours. An OD_{600} of 9.3 was reached past 16 hours, indicating that higher cell densities can

be obtained by using only ampicillin in the medium and starting from an overnight culture with higher OD_{600} value.

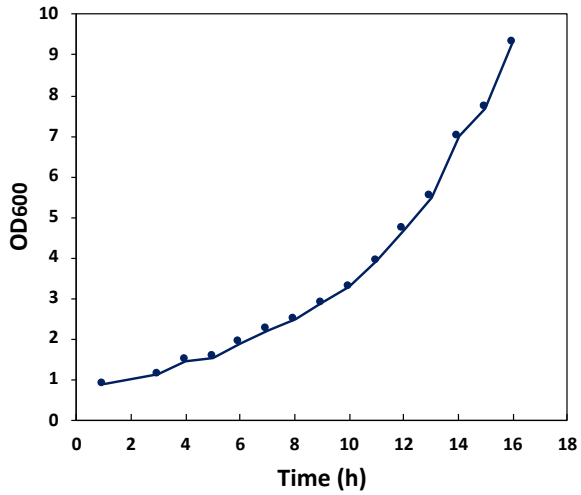


Figure 28. Growth curve of plysS cells in M9+ medium without chloramphenicol. The monitored culture was inoculated from an overnight pre-culture with an OD_{600} of 6.6.

These are valuable observations for future improvement of the PDI deuteration protocol. Besides inducing protein production at a lower temperature (e.g., 20 ℃ or 25℃), it is recommended to test a wider range of OD_{600} values, as we showed that is possible to grow the cells to an OD_{600} close to 10 while still in the exponential phase. By inducing cells at higher densities, cell biomass is increased and therefore the amount of protein could also increase proportionally. However, it must be kept in mind that if the protein is degraded or aggregates, induction at lower OD_{600} is better.

4.4 CT-PDI complex formation studies by SEC

CT disassembly by PDI has been extensively investigated in the literature[48,72,75,76,81], however, detailed structural information of the mechanism of toxin disassembly is still missing. Therefore, in this thesis we aimed at studying such interaction by a combination of SAXS and SANS. To test if CT and PDI could form a stable complex prior to SAS studies, mixtures of both proteins were run on a size-exclusion column, where samples are separated according to their size. If a complex is formed, a new peak eluting at an earlier retention volume is expected due to the larger size of the complex compared to the individual proteins. No complex was observed by SEC under the three conditions evaluated (Figure 29 and Figure 31). Toxin disassembly was expected to occur for the condition where CT was nicked, however, this was not observed either (Figure 29.B).

Conditions were selected based on four observations from studies performed by our collaborators at UCF (presented in detail in the Introduction of this thesis: **1)** reduced, but not oxidized PDI binds to CT and disassembles the toxin; **2)** disassembly is less efficient at low temperatures (< 10 ℃) than at 25 ℃ and 37 ℃; **3)** PDI can only disassemble a “nicked” toxin

(where CTA is cleaved into CTA1 and CTA2) and **4)** unfolding of free CTA1 is prevented at pH 6.5 or at low temperature (< 10 ℃). Therefore, we chose to carry out two SEC experiments at 10 ℃ with proteins in Bis-Tris buffer at pH 6.5, containing 200 mM NaCl and 20 mM TCEP, with either an un-nicked (condition 1) or a nicked CT (condition 2). Since the recombinant toxin is not nicked (at least not completely, see Section 4.1.2), CT was nicked by limited trypsin digestion as described by Jobling *et al*.[129] (Appendix Section H, Figure S2). TCEP was chosen as it is more resistant to air oxidation and is more stable over time when compared to other reducing agents like DTT and GSH[156,157]. Besides reducing PDI disulfide bonds, TCEP also reduces the disulfide bond between Cys187 and Cys199 that connects CTA1 and CTA2. The proteins were mixed on ice and incubation time was kept at minimum (< 3 min) because SPR experiments by Taylor *et al*. showed that PDI-driven CT disassembly occurs within minutes of mixing the proteins, and we wanted to avoid this when the toxin was nicked[48,72].

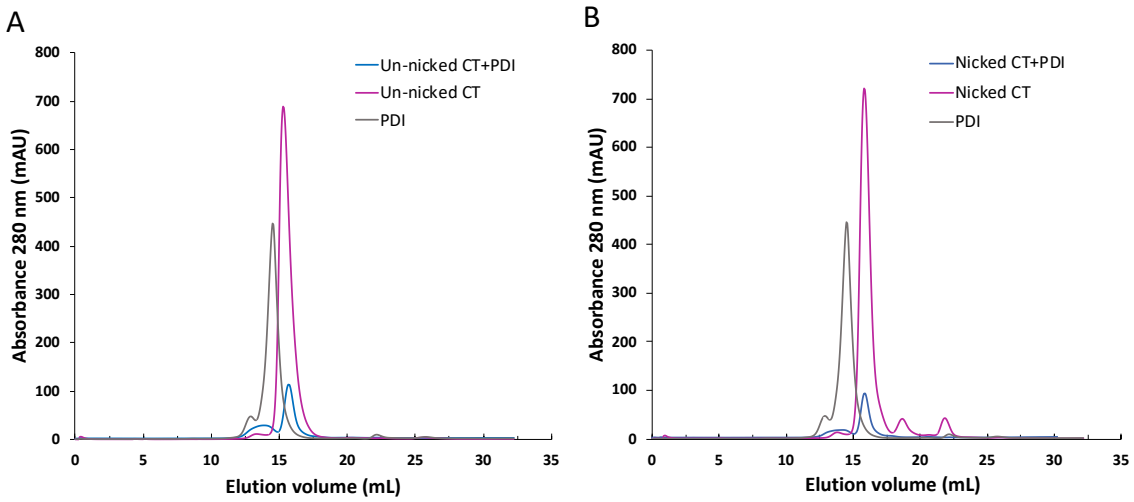


Figure 29. CT-PDI complex formation tests at pH 6.5 by SEC using an un-nicked (A) and nicked (B) CT. No complex is formed as two peaks eluting with similar retention volumes to CT and PDI were observed. Toxin disassembly is not observed either.

There could be several reasons why we did not observe formation of a complex in neither of these conditions. First, the incubation times of the mixtures were too short. In solution, the interaction can be slower since the protein molecules may take longer to find the correct orientation, while the immobilization of the toxin in SPR and the additional flow provided by the perfusion buffer in this technique can lead to a more efficient interaction. Second, the pH and temperature intentionally selected in our experiment to slow down disassembly could have made the interaction even less efficient. Third, assuming that the incubation time was enough, and a weak complex is formed, it is likely that the complex falls apart due to the movement and pressure on the SEC column.

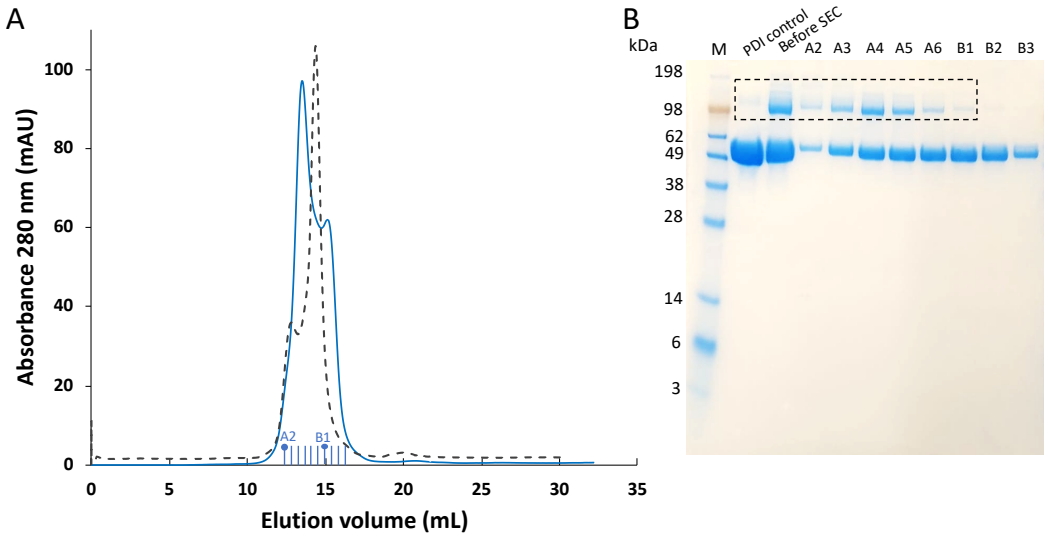


Figure 30. Purification of EDC-treated PDI. (A) Size-exclusion chromatogram and (B) SDS-PAGE analysis of SEC fractions A2-B3. The lanes to the left of the gel were loaded with a molecular weight marker (M), not EDC-treated PDI (labeled "PDI control") and EDC-PDI before SEC. Gel bands inside the dashed rectangle correspond to PDI oligomers. Fractions B1-B3 corresponding to PDI monomers were pooled and used for experiments.

The CT-PDI interaction, PDI was treated with EDC, a "zero-length" cross-linking reagent that generates intra-molecular bonds between the primary amines of basic residues and the carboxyl groups of acidic residues without introducing a spacer arm[158]. Taylor *et al*. previously showed by SPR that EDC-treated PDI can bind to CT, but does not have the ability to displace CTA1 from the rest of the toxin[72]. Therefore, we repeated the SEC experiment using an EDC-treated PDI (Figure 30 and Figure 31) and an un-nicked toxin in PBS pH 7.4. Since TCEP is unstable in neutral phosphate solutions, DTT was used as reducing agent. These conditions did not lead to formation of a complex, possibly due to the reasons already discussed above: too short incubation time, low temperature, or dissociation of the potential complex during SEC. A strong interaction between the proteins is required so that a complex stays together while it moves through the column. The interaction between CT and PDI is dynamic and unless the complex is stabilized by other means like inter cross-linking, it is likely that a complex cannot be isolated by SEC. However, this is not an issue for SAXS, as the proteins can be mixed prior to the measurements, and the conditions of the experiment do not disturb weak interactions.

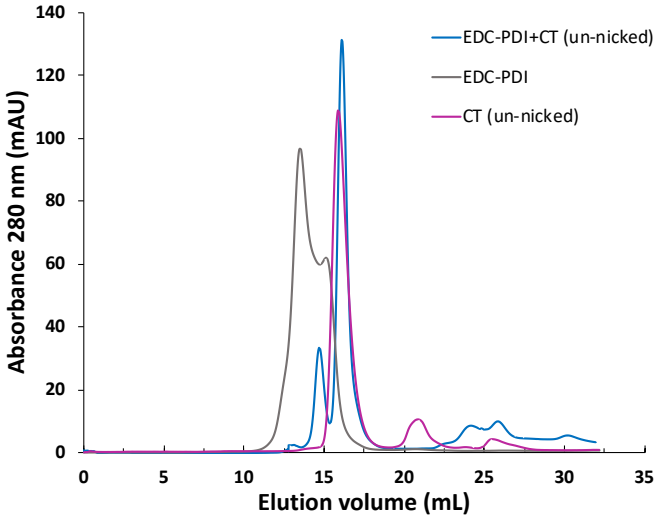


Figure 31. CT/EDC-PDI complex formation test at pH 7.4 using an un-nicked CT. No complex formation is observed.

4.5 SAXS experiments

SAXS experiments were carried out under 6 different conditions to determine if CT and PDI form a complex or at least interact (Appendix Section J). All mixtures contained reduced PDI and a nicked or un-nicked reduced toxin. The scattering of the individual proteins under the exact same conditions was also measured. If CT and PDI interact, the scattering of the mixture should differ significantly from the sum of the scattering of the individual proteins. Particularly, an increase in the scattering intensity at the low-q region of the scattering curve is expected, as the scattering is proportional to the volume squared of the particles, thus larger particles will scatter more than smaller particles (see equation (1), *Introduction Section 1.6.3*). In addition, if the interaction leads to conformational changes in the proteins, changes in the mid-q region of the scattering plot could also be observed, which are related to the shape and internal structure of the particles.

The interaction between pLT and PDI was investigated under only one condition, but due to severe aggregation (Appendix Section K, Figure S10), no structural information was obtained.

In addition to the mixtures, the individual proteins were also characterized by SAXS. This is discussed in the following sections.

4.5.1 Solution structure of oxidized and reduced PDI

The crystal structures of oxidized and reduced PDI show that the protein adopts an open U-shaped conformation when it is oxidized, whereas it has a “closed” U-shape conformation when reduced[67]. Our SAXS measurements also revealed redox-related conformational changes of PDI in solution. From the Guinier analysis, the radius of gyration R_g of oxidized and

reduced PDI in PBS, pH 7.4 was estimated to be 34.38 $\pm$ 0.17 Å and 33.16 $\pm$ 0.16 Å (Table 4) respectively, indicating that oxidized PDI adopts a more extended conformation than reduced PDI, consistent with the crystal structures. Nevertheless, these values are slightly different than those obtained by Okumura *et al*.[68] who reported R_g of 37.2 $\pm$ 0.12 (oxidized PDI) Å and 35.8 $\pm$ 0.2 Å (reduced PDI), whereas Li *et al*[159] reported 33.2 $\pm$ 0.3 Å (oxidized PDI). Okumura's measurements were carried out in a phosphate buffer with DTT, while our measurements were performed in a phosphate buffer but with GSH. Since PDI is a flexible molecule, it is possible that it adopts different conformations in different buffer environments, which could explain why the R_g values are different.

Table 4. SAXS structural parameters of oxidized and reduced PDI. [a] Measurement temperature: 37 ℃; Buffer: PBS pH 7.4, 2 mM GSH. Protein concentration: 1.3 mg/mL

	Oxidized PDI[a]	**Reduced PDI[a]**
Guinier analysis		
R_g (Å)	34.38 +/-0.17	33.16 +/- 0.16
I(0) (cm^{-1})	0.052 +/- 0.00016	0.047 +/-0.00015
qR_g max	1.29	1.29
Range	16-65	9-68
P (r) analysis		
I(0) (cm^{-1})	0.05	0.05
R_g (Å)	36.27	35.29
d_{max}(Å)	144.62	140.05
q range (Å)	16-441	9-459
Total quality estimate from **P**RIMUS	0.77	0.78
Molecular mass estimation		
MM (kDa) from Guinier Analysis	51.2	46.2
MM (kDa) from Qp (Porod invariant)	50.4	34.2

The pair distance distribution function P(r) of PDI, which describes the distribution of distances between any two points in the molecule, has an asymmetrical bell-shape (Figure 32), indicating that PDI has an extended conformation in solution. In addition, the maximum distance in the molecule (d_{max}) is larger for oxidized PDI than reduced PDI (Table 4), confirming that the reduced form is more compact than the oxidized form.

The molecular mass estimation calculated using the forward scattering I(0) of oxidized PDI (from the Guinier analysis, equation (6) in Methods section) was 51 kDa, which is within 10% of its actual value (56 kDa). A similar value was obtained with the Porod invariant method (Table 4), confirming the monomeric state of PDI in solution. However, for reduced PDI, the calculated molecular mass had a higher error (approx. 18% by the Guinier method and 40 % by the Porod invariant). Since the molecular mass calculated by the Guinier method is proportional to the protein concentration, a lower MM would indicate that the concentration was overestimated.

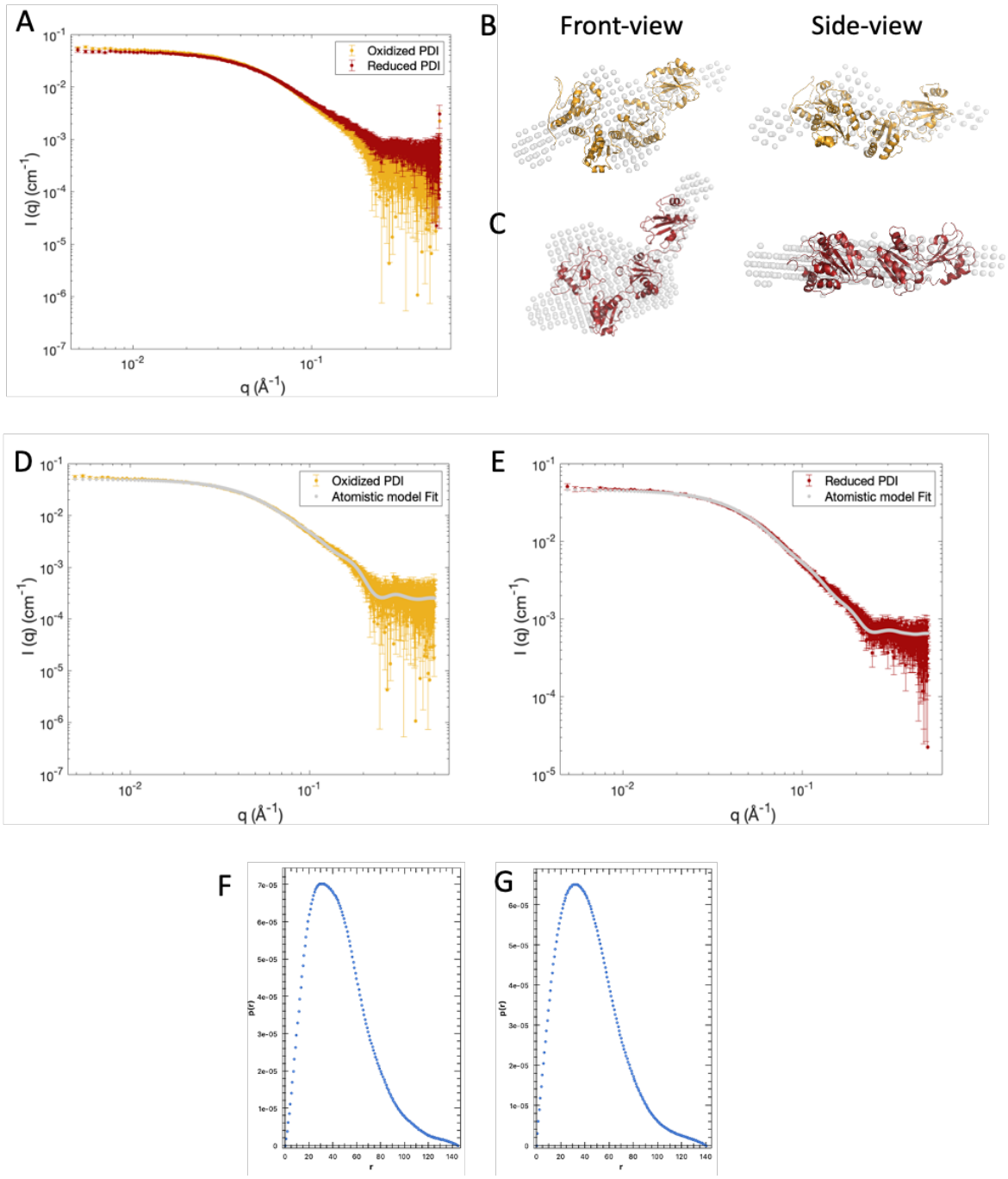


Figure 32. Solution structure of oxidized and reduced PDI (A) SAXS curves of oxidized (red) and reduced (yellow) PDI at 37 °C, pH 7.4. Concentration normalized to 1.3 mg/mL. Atomistic SAXS models of oxidized (B) and reduced (C) PDI (χ^2=1.63 [oxidized]; χ^2=1.21 [reduced]; generated with Pepsi-SAXS[136]), superimposed onto *ab initio* SAXS "envelopes. Comparison of the theoretical SAXS profile of oxidized (D) and reduced (E) PDI (calculated in CRYSOL[132] based on the atomistic model) and the experimental SAXS profiles. Pair distance distribution function of oxidized (F) and reduced (G) PDI.

To test the effect of different reducing agents, SAXS measurements of PDI at pH 6.5, using 20 mM TCEP or at pH 7.4, using 20 mM GSH were also performed (Appendix Section J). TCEP is a more stable reducing agent, while GSH is PDI's natural reducing agent in the endoplasmic reticulum. Since previous CT-PDI interaction experiments have been mainly performed with GSH[48,72,76], this was our preferred reducing agent. Whereas the data quality of the

measurements in TCEP looked good (Appendix Section J, Figure S7), the measurements with 20 mM GSH presented a steep upturn in the low-q region. Typically, aggregation effects are removed by measuring different concentrations of the protein and merging the low-q data from the low concentration samples (where protein aggregation is assumed to be less prominent), with high-q data of high-concentration samples (which have better statistics). In our case, however, the upturn became worse at lower concentrations, suggesting that either radiation damage occurred very early, as we did not see a significant difference between the first frames collected and later frames, or perhaps the presence of other types of aggregates (e.g., from the buffer) whose scattering has stronger impact when the protein concentration is lower. As we did not add GSH to the samples prior to synchrotron SAXS measurements (all samples were sent to the ESRF, and reducing agents were added *in situ*), we did not know if 20 mM GSH could cause protein aggregation. This hypothesis was easily tested by adding the reducing agent to PDI in our home lab, where we saw immediate precipitation. For this reason, the follow-up experiments were carried out with 2 mM GSH only, which provided a molar excess of at least 50-fold, sufficient for PDI reduction and interaction with CT[72].

For the highest-quality SAXS data of PDI (shown in Figure 32), three-dimensional reconstructions of the reduced and oxidized protein were generated (Figure 32 B-C). We also obtained an atomistic model of reduced and oxidized PDI by fitting the crystal structure to the SAXS data using PepsiSAXS[136]. The χ^2 value of both models is close to 1 (Table 5), suggesting that they fit well to the experimental data. These models suggested that PDI has different conformations when oxidized and reduced (Figure 32 B-C). Interestingly, the side view of the oxidized model showed that PDI has a flask-disk shape, whereas the reduced form does not. This is consistent with the crystal structure, where all domains of oxidized PDI are in the same plane, while one of the domains in reduced PDI is twisted out of the plane[64].

Table 5. PDI *ab initio* model-fitting and atomistic modelling results.

	Oxidized PDI	**Reduced PDI**
DAMAVER/DAMMIN (*ab initio* model)		
χ^2 of average model	1.064	1.165
PepsiSAXS (atomistic model)		
Estimated R_g from crystal structure	33.06	29.04
Estimated resoltuion (Å)	12.57	12.57
Experimental R_g from curve	30.29	28.71
Final χ^2	1.21	1.63

4.5.2 Solution structure of CT

To analyze if the solution structure of a completely activated CT (nicked and reduced) has significant changes compared to an inactivated toxin (un-nicked and not reduced), SAXS measurements were performed. As expected, the nicked and un-nicked toxin have very similar structures in solution as seen by the practically complete overlap of their SAXS

scattering profiles (Figure 33.A). The calculated R_g values were around 28 Å for both toxins (Table 6) and the calculated molecular weight was was within 5-10% of the actual molecular mass of CT (approx. 85 kDa). Nevertheless, for the measurements performed at 37 ℃ using 2 mM GSH, the molecular weight estimation disagrees by 20-40 % from the actual value (Table 6), indicating that the concentration was underestimated. The error is also associated with the steep upturn in the low-q region of these samples (Appendix Section K, Figure S9), similar to some of the PDI samples. Interestingly, the upturn became worse at lower concentrations. This was observed both when the measurements were carried out using 20 mM GSH (not shown) and later when the GSH was reduced to 2 mM, so corrections of the low-q data by measuring a dilution series of the protein were not useful in our case. GSH did not seem to be the problem in this case, but it is possible that the temperature influenced the stability and the monodispersity of CT. This should be further investigated by other methods like Dynamic light scattering (DLS), which can give information on protein polydispersity.

Table 6. SAXS structural parameters of CT. [a] Measurement temperature: 10 ℃; buffer: 20 mM Bis-Tris, pH 6.5, 200 mM NaCl. [b] Measurement temperature: 37 ℃; buffer: PBS pH 7.4, 2 mM GSH. [b]CT concentration= 2 mg/mL

	CT (un-nicked)[a]	**CT (nicked)[a]**	**CT (un-nicked)[b]**	**CT (un-nicked)[b]**
Reducing agent	None	20 mM TCEP	2 mM GSH	None
Guinier analysis				
R_g (Å)	28.19 +/- 0.06	27.87 +/- 0.09	28.3 +/-0.1	28.98
I(0) (cm^{-1})	0.12 +/- 0.00015	0.097 +/- 0.00019	0.096 +/- 0.00021	0.0099 +/- 0.00023
qR_g max	1.29	1.30	1.25	1.3
Range	14-81	22-83	22-78	33-79
P (r) analysis				
I(0) (cm^{-1})	0.12	0.10	0.1	0.1
R_g (Å)	28.24	27.83	29.21	29.38
d_{max}(Å)	103.65	90.75	103.17	106.35
q range (Å)	14-370	22-550	22-539	33-528
Total quality estimate from **P**RIMUS[131]	0.86	0.91	0.83	0.81
Molecular mass estimation				
MM (kDa) from Guinier Analysis	76.73	62.02	63.94	62.02
MM (kDa) from Porod invariant	81.04	74.63	49.67	61.67

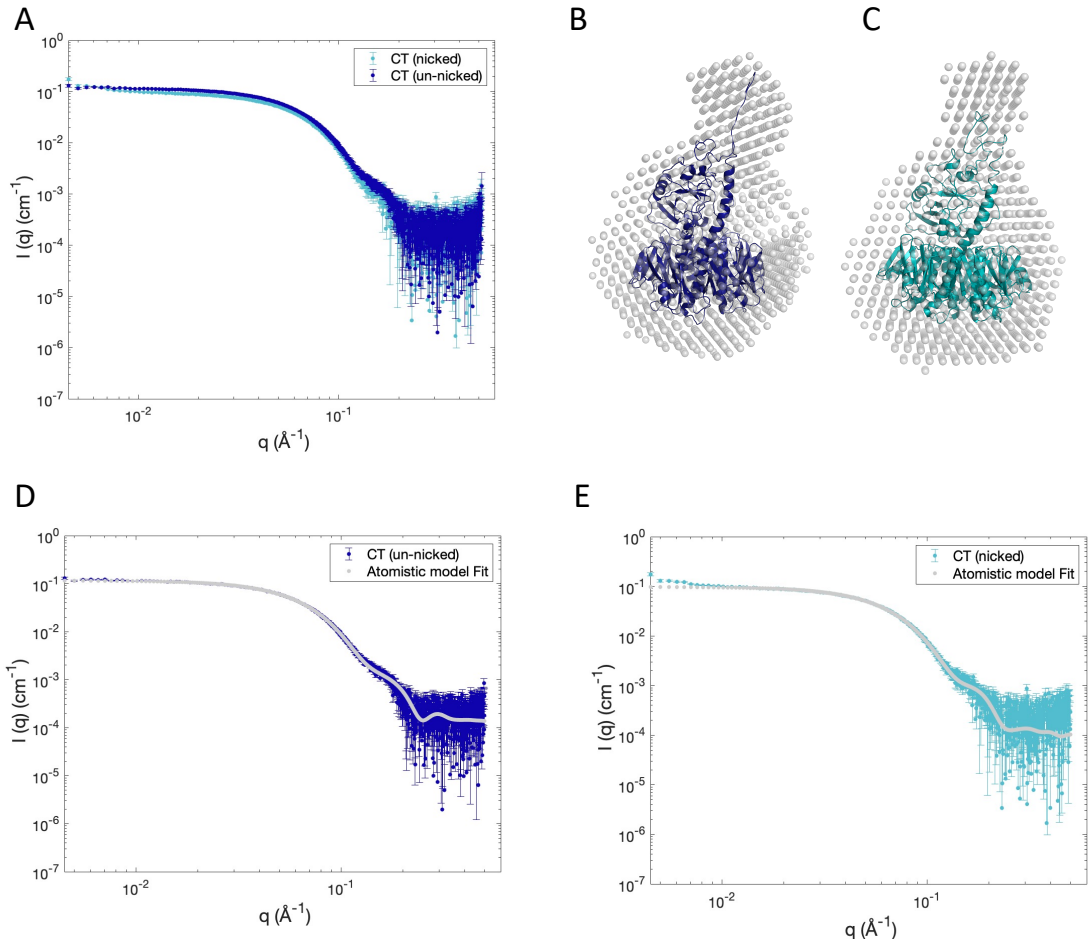


Figure 33. Solution structure of CT (A) SAXS curves of un-nicked (dark blue) and nicked (light blue) cholera toxin at 10 °C, pH 6.5. Concentration normalized to 2 mg/mL. Atomistic SAXS models of un-nicked (B) and nicked (C) CT (χ^2=1.7 [un-nicked]; χ^2=1.38 [nicked]; generated with Pepsi-SAXS[136]) superimposed to *ab initio* SAXS "envelopes" (average models of 20 calculated models in DAMMIF[135]). (D) Comparison of the theoretical SAXS profile of un-nicked and nicked CT (calculated based on the atomistic model) and the experimental SAXS profiles.

For the highest quality data (Figure 33), an *ab initio* model of the nicked and un-nicked toxin as well as an atomistic model based on the crystal structure was calculated. The theoretical R_g values from the crystal structure are slightly lower than the experimental values (Table 6-7), indicating that CT is less compact in solution. These models suggest that the loop connecting CTA1 and CTA2 in the un-nicked toxin has remarkable flexibility and appears to be extended. For the nicked toxin, this loop is cleaved and we would expect more flexibility, although the model does not seem to suggest this.

Table 7. CT *ab initio* model fitting and atomistic modelling results

	CT (un-nicked)	CT (nicked)
DAMAVER/DAMMIN		
χ^2	1.037	1.029
PepsiSAXS		
Rg from crystal structure	26.7	26.7
Estimated resoltuion (Å)	12.57	12.57
Experimental Rg from curve	30.30	46.0
Final χ^2	1.17	1.38

4.5.3 CT-PDI complex formation studies

As the main aim of this thesis was to investigate the interaction between CT and PDI by SAXS (and SANS), several measurements at different temperatures, pH and types of reducing agents were performed to evaluate if an interaction or a complex could be obtained. Conditions were selected according to previous experiments where CT and PDI have been shown to interact (see Introduction and Section 4.4) as well as observations from our preliminary SEC experiments. All samples were sent frozen to ESRF, where they were reduced *in situ* with either TCEP or GSH and measured.

To follow-up on our SEC experiments, our initial SAXS measurements were performed at 10 ℃ in a Bis-Tris buffer, pH 6.5, containing TCEP. To test our hypothesis that a weak complex could have formed, but fallen apart during SEC, the samples were treated in the same way as for the SEC experiment and mixed just before the SAXS measurement. If a complex was formed, an increase in the scattering intensity at the low-q region of the scattering curve is expected compared to the scattering of the sum of the individual proteins (linear combination). Since a nicked CT was used for this experiment, it was also expected that a small proportion could be disassembled by PDI. Nevertheless, no interaction seemed to occur, as the linear combination fit well to the experimental data (Figure 36.B). In addition, the R_g value of the mixture was 32.05 $\pm$ 0.07 Å (Table 8), which is too low to correspond to the CT-PDI complex. In fact, this value is consistent with the average R_g of PDI and CT, suggesting that the scattering of the mixture is the contribution of the individual proteins and not the complex.

To get insight into how the scattering of a possible CT-PDI complex would look and compare it to our data, docking models were calculated using the web servers ClusPro[138] and HDOCK[139]. The top 20 docking models with the highest scores provided by these servers were manually analyzed to see how well they could represent a CT-PDI complex.

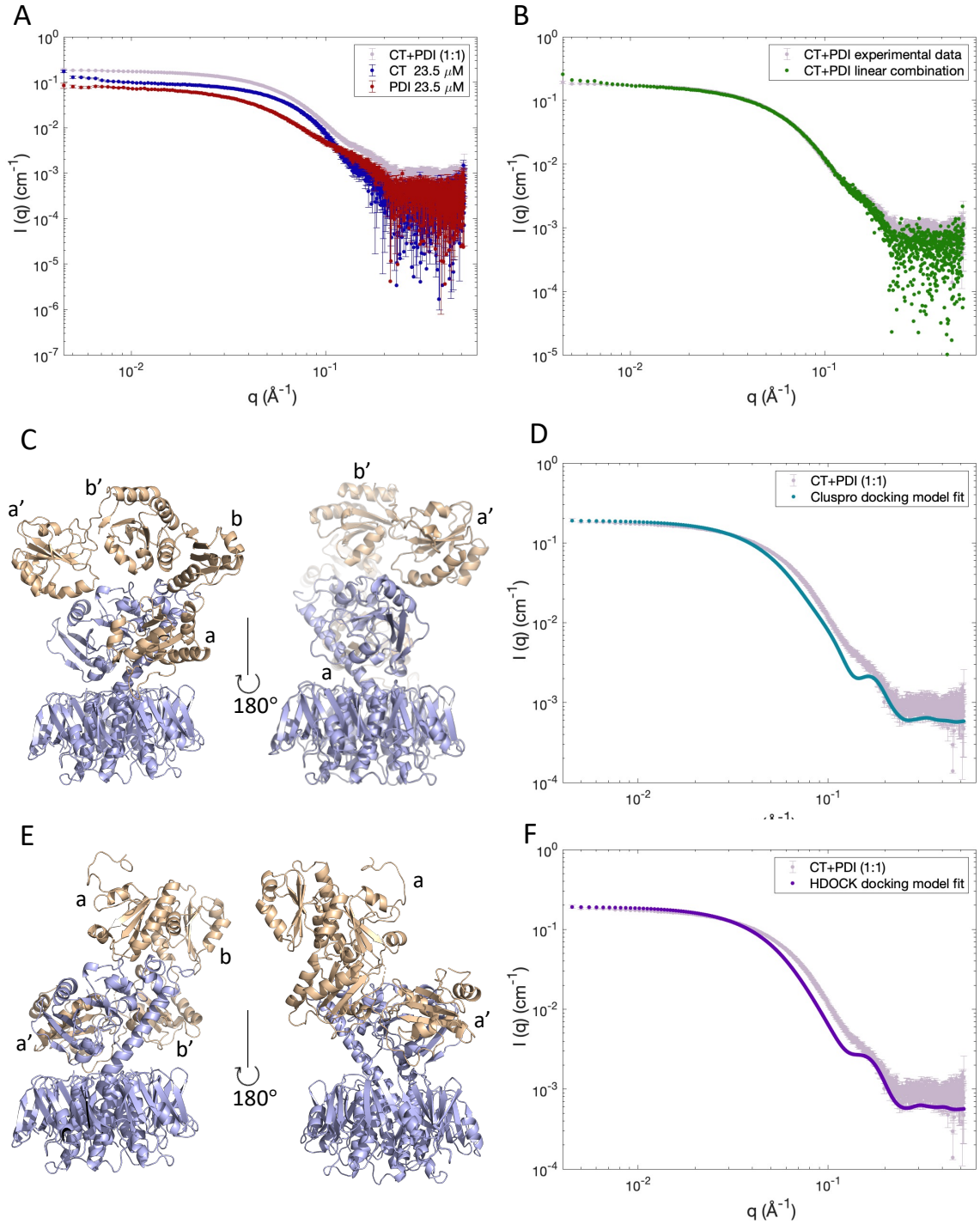


Figure 34. SAXS analysis of a CT and PDI mixture at 10 ℃ in a bis-tris buffer, pH 6.5 containing 20 mM TCEP. (A) SAXS curves of CT (blue), PDI (red) and a 1:1 molar ratio mixture of the proteins (light purple). (B) Linear combination fit (green) of CT-PDI mixture. (C) Docking model of CT-PDI generated with ClusPro[138]. (D) SAXS curve of the CT-PDI mixture fitted using Cluspro docking model (χ^2=181.58; fit generated with CRYSOL[132]). (E) Docking model of CT-PDI generated with HDOCK[139] and (F) CRYSOL-fitted CT-PDI mixture with HDOCK model (χ^2=241.55).

Models where PDI was docked at the bottom of the pentamer were discarded, as this orientation places PDI too far away from CTA1, making it difficult to promote CTA1 displacement. In addition, the SPR and ELISA experiments immobilized the toxin by its B-

pentamer on a surface. As disassembly occurs under these conditions, it means that it would not be possible for PDI to dock to CT from the bottom of the B-pentamer. In this way, only 5 models were selected, and their theoretical scattering profiles were calculated with CRYSOL. The top 2 models with the best fit to the experimental data are shown in Figure 34.C and 34.E. Nevertheless, both fits differ significantly from the experimental curves with χ^2 values remarkably higher (>180), indicating that this model of the CT-PDI complex does not represent the experimental mixture. It is worth noting that HDOCK can also include SAXS data in the docking analysis, however, the program showed a warning message stating that the molecular weight and R_g values of the experimental SAXS data were too small to correspond to a complex of the proteins (Appendix Section L).

Even though CT and PDI can interact at low temperatures[48], acidic pH and when TCEP is used as reducing agent (Teter Group at UCF, personal communication), we did not observe an interaction by SEC and SAXS under these conditions. As most experiments demonstrating CT-PDI interaction have been carried out at 37 ℃, pH 7.4 and with GSH to mimic conditions in the ER, we chose this temperature and reducing agent as the following conditions to be evaluated by SAXS. All mixtures were incubated for at least 30 min prior to the measurements. To ensure that PDI was reduced, the initial set of samples measured was incubated with 20 mM GSH, however, most of our data presented steep upturns in the low-q region, making these samples impossible to interpret (Appendix Section K, Figure S10). The possible reason, as already discussed in Section 4.5.1, was aggregation due to high concentrations of GSH not amenable with the proteins.

Based on this observation, we decided to reduce the concentration of GSH to 2 mM, which did not result in visible precipitation of the proteins. The best data are shown in Figure 35, where slight upturns of the scattering of CT and CT/PDI mixture were still visible. The data were analyzed in the same way as the previous samples and showed that no interaction occured (Figure 35.B-D). The best fit was obtained with the linear combination of the proteins rather than fits from the best ClusPro and HDOCK models (Figure 35.C-E), indicating that no interaction was taking place. Finally, we decided to test if treating PDI with quercitin-3-rutenoside (Q3R), a flavonol that binds the b'x region of PDI[82] could result in a stable complex. Q3R-treated PDI can still bind to CT but lacks its ability to displace CTA1 from CT[81]. Although the molecular mechanism of PDI inhibition by Q3R remains unknown, previous SAXS data has shown that Q3R-treated PDI exhibits a more compact conformation when compared to native PDI[82]. Based on this observation, Guyette *et al*. suggested that the conformational change of PDI upon Q3R addition prevents PDI from disassembling CT. Our SAXS experiments, however, did not show formation of a complex between Q3R-treated PDI and CT (Figure 36.A-C). It is recommended that these Q3R-experiments are repeated using different molar ratios of Q3R and perhaps incubating the proteins for longer.

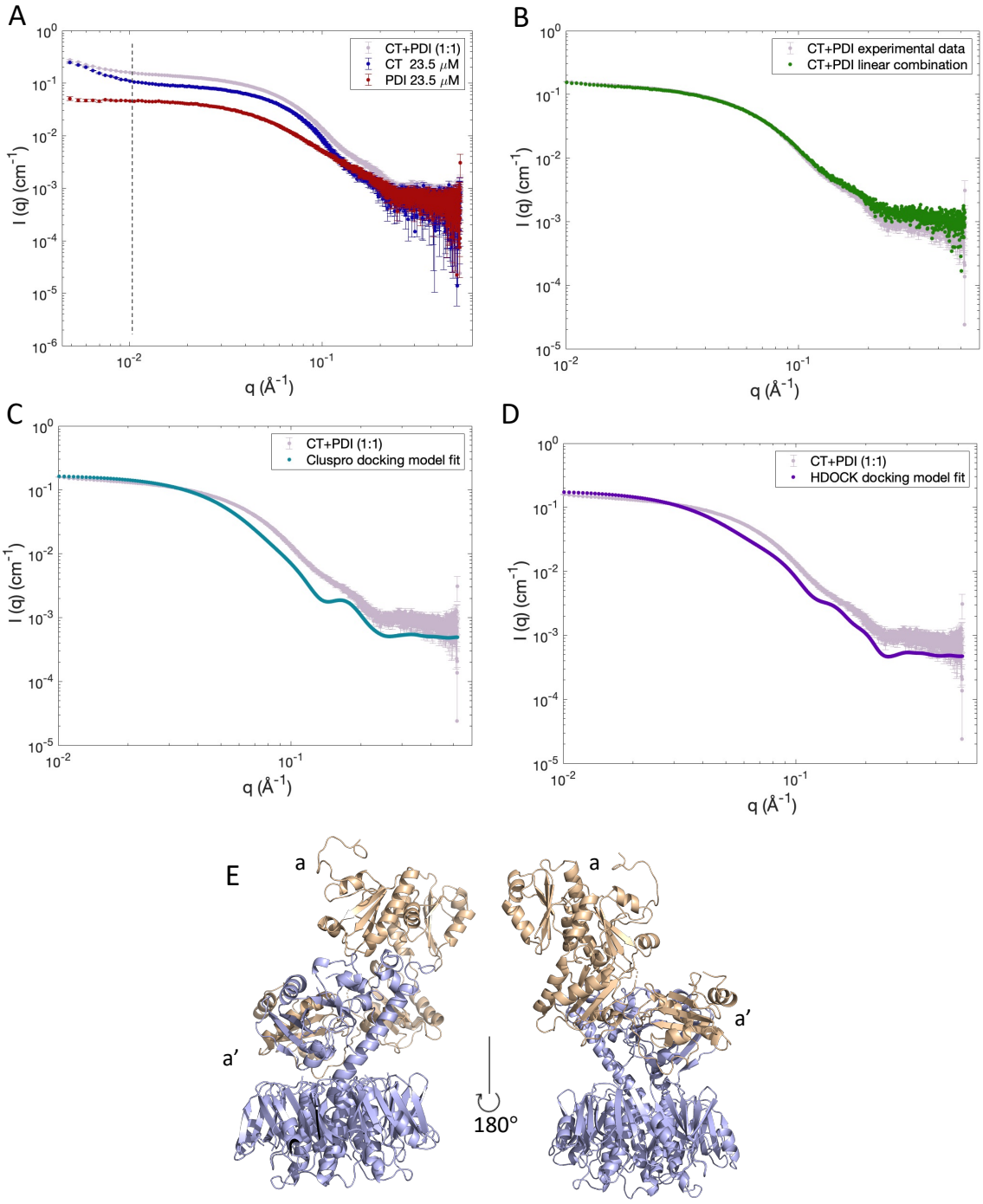


Figure 35. SAXS analysis of a CT-PDI mixture at 37 ℃ in PBS, pH 7.4 containing 2 mM GSH. (A) SAXS curves of CT (blue), PDI (red) and a 1:1 molar ratio mixture of the proteins (light purple). The dashed line indicates where the data were truncated to be analyzed by different fits. (B) Linear combination fit (green) of CT-PDI mixture. (C) SAXS curve of the mixture fitted using Cluspro docking model (shown in Figure xx) (χ^2=443.87; fit generated with CRYSOL[132]). (D) CRYSOL-fitted CT-PDI mixture with HDOCK model (χ^2=435.28). (E) Docking model of CT-PDI generated with HDOCK[139].

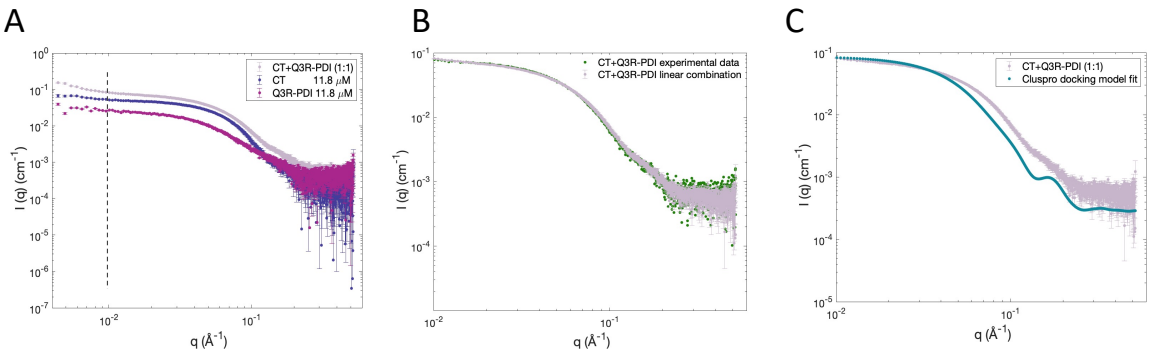


Figure 36. SAXS analysis of a mixture containing Q3R-treated PDI and CT. (A) SAXS curves of CT (purple), Q3R-PDI (magenta) and Q3R-PDI/CT mixture (light purple). The dashed line indicates where the data was truncated for further analysis with theoretical model fits. (B) Linear combination fit (green) of Q3R-PDI/CT mixture. (C) SAS curve of the mixture fitted with the Cluspro docking model (Figure 34) (χ^2=195.11).

Although the interaction between CT and PDI has been demonstrated by various techniques[48,72,76], none of the conditions that we tested resulted in a complex or an interaction. There are several possible explanations. As briefly mentioned earlier in this thesis, most experiments showing efficient disassembly involved immobilized toxin, while all our experiments were performed in solution. It is possible that the interaction requires immobilization of the toxin to a surface, which will not only position CT in the "correct" orientation but could provide additional force to facilitate displacement of the CTA1 subunit by PDI, similar to when removing the cork of a wine bottle: placing the bottle on a flat surbace makes it easier to remove the cork than when having the bottle in the air.

It would also be interesting to monitor CT-PDI interaction with time-resolved SAXS and determine if an interaction takes place and how long does it take to see any changes in the scattering. Furthermore, it is recommended to replicate the CT-disassembly ELISA experiment performed by our collaborators at UCF[76], to confirm that the interaction can be replicated with the recombinant proteins produced in our lab.

5. Conclusion and future perspectives

AB_5 toxins like CT and pLT must be disassembled in order to become toxic to cells. Displacement of the A1 subunit of these toxins is physically mediated by PDI, which unfolds upon binding to the toxin and provides the force to disrupt the non-covalent interactions holding the reduced holotoxin together[72]. Nevertheless, due to the flexibility of these proteins and dynamic nature of their interaction, structural mechanistic details are still not well understood. The aim of this thesis was to investigate the interaction of these toxins with PDI by a small-angle X-ray scattering techniques, which are e suitable for the study of dynamic systems.

The first part of this project focused on producing mg amounts of pure CT, pLT and PDI for SAXS and SANS experiments. We developed a new protocol to produce the toxins using *Vibrio natriegens* as expression host, which improved the yields by at least 10-fold compared to the previous protocol for production in *E. coli*. PDI (hydrogenated) was successfully produced with a slightly modified protocol that resulted in higher protein purity. While pilot expression of deuterated PDI yielded very little protein for SANS, our work provided a solid basis for further development of deuteration trials. It is recommended to use a temperature of 25 or 20 ℃ for overnight induction instead of 37 ℃, as the protein could aggregate if produced too quickly.

To evaluate the quality of the cholera toxin produced in Vmax, a customized crystallization screen was set-up and crystals were obtained. The aspect of the crystals suggested that , this would be a good indication that our new Vmax protocol can yield homogenous protein. CT was also characterized by SAXS. The protein was demonstrated to be folded and have a compact structure in solution. PDI, in contrast, was shown to have a more extended conformation in solution when compared to the crystal structure, consistent with previous PDI SAXS experiments[68,159].

Under the conditions tested in this thesis, no interaction between CT (nicked or un-nicked) and PDI (native, cross-linked or treated with Q3R) was observed by SEC or SAXS. Since our choice of conditions resembled those used in previous experiments, where the interaction has been demonstrated, we hypothesize that the interaction in solution is inefficient and requires either long incubation times (> 30 minutes) or immobilization of the toxin to a surface. This may position the toxin in a suitable orientation for interaction with PDI or increase the force provided by PDI to facilitate the displacement of the A1 subunit. Time-resolved SAXS could be an option to monitor in real time the interaction between CT and PDI and determine possible protein conformational changes that occur. It is also recommended to replicate the CT-disassembly ELISA experiment performed by our collaborators at UCF[76] to rule out that the lack of interaction is associated with the recombinant proteins produced in different expression hosts.

This work made important progress in terms of production of CT, pLT and PDI for SAXS/SANS studies and provides the basis for future interaction experiments.

6. References

[1] Holmes, K. K.; Bertozzi, S.; Bloom, B. R.; Jha, P. Major Infectious Diseases: Key Messages from Disease Control Priorities. In *Major Infectious Diseases*; International Bank for Reconstruction and Development / The World Bank: Washington (DC), 2017; pp 1–27. https://doi.org/10.1596/978-1-4648-0524-0.

[2] World Health Organization. *World Health Statistics 2020. Monitoring Health for the SDGs, Sustainable Development Goals*; Geneva, 2020.

[3] World Health Organization. Impact of COVID-19 on people's livelihoods, their health and our food systems https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people%27s-livelihoods-their-health-and-our-food-systems (accessed Mar 4, 2021).

[4] World Health Organization. The top 10 causes of death https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (accessed Jan 24, 2021).

[5] World Health Organization. Diarrhoeal disease https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease (accessed Jan 24, 2021).

[6] Leitner, D. R.; Lichtenegger, S.; Temel, P.; Zingl, F. G.; Ratzberger, D.; Roier, S.; Schild-Prüfert, K.; Feichter, S.; Reidl, J.; Schild, S. A Combined Vaccine Approach against Vibrio Cholerae and ETEC Based on Outer Membrane Vesicles. *Front. Microbiol.* **2015**, *6* (JUL), 1–17. https://doi.org/10.3389/fmicb.2015.00823.

[7] Zampini, M.; Pruzzo, C.; Bondre, V. P.; Tarsi, R.; Cosmo, M.; Bacciaglia, A.; Chhabra, A.; Srivastava, R.; Srivastava, B. S. Vibrio Cholerae Persistence in Aquatic Environments and Colonization of Intestinal Cells: Involvement of a Common Adhesion Mechanism. *FEMS Microbiol. Lett.* **2005**, *244* (2), 267–273. https://doi.org/10.1016/j.femsle.2005.01.052.

[8] Sack, D. A.; Sack, R. B.; Nair, G. B.; Siddique, A. Cholera. *Lancet* **2004**, *363* (9404), 223–233.

[9] World Health Organization. Cholera https://www.who.int/news-room/fact-sheets/detail/cholera (accessed Jan 26, 2021).

[10] The Lancet Gastroenterology & Hepatology. Health Catastrophe: The Toll of Cholera in Yemen. *Lancet Gastroenterol. Hepatol.* **2017**, *2* (9), 619. https://doi.org/10.1016/S2468-1253(17)30224-8.

[11] Shaikh, H.; Lynch, J.; Kim, J.; Excler, J. L. Current and Future Cholera Vaccines. *Vaccine* **2020**, *38*, A118–A126. https://doi.org/10.1016/j.vaccine.2019.12.011.

[12] De, R.; Ghosh, J. B.; Sen Gupta, S.; Takeda, Y.; Nair, G. B. The Role of Vibrio Cholerae Genotyping in Africa. *J. Infect. Dis.* **2013**, *208* (SUPPL. 1). https://doi.org/10.1093/infdis/jit199.

[13] Albert, M. J.; Ansaruzzaman, M.; Bardhan, P. K.; Faruque, A. S. G.; Faruque, S. M.; Islam, M. S.; Mahalanabis, D.; Sack, R. B.; Salam, M. A.; Siddique, A. K.; Yunus, M. D.; Zaman, K. Large Epidemic of Cholera-like Disease in Bangladesh Caused by Vibrio Cholerae 0139 Synonym Bengal. *Lancet* **1993**, *342* (8868), 387–390. https://doi.org/10.1016/0140-6736(93)92811-7.

[14] Sánchez, J.; Holmgren, J. Cholera Toxin Structure, Gene Regulation and Pathophysiological and Immunological Aspects. *Cell. Mol. Life Sci.* **2008**, *65* (9), 1347–1360. https://doi.org/10.1007/s00018-008-7496-5.

[15] Safa, A.; Sultana, J.; Cam, P. D.; Mwansa, J. C.; Kong, R. Y. C. Vibrio Cholerae O1 Hybrid El Tor Strains, Asia and Africa. *Emerg. Infect. Dis.* **2008**, *14* (6), 987–988. https://doi.org/10.3201/eid1406.080129.

[16] Binh, M. N.; Je, H. L.; Ngo, T. C.; Seon, Y. C.; Nguyen, T. H.; Dang, D. A.; Hye, R. L.; Ansaruzzaman, M.; Endtz, H. P.; Chun, J.; Lopez, A. L.; Czerkinsky, C.; Clemens, J. D.; Dong, W. K. Cholera Outbreaks Caused by an Altered Vibrio Cholerae O1 El Tor Biotype Strain Producing Classical Cholera Toxin B in Vietnam in 2007 to 2008. *J. Clin. Microbiol.* **2009**, *47* (5), 1568–1571. https://doi.org/10.1128/JCM.02040-08.

[17] Chowdhury, F.; Mather, A. E.; Begum, Y. A.; Asaduzzaman, M.; Baby, N.; Sharmin, S.; Biswas, R.; Ikhtear Uddin, M.; LaRocque, R. C.; Harris, J. B.; Calderwood, S. B.; Ryan, E. T.; Clemens, J. D.; Thomson, N. R.; Qadri, F. Vibrio Cholerae Serogroup O139: Isolation from Cholera Patients and Asymptomatic Household Family Members in Bangladesh between 2013 and 2014. *PLoS Negl. Trop. Dis.* **2015**, *9* (11), e0004183.

[18] Centers for Disease Control and Prevention, C. Enterotoxigenic E.coli https://www.cdc.gov/ecoli/etec.html (accessed Jan 26, 2021).

[19] Qadri, F.; Svennerholm, A.-M.; Faruque, A. S. G.; Bradley, R. . Enterotoxigenic Escherichia Coli in Developing Countries: Epidemiology, Microbiology, Clinical Features, Treatment, and Prevention. *Clin. Microbiol. Rev.* **2005**, *18* (3), 465–483.

[20] Duan, Q.; Xia, P.; Nandre, R.; Zhang, W.; Zhu, G. Review of Newly Identified Functions Associated With the Heat-Labile Toxin of Enterotoxigenic Escherichia Coli. *Front. Cell. Infect. Microbiol.* **2019**, *9* (August), 1–11. https://doi.org/10.3389/fcimb.2019.00292.

[21] Butt, S.; Saleh, M.; Gagnon, J. Impact of the Escherichia Coli Heat-Stable Enterotoxin b (STb) on Gut Health and Function. *Toxins (Basel).* **2020**, *12* (12), 1–12. https://doi.org/10.3390/toxins12120760.

[22] Summers, D. *The Biology of Plasmids*; Blackwell Science: Oxford, 1996; Vol. 71. https://doi.org/10.1016/s0168-9525(97)84646-x.

[23] Wang, H.; Zhong, Z.; Luo, Y.; Cox, E.; Devriendt, B. Heat-Stable Enterotoxins of Enterotoxigenic Escherichia Coli and Their Impact on Host Immunity. *Toxins* . 2019. https://doi.org/10.3390/toxins11010024.

[24] Taxt, A.; Aasland, R.; Sommerfelt, H.; Nataro, J.; Puntervoll, P. Heat-Stable Enterotoxin of Enterotoxigenic Escherichia Coli as a Vaccine Target. *Infect. Immun.* **2010**, *78* (5), 1824–1831. https://doi.org/10.1128/IAI.01397-09.

[25] Seo, H.; Zhang, W. Development of Effective Vaccines for Enterotoxigenic Escherichia Coli. *Lancet Infect. Dis.* **2020**, *20* (2), 150–152. https://doi.org/10.1016/S1473-3099(19)30631-0.

[26] Byrd, W.; Boedeker, E. C. Attenuated Escherichia Coli Strains Expressing the Colonization Factor Antigen I (CFA/I) and a Detoxified Heat-Labile Enterotoxin (LThK63) Enhance Clearance of ETEC from the Lungs of Mice and Protect Mice from Intestinal ETEC Colonization and LT-Induced F. *Vet. Immunol. Immunopathol.* **2013**, *152* (1), 57–67. https://doi.org/https://doi.org/10.1016/j.vetimm.2012.10.001.

[27] Clemens, J. D.; Sack, D. A.; Harris, J. R.; Chakraborty, J.; Neogy, P. K.; Stanton, B.; Huda, N.; Khan, M. U.; Kay, B. A.; Khan, M. R.; Ansaruzzaman, M.; Yunus, M.; Rao, M. R.; Svennerholm, A.-M.; Holmgren, J. Cross-Protection by B Subunit-Whole Cell Cholera Vaccine against Diarrhea Associated with Heat-Labile Toxin-Producing Enterotoxigenic Escherichia Coli: Results of a Large-Scale Field Trial. *J. Infect. Dis.* **1988**, *158* (2), 372–377.

[28] Heggelund, J. E.; Bjørnestad, V. A.; Krengel, U. Vibrio Cholerae and Escherichia Coli Heat-Labile Enterotoxins and Beyond. In *The Comprehensive Sourcebook of Bacterial Protein Toxins*; Elsevier Ltd, 2015; pp 195–229.

[29] Wilson, B. A.; Winkler, M. E.; Ho, B. T. *Bacterial Pathogenesis:A Molecular Approach*;

ASM press: Washington, dc, 2019.
[30] Casadevall, A.; Pirofski, L. A. Virulence Factors and Their Mechanisms of Action: The View from a Damage-Response Framework. *J. Water Health* **2009**, *7* (SUPPL. 1), 2–18. https://doi.org/10.2166/wh.2009.036.
[31] do Vale, A.; Cabanes, D.; Sousa, S. Bacterial Toxins as Pathogen Weapons Against Phagocytes. *Front. Microbiol.* **2016**, *7* (42), 1–21.
[32] Duport, C.; Alpha-Bazin, B.; Armengaud, J. Advanced Proteomics as a Powerful Tool for Studying Toxins of Human Bacterial Pathogens. *Toxins (Basel).* **2019**, *11* (10). https://doi.org/10.3390/toxins11100576.
[33] Roy, V.; Ghani, K.; Caruso, M. A Dominant-Negative Approach That Prevents Diphthamide Formation Confers Resistance to Pseudomonas Exotoxin A and Diphtheria Toxin. *PLoS One* **2010**, *5* (12). https://doi.org/10.1371/journal.pone.0015753.
[34] Carbonetti, N. H. Pertussis Toxin and Adenylate Cyclase Toxin: Key Virulence Factors Of. *Rev. Lit. Arts Am.* **2010**, 455–469.
[35] Michalska, M.; Wolf, P. Pseudomonas Exotoxin A: Optimized by Evolution for Effective Killing. *Front. Microbiol.* **2015**, *6* (SEP), 1–7. https://doi.org/10.3389/fmicb.2015.00963.
[36] Yamamoto, T.; Yokota, T.; Suyama, A.; Mori, N.; Wada, A. Gene Expression in the Polycistronic Operons of Escherichia Coli Heat-Labile Toxin and Cholera Toxin: A New Model of Translational Control. *FEBS J.* **1985**, *181* (2), 377–380.
[37] Reichow, S. L.; Korotkov, K. V.; Hol, W. G. J.; Gonen, T. Structure of the Cholera Toxin Secretion Channel in Its Closed State. *Nat. Struct. Mol. Biol.* **2010**, *17* (10), 1226–1232. https://doi.org/10.1038/nsmb.1910.
[38] Sixma, T. K.; Pronk, S. E.; Kalk, K. H.; Wartna, E. S.; Van Zanten, B. A. M.; Witholt, B.; Hoi, W. G. J. Crystal Structure of a Cholera Toxin-Related Heat-Labile Enterotoxin from E. Coli. *Nature* **1991**, *351* (6325), 371–377. https://doi.org/10.1038/351371a0.
[39] O'Neal, C. J.; Amaya, E. I.; Jobling, M. G.; Holmes, R. K.; Hol, W. G. J. Crystal Structures of an Intrinsically Active Cholera Toxin Mutant Yield Insight into the Toxin Activation Mechanism. *Biochemistry* **2004**, *43* (13), 3772–3782.
[40] Chaudhuri, K.; Chatterjee, S. N. Cholera Toxin (CT): Structure. In *Cholera Toxins*; Springer-Verlag Berlin Heidelberg, 2009; pp 105–123. https://doi.org/10.1007/978-3-540-88452-1_7.
[41] O'Neal, C. J.; Amaya, E. I.; Jobling, M. G.; Holmes, R. K.; Hol, W. G. J. Crystal Structures of an Intrinsically Active Cholera Toxin Mutant Yield Insight into the Toxin Activation Mechanism. *Biochemistry* **2004**, *43*, 3772–3782.
[42] Sixma, T. K.; Kalk, K. H.; van Zanten, B. A. M.; Dauter, Z.; Kingma, J.; Witholt, B.; Hol, W. G. J. Refined Structure of Escherichia Coli Heat-Labile Enterotoxin, a Close Relative of Cholera Toxin. *Journal of Molecular Biology*. 1993, pp 890–918. https://doi.org/10.1006/jmbi.1993.1209.
[43] Horstman, A. L.; Kuehn, M. J. Enterotoxigenic Escherichia Coli Secretes Active Heat-Labile Enterotoxin via Outer Membrane Vesicles. *J. Biol. Chem.* **2000**, *275* (17), 12489–12496. https://doi.org/10.1074/jbc.275.17.12489.
[44] Guyette, J. Unraveling PDI and Its Interaction with AB Toxins, University of Central Florida, 2019.
[45] A.Merritt, E.; Sarfaty, S.; Van Den Akker, F.; L'hoir, C.; Marital, J. A.; Hol, W. G. J. Crystal Structureof Cholera Toxin B-Pentamer Bound to Receptor GMl Pentasaccharide. *Protein Sci.* **1994**, *3*, 166–175.

[46] Chinnapen, D. J. F.; Chinnapen, H.; Saslowsky, D.; Lencer, W. I. Rafting with Cholera Toxin: Endocytosis and Trafficking from Plasma Membrane to ER. *FEMS Microbiol. Lett.* **2007**, *266* (2), 129–137. https://doi.org/10.1111/j.1574-6968.2006.00545.x.

[47] Wernick, N. L. B.; Chinnapen, D. J. F.; Cho, J. A.; Lencer, W. I. Cholera Toxin: An Intracellular Journey into the Cytosol by Way of the Endoplasmic Reticulum. *Toxins (Basel).* **2010**, *2* (3), 310–325. https://doi.org/10.3390/toxins2030310.

[48] Taylor, M.; Banerjee, T.; Ray, S.; Tatulian, S. A.; Teter, K. Protein-Disulfide Isomerase Displaces the Cholera Toxin A1 Subunit from the Holotoxin without Unfolding the A1 Subunit. *J. Biol. Chem.* **2011**, *286* (25), 22090–22100. https://doi.org/10.1074/jbc.M111.237966.

[49] Pande, A. H.; Scaglione, P.; Taylor, M.; Nemec, K. N.; Tuthill, S.; Moe, D.; Holmes, R. K.; Tatulian, S. A.; Teter, K. Conformational Instability of the Cholera Toxin A1 Polypeptide. *J. Mol. Biol.* **2007**, *374* (4), 1114–1128. https://doi.org/10.1016/j.jmb.2007.10.025.

[50] Rodighiero, C.; Tsai, B.; Rapoport, T. A.; Lencer, W. I. Role of Ubiquitination in Retro-Translocation of Cholera Toxin and Escape of Cytosolic Degradation. *EMBO Rep.* **2002**, *3* (12), 1222–1227. https://doi.org/10.1093/embo-reports/kvf239.

[51] Burress, H.; Taylor, M.; Banerjee, T.; Tatulian, S. A.; Teter, K. Co- and Post-Translocation Roles for HSP90 in Cholera Intoxication. *J. Biol. Chem.* **2014**, *289* (48), 33644–33654. https://doi.org/10.1074/jbc.M114.609800.

[52] Burress, H.; Kellner, A.; Guyette, J.; Tatulian, S. A.; Teter, K. HSC70 and HSP90 Chaperones Perform Complementary Roles in Translocation of the Cholera Toxin A1 Subunit from the Endoplasmic Reticulum to the Cytosol. *J. Biol. Chem.* **2019**, *294* (32), 12122–12131. https://doi.org/10.1074/jbc.RA119.008568.

[53] Kahn, R. A.; Gilman, A. G. Purification of a Protein Cofactor Required for ADP-Ribosylation of the Stimulatory Regulatory Component of Adenylate Cyclase by Cholera Toxin. *J. Biol. Chem.* **1984**, *259* (10), 6228–6234. https://doi.org/https://doi.org/10.1016/S0021-9258(20)82130-9.

[54] O'Neal, C. J.; Jobling, M. G.; Holmes, R. K.; Hol, W. G. J. Structural Basis for the Activation of Cholera Toxin by Human ARF6-GTP. *Science (80-.).* **2005**, *309* (5737), 1093–1096. https://doi.org/10.1126/science.1113398.

[55] Gill, D. M.; Meren, R. ADP-Ribosylation of Membrane Proteins Catalyzed by Cholera Toxin: Basis of the Activation of Adenylate Cyclase. *Proc. Natl. Acad. Sci.* **1978**, *75* (7), 3050 LP – 3054. https://doi.org/10.1073/pnas.75.7.3050.

[56] Cassel, D.; Selinger, Z. Mechanism of Adenylate Cyclase Activation by Cholera Toxin: Inhibition of GTP Hydrolysis at the Regulatory Site. *Proc. Natl. Acad. Sci.* **1977**, *74* (8), 3307 LP – 3311. https://doi.org/10.1073/pnas.74.8.3307.

[57] Cheng, S. H.; Rich, D. P.; Marshall, J.; Gregory, R. J.; Welsh, M. J.; Smith, A. E. Phosphorylation of the R Domain by CAMP-Dependent Protein Kinase Regulates the CFTR Chloride Channel. *Cell* **1991**, *66* (5), 1027–1036. https://doi.org/https://doi.org/10.1016/0092-8674(91)90446-6.

[58] Moss, J.; Richardson, S. H. Activation of Adenylate Cyclase by Heat-Labile Escherichia Coli Enterotoxin. Evidence for ADP-Ribosyltransferase Activity Similar to That of Choleragen. *J. Clin. Invest.* **1978**, *62* (2), 281–285. https://doi.org/10.1172/JCI109127.

[59] Spangler, B. D. Structure and Function of Cholera Toxin and the Related Escherichia Coli Heat-Labile Enterotoxin. *Microbiol. Rev.* **1992**, *56* (4), 622 LP – 647.

[60] Hartl, F. U.; Bracher, A.; Hayer-Hartl, M. Molecular Chaperones in Protein Folding and Proteostasis. *Nature* **2011**, *475* (7356), 324–332.

https://doi.org/10.1038/nature10317.
[61] Spooner, R. A.; Hart, P. J.; Cook, J. P.; Pietroni, P.; Rogon, C.; Höhfeld, J.; Roberts, L. M.; Lord, J. M. Cytosolic Chaperones Influence the Fate of a Toxin Dislocated from the Endoplasmic Reticulum. *Proc. Natl. Acad. Sci.* **2008**, *105* (45), 17408 LP – 17413. https://doi.org/10.1073/pnas.0809013105.
[62] Schuster, M.; Schnell, L.; Feigl, P.; Birkhofer, C.; Mohr, K.; Roeder, M.; Carle, S.; Langer, S.; Tippel, F.; Buchner, J.; Fischer, G.; Hausch, F.; Frick, M.; Schwan, C.; Aktories, K.; Schiene-Fischer, C.; Barth, H. The Hsp90 Machinery Facilitates the Transport of Diphtheria Toxin into Human Cells. *Sci. Rep.* **2017**, *7* (1), 613. https://doi.org/10.1038/s41598-017-00780-x.
[63] Yu, M.; Haslam, D. B. Shiga Toxin Is Transported from the Endoplasmic Reticulum Following Interaction with the Luminal Chaperone HEDJ/ERdj3. *Infect. Immun.* **2005**, *73* (4), 2524 LP – 2532. https://doi.org/10.1128/IAI.73.4.2524-2532.2005.
[64] Wang, L.; Wang, X.; Wang, C. C. Protein Disulfide-Isomerase, a Folding Catalyst and a Redox-Regulated Chaperone. *Free Radic. Biol. Med.* **2015**, *83*, 305–313. https://doi.org/10.1016/j.freeradbiomed.2015.02.007.
[65] Khan, H. A.; Mutus, B. Protein Disulfide Isomerase a Multifunctional Protein with Multiple Physiological Roles. *Front. Chem.* **2014**, *2* (AUG), 1–9. https://doi.org/10.3389/fchem.2014.00070.
[66] Guyette, J.; Evangelista, B.; Tatulian, S. A.; Teter, K. Stability and Conformational Resilience of Protein Disulfide Isomerase. *Biochemistry* **2019**, *58* (34), 3572–3584. https://doi.org/10.1021/acs.biochem.9b00405.
[67] Wang, C.; Li, W.; Ren, J.; Fang, I.; Ke, H.; Gong, W.; Feng, W.; Wang1, C. Structural Insights into the Redox-Regulated Dynamic Conformations of Human Protein Disulfide Isomerase. *Antioxid. Redox Signal.* **2013**, *19* (1), 36–45.
[68] Okumura, M.; Noi, K.; Kanemura, S.; Kinoshita, M.; Saio, T.; Inoue, Y.; Hikima, T.; Akiyama, S.; Ogura, T.; Inaba, K. Dynamic Assembly of Protein Disulfide Isomerase in Catalysis of Oxidative Folding. *Nat. Chem. Biol.* **2019**, *15* (5), 499–509. https://doi.org/10.1038/s41589-019-0268-8.
[69] Cai, H.; Wang, C. C.; Tsou, C. L. Chaperone-like Activity of Protein Disulfide Isomerase in the Refolding of a Protein with No Disulfide Bonds. *J. Biol. Chem.* **1994**, *269* (40), 24550–24552. https://doi.org/https://doi.org/10.1016/S0021-9258(17)31426-6.
[70] Khan, H. A.; Mutus, B. Protein Disulfide Isomerase a Multifunctional Protein with Multiple Physiological Roles. *Front. Chem.* **2014**, *2* (70), 1–9. https://doi.org/10.3389/fchem.2014.00070.
[71] Serrano, A.; Qiao, X.; Matos, J. O.; Farley, L.; Cilenti, L.; Chen, B.; Tatulian, S. A.; Teter, K. Reversal of Alpha-Synuclein Fibrillization by Protein Disulfide Isomerase. *Front. Cell Dev. Biol.* **2020**, *8*, 726. https://doi.org/10.3389/fcell.2020.00726.
[72] Taylor, M.; Burress, H.; Banerjee, T.; Ray, S.; Curtis, D.; Tatulian, S. A.; Teter, K. Substrate-Induced Unfolding of Protein Disulfide Isomerase Displaces the Cholera Toxin A1 Subunit from Its Holotoxin. *PLoS Pathog.* **2014**, *10* (2). https://doi.org/10.1371/journal.ppat.1003925.
[73] Orlandi, P. A. Protein-Disulfide Isomerase-Mediated Reduction of the A Subunit of Cholera Toxin in a Human Intestinal Cell Line. *J. Biol. Chem.* **1997**, *272* (7), 4591–4599.
[74] Majoul, I.; Ferrari, D.; Söling, H. D. 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.* **1997**, *401* (2–3), 104–108.

https://doi.org/10.1016/s0014-5793(96)01447-0.
[75] Tsai, B.; Rodighiero, C.; Lencer, W. I.; Rapoport, T. A. Protein Disulfide Isomerase Acts as a Redox-Dependent Chaperone to Unfold Cholera Toxin. *Cell* **2001**, *104* (6), 937–948. https://doi.org/10.1016/S0092-8674(01)00289-6.
[76] Cherubin, P.; Guyette, J.; Taylor, M.; O'Donnell, M.; Herndon, L.; Burress, H.; Riad, A.; Tatulian, S. A.; Teter, K. Protein Disulfide Isomerase Does Not Act as an Unfoldase in the Disassembly of Cholera Toxin. *Biosci. Rep.* **2018**, *38* (5), 1–15. https://doi.org/10.1042/BSR20181320.
[77] Banerjee, T.; Pande, A.; Jobling, M. G.; Taylor, M.; Massey, S.; Holmes, R. K.; Tatulian, S. A.; Teter, K. Contribution of Subdomain Structure to the Thermal Stability of the Cholera Toxin A1 Subunit. *Biochemistry* **2010**, *49* (41), 8839–8846. https://doi.org/10.1021/bi101201c.
[78] Massey, S.; Banerjee, T.; Pande, A. H.; Taylor, M.; Tatulian, S. A.; Teter, K. Stabilization of the Tertiary Structure of the Cholera Toxin A1 Subunit Inhibits Toxin Dislocation and Cellular Intoxication. *J. Mol. Biol.* **2009**, *393* (5), 1083–1096. https://doi.org/10.1016/j.jmb.2009.09.013.
[79] Tadesse, L.; Nazarbaghi, R.; Walters, L. Isotopically Enhanced Infrared Spectroscopy: A Novel Method for Examining Secondary Structure at Specific Sites in Conformationally Heterogeneous Peptides. *J. Am. Chem. Soc.* **1991**, *113* (18), 7036–7037. https://doi.org/10.1021/ja00018a052.
[80] Zhang, M.; Vogel, H. J.; Fabian, H.; Mantsch, H. H. Isotope-Edited Fourier Transform Infrared Spectroscopy Studies of Calmodulin's Interaction with Its Target Peptides. *Biochemistry* **1994**, *33* (36), 10883–10888. https://doi.org/10.1021/bi00202a006.
[81] Guyette, J.; Cherubin, P.; Serrano, A.; Taylor, M.; Abedin, F.; O'Donnell, M.; Burress, H.; Tatulian, S. A.; Teter, K. Quercetin-3-Rutinoside Blocks the Disassembly of Cholera Toxin by Protein Disulfide Isomerase. *Toxins (Basel).* **2019**, *11* (8), 458. https://doi.org/10.3390/toxins11080458.
[82] Lin, L.; Gopal, S.; Sharda, A.; Passam, F.; Bowley, S. R.; Stopa, J.; Xue, G.; Yuan, C.; Furie, B. C.; Flaumenhaft, R.; Huang, M.; Furie, B. Quercetin-3-Rutinoside Inhibits Protein Disulfide Isomerase by Binding to Its B′x Domain. *J. Biol. Chem.* **2015**, *290* (39), 23543–23552. https://doi.org/10.1074/jbc.M115.666180.
[83] Petsko, G. A.; Ringe, D. *Protein Structure and Function*; Lawrence, E., Robertson, M., Eds.; New Science Press Ltd: London, 2004.
[84] Smyth, M. S.; Martin, J. H. J. X-Ray Crystallography. *Mol. Pathol.* **2000**, *53* (1), 8–14. https://doi.org/10.1136/mp.53.1.8.
[85] Holcomb, J.; Spellmon, N.; Zhang, Y.; Doughan, M.; Li, C.; Yang, Z. Protein Crystallization: Eluding the Bottleneck of X-Ray Crystallography. *AIMS Biophys.* **2017**, *4* (4), 557–575. https://doi.org/10.3934/biophy.2017.4.557.
[86] Hampton Research. Crystal Growth 101.
[87] Danley, D. Crystallization to Obtain Protein-Ligand Complexes for Structure-Aided Drug Design. *Acta Crystallogr. Sect. D* **2006**, *62* (6), 569–575.
[88] Cowtan, K. Phase Problem in X-Ray Crystallography, and Its Solution. *Encycl. Life Sci.* **2003**, 1–5. https://doi.org/10.1038/npg.els.0002722.
[89] Taylor, G. The Phase Problem. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **2003**, *59* (11), 1881–1890. https://doi.org/10.1107/S0907444903017815.
[90] Bhella, D. Cryo-Electron Microscopy: An Introduction to the Technique, and Considerations When Working to Establish a National Facility. *Biophys. Rev.* **2019**, *11* (4), 515–519. https://doi.org/10.1007/s12551-019-00571-w.

[91] Liu, S.; Hattne, J.; Reyes, F. E.; Sanchez-Martinez, S.; Jason de la Cruz, M.; Shi, D.; Gonen, T. Atomic Resolution Structure Determination by the Cryo-EM Method MicroED. *Protein Sci.* **2017**, *26* (1), 8–15. https://doi.org/https://doi.org/10.1002/pro.2989.

[92] Yip, K. M.; Fischer, N.; Paknia, E.; Chari, A.; Stark, H. Atomic-Resolution Protein Structure Determination by Cryo-EM. *Nature* **2020**, *587* (7832), 157–161. https://doi.org/10.1038/s41586-020-2833-4.

[93] Nakane, T.; Kotecha, A.; Sente, A.; McMullan, G.; Masiulis, S.; Brown, P. M. G. E.; Grigoras, I. T.; Malinauskaite, L.; Malinauskas, T.; Miehling, J.; Uchański, T.; Yu, L.; Karia, D.; Pechnikova, E. V; de Jong, E.; Keizer, J.; Bischoff, M.; McCormack, J.; Tiemeijer, P.; Hardwick, S. W.; Chirgadze, D. Y.; Murshudov, G.; Aricescu, A. R.; Scheres, S. H. W. Single-Particle Cryo-EM at Atomic Resolution. *Nature* **2020**, *587* (7832), 152–156. https://doi.org/10.1038/s41586-020-2829-0.

[94] Henderson, R.; Clarke, B. C. Cryo-Protection of Protein Crystals against Radiation Damage in Electron and X-Ray Diffraction. *Proc. R. Soc. London. Ser. B Biol. Sci.* **1990**, *241* (1300), 6–8. https://doi.org/10.1098/rspb.1990.0057.

[95] Serna, M. Hands on Methods for High Resolution Cryo-Electron Microscopy Structures of Heterogeneous Macromolecular Complexes . *Frontiers in Molecular Biosciences* . 2019, p 33.

[96] Wu, M.; Lander, G. C. How Low Can We Go? Structure Determination of Small Biological Complexes Using Single-Particle Cryo-EM. *Curr. Opin. Struct. Biol.* **2020**, *64*, 9–16. https://doi.org/10.1016/j.sbi.2020.05.007.

[97] Merk, A.; Bartesaghi, A.; Banerjee, S.; Falconieri, V.; Rao, P.; Davis, M. I.; Pragani, R.; Boxer, M. B.; Earl, L. A.; Milne, J. L. S.; Subramaniam, S. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell* **2016**, *165* (7), 1698–1707. https://doi.org/10.1016/j.cell.2016.05.040.

[98] Fan, X.; Wang, J.; Zhang, X.; Yang, Z.; Zhang, J.-C.; Zhao, L.; Peng, H.-L.; Lei, J.; Wang, H.-W. Single Particle Cryo-EM Reconstruction of 52 KDa Streptavidin at 3.2 Angstrom Resolution. *Nat. Commun.* **2019**, *10* (1), 2386. https://doi.org/10.1038/s41467-019-10368-w.

[99] Herzik, M. A.; Wu, M.; Lander, G. C. High-Resolution Structure Determination of Sub-100 KDa Complexes Using Conventional Cryo-EM. *Nat. Commun.* **2019**, *10* (1), 1–9. https://doi.org/10.1038/s41467-019-08991-8.

[100] Zhang, K.; Li, S.; Kappel, K.; Pintilie, G.; Su, Z.; Mou, T.-C.; Schmid, M. F.; Das, R.; Chiu, W. Cryo-EM Structure of a 40 KDa SAM-IV Riboswitch RNA at 3.7 Å Resolution. *Nat. Commun.* **2019**, *10* (1), 5511. https://doi.org/10.1038/s41467-019-13494-7.

[101] Scarff, C. A.; Fuller, M. J. G.; Thompson, R. F.; Iadanza, M. G. Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems. *JoVE* **2018**, No. 132, e57199. https://doi.org/doi:10.3791/57199.

[102] Weiss, T. M. Small Angle Scattering: Historical Perspective and Future Outlook. In *Advances in experimental medicine and biology*; Springer Nature: Singapore, 2017; Vol. 1009, pp 1–10. https://doi.org/10.1007/978-981-10-6038-0_1.

[103] Jeffries, C. M.; Graewert, M. A.; Blanchet, C. E.; Langley, D. B.; Whitten, A. E.; Svergun, D. I. Preparing Monodisperse Macromolecular Samples for Successful Biological Small-Angle X-Ray and Neutron-Scattering Experiments. *Nat. Protoc.* **2016**, *11* (11), 2122–2153. https://doi.org/10.1038/nprot.2016.113.

[104] Prior, C.; Davies, O. R.; Bruce, D.; Pohl, E. Obtaining Tertiary Protein Structures by the Ab Initio Interpretation of Small Angle X-Ray Scattering Data. *J. Chem. Theory Comput.* **2020**, *16* (3), 1985–2001. https://doi.org/10.1021/acs.jctc.9b01010.

[105] Feigin, L. .; Svergun, D. I. *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*; Plenum Press: New York, 1987.
[106] Dmitri I. Svergun; Koch, M. H. J.; Timmins, P. A.; May, R. P. *Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules*, 1st ed.; Oxford University Press, 2013.
[107] Gräwert, M.; Svergun, D. A Beginner's Guide to Solution Small-Angle X-Ray Scattering (SAXS). *Biochem. (Lond).* **2020**, *42* (1), 36–42. https://doi.org/10.1042/BIO04201036.
[108] Da Vela, S.; Svergun, D. I. Methods, Development and Applications of Small-Angle X-Ray Scattering to Characterize Biological Macromolecules in Solution. *Curr. Res. Struct. Biol.* **2020**, *2* (August), 164–170. https://doi.org/10.1016/j.crstbi.2020.08.004.
[109] Brennich, M.; Pernot, P.; Round, A. How to Analyze and Present SAS Data for Publication. In *Biological Small Angle Scattering: Techniques, Strategies and Tips*; Chaudhuri, B., Muñoz, I. G., Qian, S., Urban, V. S., Eds.; Springer Singapore: Singapore, 2017; pp 47–64. https://doi.org/10.1007/978-981-10-6038-0_4.
[110] Krueger, S. Designing and Performing Biological Solution Small-Angle Neutron Scattering Contrast Variation Experiments on Multi-Component Assemblies. In *Biological Small Angle Scattering: Techniques, Strategies and Tips*; Chaudhuri, B., Muñoz, I. G., Qian, S., Urban, V. S., Eds.; Springer Singapore: Singapore, 2017; pp 65–85. https://doi.org/10.1007/978-981-10-6038-0_5.
[111] Svergun, D. I.; Koch, M. H. J. Small-Angle Scattering Studies of Biological Macromolecules in Solution. *Reports Prog. Phys.* **2003**, *66* (10), 1735–1782. https://doi.org/10.1088/0034-4885/66/10/r05.
[112] Kikhney, A. G.; Svergun, D. I. A Practical Guide to Small Angle X-Ray Scattering (SAXS) of Flexible and Intrinsically Disordered Proteins. *FEBS Lett.* **2015**, *589* (19PartA), 2570–2577. https://doi.org/https://doi.org/10.1016/j.febslet.2015.08.027.
[113] Heller, W. T. Small-Angle Neutron Scattering and Contrast Variation: A Powerful Combination for Studying Biological Structures. *Acta Crystallogr. Sect. D* **2010**, *66* (11), 1213–1217. https://doi.org/https://doi.org/10.1107/S0907444910017658.
[114] Dunne, O.; Weidenhaupt, M.; Callow, P.; Martel, A.; Moulin, M.; Perkins, S. J.; Haertlein, M.; Forsyth, V. T. Matchout Deuterium Labelling of Proteins for Small-Angle Neutron Scattering Studies Using Prokaryotic and Eukaryotic Expression Systems and High Cell-Density Cultures. *Eur. Biophys. J.* **2017**, *46* (5), 425–432. https://doi.org/10.1007/s00249-016-1186-2.
[115] Douzi, B. Protein–Protein Interactions: Surface Plasmon Resonance. In *Bacterial Protein Secretion Systems: Methods and Protocols*; Journet, L., Cascales, E., Eds.; Springer New York: New York, NY, 2017; pp 257–275. https://doi.org/10.1007/978-1-4939-7033-9_21.
[116] Karlsson, R.; Michaelsson, A.; Mattsson, L. Kinetic Analysis of Monoclonal Antibody-Antigen Interactions with a New Biosensor Based Analytical System. *J. Immunol. Methods* **1991**, *145* (1), 229–240. https://doi.org/https://doi.org/10.1016/0022-1759(91)90331-9.
[117] Yu, H. W.; Halonen, M. J.; Pepper, I. L. *Immunological Methods*; Elsevier Inc., 2015. https://doi.org/10.1016/B978-0-12-394626-3.00012-0.
[118] Jerabek-Willemsen, M.; André, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D. MicroScale Thermophoresis: Interaction Analysis and Beyond. *J.*

Mol. Struct. **2014**, *1077*, 101–113. https://doi.org/10.1016/j.molstruc.2014.03.009.
[119] Ludwig, C. Diffusion Zwischen Ungleich Erwärmten Orten Gleich Zusammengesetzter Lösungen. *Sitzungber Bayer Akad Wiss Wien Math- Naturwiss Kl* **1856**, *20*.
[120] NanoTemper Technologies GmbH. User Manual for the Monolith NT.115.
[121] Mueller, A. M.; Breitsprecher, D.; Duhr, S.; Baaske, P.; Schubert, T.; Längst, G. MicroScale Thermophoresis: A Rapid and Precise Method to Quantify Protein–Nucleic Acid Interactions in Solution. In *Functional Genomics: Methods and Protocols*; Kaufmann, M., Klinger, C., Savelsbergh, A., Eds.; Springer New York: New York, NY, 2017; pp 151–164. https://doi.org/10.1007/978-1-4939-7231-9_10.
[122] Harris, G.; Scott, D. J. Isothermal Titration Calorimetry: Feeling the Heat. *Biochem. (Lond).* **2019**, *41* (2), 4–7. https://doi.org/10.1042/bio04102004.
[123] Archer, W. R.; Schulz, M. D. Isothermal Titration Calorimetry: Practical Approaches and Current Applications in Soft Matter. *Soft Matter* **2020**, *16* (38), 8760–8774. https://doi.org/10.1039/d0sm01345e.
[124] ThermoFisher Scientfic. Overview of ELISA https://www.thermofisher.com/no/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-elisa.html (accessed May 1, 2021).
[125] Weng, Z.; Zhao, Q. Utilizing ELISA to Monitor Protein-Protein Interaction. In *Protein-Protein Interactions. Methods in Molecular Biology*; Meyerkord, C., Fu, H., Eds.; Humana Press: New York, NY, 2015; pp 341–352.
[126] Miura*, K. An Overview of Current Methods to Confirm Protein-Protein Interactions. *Protein & Peptide Letters*. 2018, pp 728–733. https://doi.org/http://dx.doi.org/10.2174/0929866525666180821122240.
[127] Jobling, M. G.; Palmer, L. M.; Erbe, J. L.; Holmes, R. K. Construction and Characterization of Versatile Cloning Vectors for Efficient Delivery of Native Foreign Proteins to the Periplasm of Escherichia Coli. *Plasmid* **1997**, *38* (3), 158–173. https://doi.org/10.1006/plas.1997.1309.
[128] Dertzbaugh, M. T.; Cox, L. M. The Affinity of Cholera Toxin for Ni2+ Ion. *Protein Eng.* **1998**, *11* (7), 577–581. https://doi.org/10.1093/protein/11.7.577.
[129] Jobling, M. G.; Gotow, L. F.; Yang, Z.; Holmes, R. K. A Mutational Analysis of Residues in Cholera Toxin A1 Necessary for Interaction with Its Substrate, the Stimulatory G Protein Gsα. *Toxins (Basel).* **2015**, *7* (3), 919–935. https://doi.org/10.3390/toxins7030919.
[130] Pedersen, J. A Flux- and Background-Optimized Version of the NanoSTAR Small-Angle X-Ray Scattering Camera for Solution Scattering. *J. Appl. Crystallogr.* **2004**, *37* (3), 369–380.
[131] Konarev, P. V; Volkov, V. V; Sokolova, A. V; Koch, M. H. J.; Svergun, D. I. PRIMUS: A Windows PC-Based System for Small-Angle Scattering Data Analysis. *J. Appl. Crystallogr.* **2003**, *36* (5), 1277–1282. https://doi.org/https://doi.org/10.1107/S0021889803012779.
[132] Franke, D.; Petoukhov, M. V; Konarev, P. V; Panjkovich, A.; Tuukkanen, A.; Mertens, H. D. T.; Kikhney, A. G.; Hajizadeh, N. R.; Franklin, J. M.; Jeffries, C. M.; Svergun, D. I. ATSAS 2.8: A Comprehensive Data Analysis Suite for Small-Angle Scattering from Macromolecular Solutions. *J. Appl. Crystallogr.* **2017**, *50* (4), 1212–1225.
[133] Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.;

Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat. Methods* **2019**, *16* (7), 595–602. https://doi.org/10.1038/s41592-019-0459-y.

[134] Volkov, V. V; Svergun, D. I. Uniqueness of Ab Initio Shape Determination in Small-Angle Scattering. *J. Appl. Crystallogr.* **2003**, *36* (3), 860–864.

[135] Franke, D.; Svergun, D. I. DAMMIF, a Program for Rapid Ab-Initio Shape Determination in Small-Angle Scattering. *J. Appl. Crystallogr.* **2009**, *42* (2), 342–346.

[136] Grudinin, S.; Garkavenko, M.; Kazennov, A. Pepsi-SAXS: An Adaptive Method for Rapid and Accurate Computation of Small-Angle X-Ray Scattering Profiles. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73* (5), 449–464. https://doi.org/10.1107/S2059798317005745.

[137] Kozin, M. B.; Svergun, D. I. Automated Matching of High- and Low-Resolution Structural Models. *J. Appl. Crystallogr.* **2001**, *34* (1), 33–41.

[138] Kozakov, D.; Hall, D. R.; Xia, B.; Porter, K. A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro Web Server for Protein-Protein Docking. *Nat. Protoc.* **2017**, *12* (2), 255–278. https://doi.org/10.1038/nprot.2016.169.

[139] Yan, Y.; Tao, H.; He, J.; Huang, S. Y. The HDOCK Server for Integrated Protein–Protein Docking. *Nat. Protoc.* **2020**, *15* (5), 1829–1852. https://doi.org/10.1038/s41596-020-0312-x.

[140] Dumon-Seignovert, L.; Cariot, G.; Vuillard, L. The Toxicity of Recombinant Proteins in Escherichia Coli: A Comparison of Overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr. Purif.* **2004**, *37* (1), 203–206. https://doi.org/10.1016/j.pep.2004.04.025.

[141] Cytiva. Superdex 200 Increase small-scale SEC columns https://www.cytivalifesciences.com/en/us/shop/chromatography/prepacked-columns/size-exclusion/superdex-200-increase-small-scale-size-exclusion-chromatography-columns-p-06190.

[142] Synthetic Genomics. Vmax™ Express Chemically Competent Cells-User Guide. 2017.

[143] Weinstock, M. T.; Hesek, E. D.; Wilson, C. M.; Gibson, D. G. Vibrio Natriegens as a Fast-Growing Host for Molecular Biology. *Nat. Methods* **2016**, *13* (10), 849–851. https://doi.org/10.1038/nmeth.3970.

[144] Pérez-Pérez, J.; Gutirrrez, J. An Arabinose-Inducible Expression Vector, PAR3, Compatible with ColEl-Derived Plasmids*. *Gene* **1995**, *158*, 141–142.

[145] Connell, T. D.; Metzger, D. J.; Wang, M.; Jobling, M. G.; Holmes, R. K. Initial Studies of the Structural Signal for Extracellular Transport of Cholera Toxin and Other Proteins Recognized by Vibrio Cholerae. *Infect. Immun.* **1995**, *63* (10), 4091–4098. https://doi.org/10.1128/iai.63.10.4091-4098.1995.

[146] Hoff, J.; Daniel, B.; Stukenberg, D.; Thuronyi, B. W.; Waldminghaus, T.; Fritz, G. Vibrio Natriegens: An Ultrafast-Growing Marine Bacterium as Emerging Synthetic Biology Chassis. *Environ. Microbiol.* **2020**, *22* (10), 4394–4408. https://doi.org/https://doi.org/10.1111/1462-2920.15128.

[147] Krengel, U.; Imberty, A. Chapter 2 - Crystallography and Lectin Structure Database.

In *Lectins-Analytical technologies*; Nilsson, C. L. B. T.-L., Ed.; Elsevier Science B.V.: Amsterdam, 2007; pp 15–50. https://doi.org/https://doi.org/10.1016/B978-044453077-6/50003-X.
[148] Hu, C.-H.; Tsou, C.-L. C-Terminal Truncation of Bovine Protein Disulfide Isomerase Increases Its Activity. *Biochem. Biophys. Res. Commun.* **1992**, *183* (2), 714–718. https://doi.org/https://doi.org/10.1016/0006-291X(92)90541-R.
[149] Yu, X.; Wang, C.-C.; Tsou, C.-L. Association and Dissociation of Protein Disulfide Isomerase. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1994**, *1207* (1), 109–113. https://doi.org/https://doi.org/10.1016/0167-4838(94)90058-2.
[150] Pace, M.; Dixon, J. . The Nature of Multiple Forms of Bovine Thiol: Protein Disulfide Oxidoreductase. *Int. J. Pept. Protein Res.* **1979**, *14*, 409–413.
[151] Wallis, A. K.; Sidhu, A.; Byrne, L. J.; Howard, M. J.; Ruddock, L. W.; Williamson, R. A.; Freedman, R. B. The Ligand-Binding B′ Domain of Human Protein Disulphide-Isomerase Mediates Homodimerization. *Protein Sci.* **2009**, *18* (12), 2569–2577. https://doi.org/https://doi.org/10.1002/pro.270.
[152] Bastos-Aristizabal, S.; Kozlov, G.; Gehring, K. Structural Insight into the Dimerization of Human Protein Disulfide Isomerase. *Protein Sci.* **2014**, *23* (5), 618–626. https://doi.org/https://doi.org/10.1002/pro.2444.
[153] Morjana, N. A.; McKeone, B. J.; Gilbert, H. F. Guanidine Hydrochloride Stabilization of a Partially Unfolded Intermediate during the Reversible Denaturation of Protein Disulfide Isomerase. *Proc. Natl. Acad. Sci.* **1993**, *90* (6), 2107 LP – 2111. https://doi.org/10.1073/pnas.90.6.2107.
[154] Cai, M.; Huang, Y.; Yang, R.; Craigie, R.; Clore, G. M. A Simple and Robust Protocol for High-Yield Expression of Perdeuterated Proteins in Escherichia Coli Grown in Shaker Flasks. *J. Biomol. NMR* **2016**, *66* (2), 85–91. https://doi.org/10.1007/s10858-016-0052-y.
[155] Sørensen, H. V. Of Shellfish and Men - Applying X-Ray and Neutron Techniques to Surface-Active Bacterial Colonization Factors, University of Oslo, 2021.
[156] Hampton Research. *TCEP HCl-User Guide*; 2019.
[157] Getz, E. B.; Xiao, M.; Chakrabarty, T.; Cooke, R.; Selvin, P. R. A Comparison between the Sulfhydryl Reductants Tris(2-Carboxyethyl)Phosphine and Dithiothreitol for Use in Protein Biochemistry. *Anal. Biochem.* **1999**, *273* (1), 73–80. https://doi.org/https://doi.org/10.1006/abio.1999.4203.
[158] Novák, P.; Kruppa, G. H. Intra-Molecular Cross-Linking of Acidic Residues for Protein Structure Studies. *Eur. J. Mass Spectrom.* **2008**, *14* (6), 355–365. https://doi.org/10.1255/ejms.963.
[159] Li, S. J.; Hong, X. G.; Shi, Y. Y.; Li, H.; Wang, C. C. Annular Arrangement and Collaborative Actions of Four Domains of Protein-Disulfide Isomerase: A Small Angle x-Ray Scattering Study in Solution. *J. Biol. Chem.* **2006**, *281* (10), 6581–6588. https://doi.org/10.1074/jbc.M508422200.
[160] Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta - Bioenerg.* **2007**, *1767* (9), 1073–1101. https://doi.org/https://doi.org/10.1016/j.bbabio.2007.06.004.
[161] Ismail, A. A.; van de Voort, F. R.; Sedman, J. Chapter 4 Fourier Transform Infrared Spectroscopy: Principles and Applications. In *Instrumental Methods in Food Analysis*; Paré, J. R. J., Bélanger, J. M. R. B. T.-T. and I. in A. C., Eds.; Elsevier, 1997; Vol. 18, pp 93–139. https://doi.org/https://doi.org/10.1016/S0167-9244(97)80013-3.
[162] Donohue Rolfe, A.; Keusch, G. T. Shigella Dysenteriae 1 Cytotoxin: Periplasmic

Protein Releasable by Polymyxin B and Osmotic Shock. *Infect. Immun.* **1983**, *39* (1), 270–274. https://doi.org/10.1128/iai.39.1.270-274.1983.

7. Appendices

Section A: Software, instruments, kits, disposables

Table S1. Software and suppliers

Software	Supplier
ATSAS 3.0.2	Private, reference [132]
Biorender	Private, reference
ClusPro	Private, reference [138]
HDOCK	Private, reference [139]
MATLAB ® (R2021a)	Mathworks ®
ProtParam tool	ExPasy
PyMOL 2.4.0	Schrödinger, Inc.
Rock Maker ®	Formulatrix ®
SnapGene® 5.0.7	GSL Biotech LLC
UNICORNTM 5.31	GE Healthcare
Wasp Run	Douglas Instruments

Table S2. Instruments and suppliers

Instrument	Supplier
5810R bench-top refrigerated centrifuge	Eppendorf
ÄKTA™ pure 150	GE Healthcare
ÄKTA™ purifier-900	GE Healthcare
Avanti J-26 XP centrifuge	Beckman-Coulter
Biofuge Fresco centrifuge	Haraeus
Capsulefuge TOMY PMC-060	Tomytech
CO8000 OD600 spectrophotometer	BioChrom-WPA
Dri-Block DB 2A heating block	Techne
Electrophoresis powersupply-EPS601	GE Healthcare
Eppendorf Comfort Thermomixer	Eppendorf
Formulator ® liquid handler	Formulatrix
J-Lite JLA-8.1000 Fixed-Angle Rotor	Beckman-Coulter
JA-25.50 Fixed-Angle Rotor	Beckman-Coulter
Kelvitron T incubator	Heraeus
LS Spectrometer	LS Instruments
Millipore Direct-Q5	Merck
Mini gel tank electrophoresis system	Invitrogen
Miniorbital shaker SSM1	Thermo Scientific
MS-3000 magnetic stirrer	BioSan
Multitron standard incubator-shaker	InforsHT
NanoDrop One- UV-Vis spectrophotometer	Thermo Scientific
Nanostar SAXS	Bruker Daltonics

Olympus BX41 microscope	Olympus
Oryx4 Protein Crystallization robot	Douglas Instruments
PCB scale	Kern
PowerPack HC	Bio-Rad
Prometheus NT.48 instrument	NanoTemper
Rocking platform	VWR
TC3000 PCR thermocycler	Bio-Rad Laboratories, Inc.
UV Transilluminator 2000	Bio-Rad
UltrafleXtreme instrument	Bruker Daltonics
Vapourline 80/135	VWR
Våntec-2000 detector	Bruker Daltonics
Vibra Cell Ultrasonic Processor	Sonics
VP 100 vacuum pump	VWR
Water bath (462-0556)	VWR

Table S3. Disposable materials and suppliers

Disposable	Supplier
1.5 mL Eppendorf tube	Eppendorf
15 mL Falcon tubes	Thermo Scientific
3 Lens Midi Crystallization Plate	SWISSCI
50 mL Falcon tubes	Thermo Scientific
Amicon Ultra Centrifugal Filter Units 10 000 MWCO	Merck
Amicon Ultra Centrifugal Filter Units 5 000 MWCO	Merck
BRAND PCR tubes, 0.2 mL	Merck
Bolt 4-12% Bis-Tris Plus polyacrylamide gel	Thermo Scientific
Cell scraper 25 cm blade 20 mm	VWR
Disposable cuvettes polystyrene 1.5 mL	VWR
L-shaped cell spreaders	VWR
Filter upper cup 500 mL bottle top filer 0.2 µm filter PES membrane	VWR
Filter upper cup 500 mL bottle top filer 0.45 µm filter PES membrane	VWR
Millex-GS MCE membrane 0.22 µm syringe filter	Millex
Petri dish in polystyrene	VWR
Pur-A-Lyzer ™ Midi 3 500 MWCO	Thermo Scientific
SnakeSkin Dialysis tubing™ 10 000 MWCO 22 mm diameter	Thermo Scientific

Table S4. Kits and suppliers

Table S5. Reagents (chemicals, enzymes, ladders, etc.) and suppliers

Reagent	Supplier
100 X MEM vitamin solution	HyClone
10X CutSmart Buffer	New England Biolabs
10X FastDigest Buffer	Fermentas (Thermo Scientific)
10X T4 DNA ligase Buffer	New England Biolabs
20X MES SDS Running Buffer	New England BIolabs
4X NuPAGE LDS Sample Buffer	Thermo Scientific
6X DNA loading Dye	Thermo Scientific
d_8-Glycerol	Eurisotop
Agar-agar	Merck
Ammonium chloride	VWR
Ammonium heptamolybdate tetrahydrate	Sigma
Ampicillin sodium salt	Applichem
Benzeonase nuclease	Sigma
Bis-Tris	Amresco
Boric acid	Sigma
Calcium chloride	VWR
Chloramphenicol	Fluka
Cobalt (II) chloride hexahydrate	Sigma Aldrich
Coomassi Brilliant Blue G-250	Fluka
Cupper chloride dihydrate	Sigma-Aldrich
Di-potassium hydrogen phosphate	Merck
Di-sodium hydrogen phosphate	VWR
DTT - (R,R)-1,4-dimercaptobutane-2,3-diol	Applichem
EDC-(1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride)	Thermo Scientific
EDTA disodiuem salt dihydrate	Sigma
Gentamicin sulfate	Sigma
Glycerol 95 %	Sigma
GSH-Glutathione (reduced)	Applichem
Imidazole	Sigma
Iron (II) sulfate heptahydrate	Sigma
IPTG - isopropyl beta-D-1- thiogalactopyranoside	Sigma
L-arabinose	Sigma
Lysozyme from chicken egg white	Sigma
Magnesium acetate tetrahydrate	Sigma-Aldrich
Magnesium chloride	Sigma
PEG 3350-olyethylene glycol 3350	Sigma
Peptone	Sigma
PMSF-phenylmethylsulfonyl fluoride	Fluka

Potassium sulfate	Sigma-Aldrich
Potassium dihydrogen phosphate	VWR
Protease Inhibitor cocktail	Roche
Quercetin-3-rutenoside hydrate	Sigma
Restriction enzyme BamHI	New England Biolabs
Restriction enzyme NcoI	New England Biolabs
SeeBlue Plus2 Pre-stained protein standard	Thermo Scientific
Shrimp Alkaline Phosphatase (rSAP)	New England Biolabs
SOC medium	Thermo Scientific
Sodium chloride	VWR
Sodium dihydrogen phosphate monohydrate	Merck
T4 DNA ligase	New England Biolabs
TCEP-tris(2-carboxyethyl)phosphine)	Sigma
Tris base - tris(hydroxymethyl)aminomethane	Sigma-Aldrich
Tri-sodium citrate dihydrate	Sigma
Trypsin type I from bovine pancreas	Sigma-Aldrich
Yeast extract	Sigma
Zinc sulfate	Sigma-Aldirch

Table S6. Cells and supliers

Cell line	Supplier
BL21 (DE3) pLysS Competent cells	Novagen
NEB ® 5-alpha Competent E. coli (High Efficiency)	New England Biolabs
Vmax™ Express Chemically Competent Cells	SGI-DNA
OverExpress™ C43 (DE3) Chemically Competent cells	Sigma-Aldrich

Table S7. Columns and suppliers

Columns	Supplier
HiLoad 16/60 Superdex 200 prep grade 120 mL column	GE Healthcare
HisTrap HP 5 mL column	GE Healthcare
HitrapQ™ 5 mL column	GE Healthcare
HitrapSP™ 5 mL column	GE Healthcare
Hitrap Talon® Crude 5 mL column	GE Healthcare
Superdex 200 Increase 10/30 GL	GE Healthcare

Section B: Solutions and buffers

Culture growth solutions (media, antibiotics, L-arabinose and IPTG)

Table S8. Composition of LB medium (1 L) *

	Volume	Amount
MQ-H_2O	1 L	-
Yeast extract		5 g
Peptone		10 g
NaCl		10 g

*Medium pH= 7.1. The components were dissolved in 800 mL of MQ-water and sterilized in the autoclave. Autoclaved water was then added to a final volume of 1 L.

Table S9. Composition of TB medium (2 L) *

	Volume	Amount
MQ-H_2O	2 L	-
Yeast extract		48 g
Peptone		24 g
NaCl		10 g
Glycerol	8 mL	
K_2HPO_4		1.8 g
KH_2PO_4		4.4 g

*Medium pH= 7.1. The components were dissolved sterilized in the autoclave. The phosphate salts were dissolved in 500 mL of MQ-H_2O and autoclaved separately. Glycerol was sterile-filtered, then added to autoclaved medium

Table S10. Composition of LB+v2 salts medium (1 L)

	Volume	Amount
MQ-H_2O	900 mL	-
Yeast extract		5 g
Peptone		10 g
NaCl		10 g
10 X v2 Salts	100 mL	

*Medium pH= 7.1. The components were dissolved in 800 mL of MQ-water and sterilized in the autoclave. Autoclaved water was then added to a final volume of 1 L.

Table S11. Composition of 10X v2 salts

	Volume	Concentration	Amount
MQ-H_2O	1 L	-	-
NaCl		2.04 M	119.2 g
KCl		0.042	1.566 g
MgCl2 6 H2O		0.2314	23.42

Table S12. Minimal M9+ medium (1 L)*

	Volume	Amount

MQ-H_2O/D_2O	989 mL	-
K_2HPO_4		19.0 g
KH_2PO_4		5.0 g
Na_2HPO_4		9.0 g
K_2SO_4		2.4 g
NH_4Cl		5.0 g
Glycerol/d_8-Glycerol		18.0 g
$MgCl_2$		0.95 g
100X MEM vitamins	10 mL	
Trace element solution	1 mL	

*Both hydrogenated and deuterated media contain the same components except that deuterated glycerol and deuterated water was used in deuterated M9+. Salts and deuterated glycerol (dissolved in H_2O/D_2o) were autoclaved separately. MEM vitamins, $MgCl_2$ (dissolved in 10 mL H_2O/D_2O and sterile-filtered 0.22 μm) and trace element (sterile-filtered 0.22 μm) solution were added after autoclavation.

Table S13. Trace element solution (100 mL)*

	Volume	**Amount**
MQ-H_2O	100 mL	
$FeSO_4$ ($7H_2O$)		0.6 g
$CaCl_2$ ($2H_2O$)		0.6 g
$MnCl_2$ ($4H_2O$)		0.12 g
$CoCl_2$ ($6H_2O$)		0.08 g
$ZnSO_4$ ($7H_2O$)		0.07 g
$CuCl_2$ ($2H_2O$)		0.03 g
H_3BO_4		0.002 g
$(NH_4)_6Mo_7O_{24}$ ($4H_2O$)		0.025 g
EDTA		0.5 g

*The solution was sterile-filtered (0.22 2 μm) and stored at 20 ℃.

Table S14. Composition of 100 mg/mL ampicillin solution (20 mL) *

	Volume	**Concentration**	**Amount**
MQ-H_2O	20 mL	-	-
Ampicillin	-	100 mg/mL	2 g

*The solution was sterilized by filtration (0.2 μm) and stored in 1 mL aliquots at -20 ℃.

Table S15. Composition of 25 mg/mL chloramphenicol solution (20 mL)

	Volume	**Concentration**	**Amount**
MQ-H_2O	20 mL	-	-
Chloramphenicol	-	25 mg/mL	0.5 g

*The solution was sterilized by filtration (0.2 μm) and stored in 1 mL aliquots at -20 ℃.

Table S16. Composition of 20 % L-arabinose

	Volume	Concentration	Amount
MQ-H_2O	10 mL	-	-
L-arabinose	-	20 % w/v	2 g

*The solution was sterilized by filtration (0.2 μm) and either used immediately or stored at 4 ℃ for no longer than two weeks.

Table S17. Composition of 1 M IPTG

	Volume	Concentration	Amount
MQ-H_2O	20 mL	-	-
IPTG	-	1 M	4.77 g

*The solution was sterilized by filtration (0.2 μm) and stored in 1 mL aliquots at -20 ℃.

Protein purification buffers

All buffers were filtered through a 0.22 μm PES membrane filter and degassed.

Table S18. Composition of binding buffer for Talon Affinity Chromatography (Talon A; 1L)

	Volume	Concentration	Amount
MQ-H_2O	840 mL	-	
Na_2HPO_4 0.5 M	100 mL	50 mM	-
NaCl 5 M	60 mL	300 mM	-

pH 8.0, adjusted with HCl.

Table S19. Composition of elution buffer for Talon Affinity Chromatography (Talon B; 1 L)

	Volume	Concentration	Amount
MQ-H_2O	790 mL	-	-
Na_2HPO_4 0.5 M	100 mL	50 mM	-
NaCl 5 M	60 mL	300 mM	-
Imidazole 1 M	50 mL	50 mM	-

pH 8.0, adjusted with HCl.

Table S20. Composition of binding buffer for Galactose Affinity Chromatography (Gal A; 1 L)

	Volume	Concentration	Amount
MQ-H_2O	860 mL	-	-
Na_2HPO_4 0.5 M	100 mL	50 mM	-
NaCl 5 M	40 mL	200 mM	-

pH 7.4, adjusted with HCl.

Table S21. Composition of elution buffer for Galactose Affinity Chromatography (Gal B; 250 mL)

	Volume	Concentration	Amount
MQ-H_2O	215 mL	-	-
Na_2HPO_4 0.5 M	25 mL	50 mM	-
NaCl 5 M	10 mL	200 mM	-
Galactose	-	300 mM	13.5 g

pH 7.4, adjusted with HCl.

Table S22. Composition of binding buffer for Cation Exchange and Anion Exchange Chromatography (CT purification; 1 L)

	Volume	Concentration	Amount
MQ-H_2O	900 mL	-	-
Tris-HCl 0.5 M	100 mL	50 mM	-

pH 8.0, adjusted with HCl.

Table S23. Composition of elution buffer for Cation Exchange and Anion Exchange Chromatography (CT purification; 500 mL)

	Volume	Concentration	Amount
MQ-H_2O	350 mL	-	-
Tris-HCl 0.5 M	50 mL	50 mM	-
NaCl 5 M	100 mL	1 M	-

pH 8.0, adjusted with HCl.

Table S24. Composition of binding buffer His-tag affinity chromatography (PDI purification; 500 mL)

	Volume	Concentration	Amount
MQ-H_2O	420 mL	-	-
Tris-HCl 0.5 M	50 mL	50 mM	-
NaCl 5 M	20 mL	200 mM	-
Imidazole 1 M	10 mL	20 mM	-
TCEP*	-	5 mM	0.72 g

pH 7.5, adjusted with NaOH. *For production of oxidized PDI, TCEP was not included in the buffer.

Table S25. Composition of elution buffer His-tag affinity chromatography (PDI purification; 500 mL)

	Volume	Concentration	Amount
MQ-H_2O	355 mL	-	-
Tris-HCl 0.5 M	50 mL	50 mM	-
NaCl 5 M	20 mL	200 mM	-

Imidazole 1 M	75 mL	150 mM	-
TCEP*	-	5 mM	0.72 g

pH 7.5, adjusted with NaOH. *For production of oxidized PDI, TCEP was not included in the buffer.

Table S26. Composition of binding buffer Anion Exchange chromatography (PDI purification; 1 L)

	Volume	**Concentration**	**Amount**
MQ-H_2O	950 mL	-	-
Bis-Tris 1 M	50 mL	50 mM	-
DTT*	-	5 mM	0.77 g

pH 6.0, adjusted with HCl. *For production of oxidized PDI, DTT was not included in the buffer.

Table S27. Composition of elution buffer Anion Exchange chromatography (PDI purification; 500 mL)

	Volume	**Concentration**	**Amount**
MQ-H_2O	350 mL	-	-
Bis-Tris 1 M	50 mL	50 mM	-
NaCl 5 M	100 mL	1 M	-
DTT*	-	5 mM	0.385 g

pH 6.0, adjusted with HCl. *For production of oxidized PDI, DTT was not included in the buffer.

Table S28. Composition of 10X Phosphate-Buffered saline (PBS) (1 L)

	Volume	**Concentration**	**Amount**
MQ-H_2O	1000 mL	-	-
Na_2HPO_4 0.5 M	200 mL	100 mM	-
NaCl	-	1.37 M	80 g
KH_2PO_4	-	18 mM	2.4 g
KCl		27 mM	2 g

Table S29. Composition of 1X Phosphate-Buffered saline (PBS) (1 L)

	Volume	**Concentration**	**Amount**
MQ-H_2O	900 mL	-	-
10X PBS	100 mL	1X	-

pH 7.4, adjusted with HCl.

Table S30. Composition of Bis-tris buffer for SAXS experiments (500 mL)

	Volume	**Concentration**	**Amount**
MQ-H_2O	470 mL	-	-
Bis-Tris 1 M	10 mL	20 mM	-

NaCl 5 M	20 mL	200 M	-
TCEP	-	20 mM	2.87 g

pH 6.5, adjusted with NaOH.

Table S31. Composition of PBS buffer for SAXS experiments (500 mL)

	Volume	**Concentration**	**Amount**
MQ-H_2O	400 mL	-	-
10X PBS	100 mL	1X	-
GSH	-	2 mM	0.307 g

pH 7.4, adjusted with NaOH.

Section C: Cloning reactions

Table S32. Composition of digestion reaction with NcoI

	Volume	**Concentration**	**Amount**
MQ-H_2O	43 μL	-	-

10X rCutSmart™	5 μL	1X	-
DNA (pARCT5* or pLT-pUC57)		-	1 μg
NcoI-HF (NEB), 10 U/ μL	1 μL	-	10 U

Total volume: 50 μL. Reagents added in this order and gently mixed. Incubation at 37 ℃ for 60 min. Reaction inactivated at 65 ℃ for 15 min.

Table S33. Composition of dephosphorylation reaction of gel-purified pARCT5*

	Volume	Concentration	Amount
MQ-H_2O	-	-	-
10X rCutSmart™	2 μL	1X	-
Gel-purified pARCT5*, 8 ng/μL	17 μL	-	0.1 pmol of DNA ends
rSAP, 1 U/ μL	1 μL	-	1 U

Total volume: 20 μL. Reagents added in this order and gently mixed. Incubation at 37 ℃ for 30 min. Reaction inactivated at 65 ℃ for 5 min.

Table S34. Composition of ligation-mix with vector/insert ratio 1:2

	Volume	Concentration	Amount
MQ-H_2O	8.8 μL	-	-
10X T4 DNA ligase reaction Buffer	2 μL	1X	-
Vector (pARCT5*)	6 μL	-	50 ng
Insert	2.2 μL		
T4 DNA ligase	1 μL	-	-

Total volume: 20 μL. Reagents added in this order and gently mixed. Incubation at room temperature for 10 min. Reaction inactivated at 65 ℃ for 10 min.

Table S35. Composition of ligation-mix with vector/insert ratio 1:3

	Volume	Concentration	Amount
MQ-H_2O	7.8 μL	-	-
10X T4 DNA Ligase Reaction Buffer	2 μL	1X	-
Vector (pARCT5*)	6 μL	-	50 ng
Insert	3.2 μL		
T4 DNA ligase	1 μL	-	-

Total volume: 20 μL. Reagents added in this order and gently mixed. Incubation at room temperature for 10 min. Reaction inactivated at 65 ℃ for 10 min.

Table S36. Composition of diagnostic digestions of plasmids isolated from single colonies

	Volume	Concentration	Amount
MQ-H_2O	3 μL	-	-
10X FastDigest Green Buffer	1 μL	1X	-
Plasmid from single colony	5 μL	-	200 ng
FastDigest BamHI, 10 U/ μL	1 μL	0.2 U/ μL	10 U

Total volume: 10 μL. Reagents added in this order and gently mixed. Incubation at 37 ℃ for 10 min. Reaction inactivated at 80 ℃ for 5 min.

Section D: DNA oligonucleotides

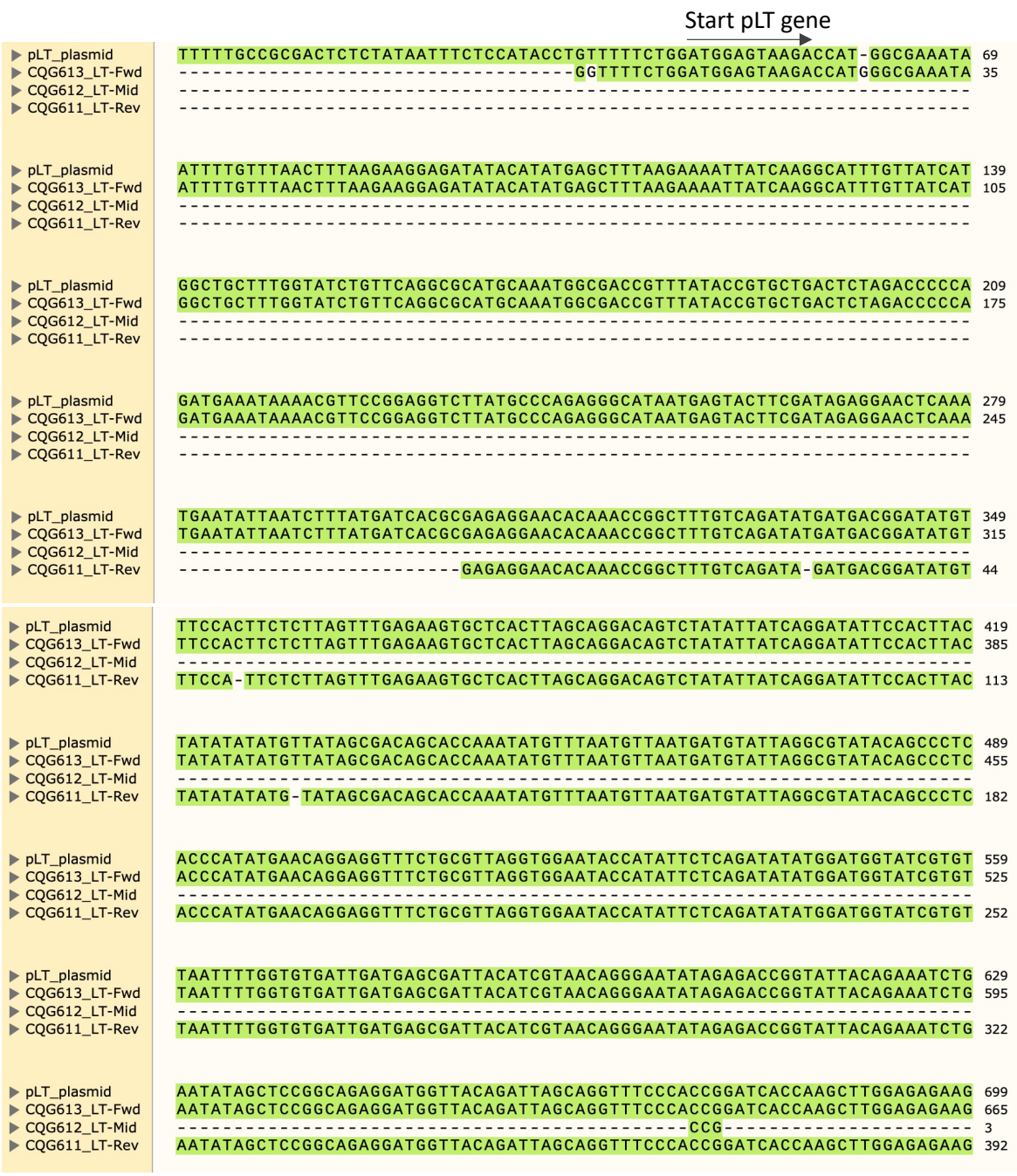
Table 37. DNA oligonucleotides

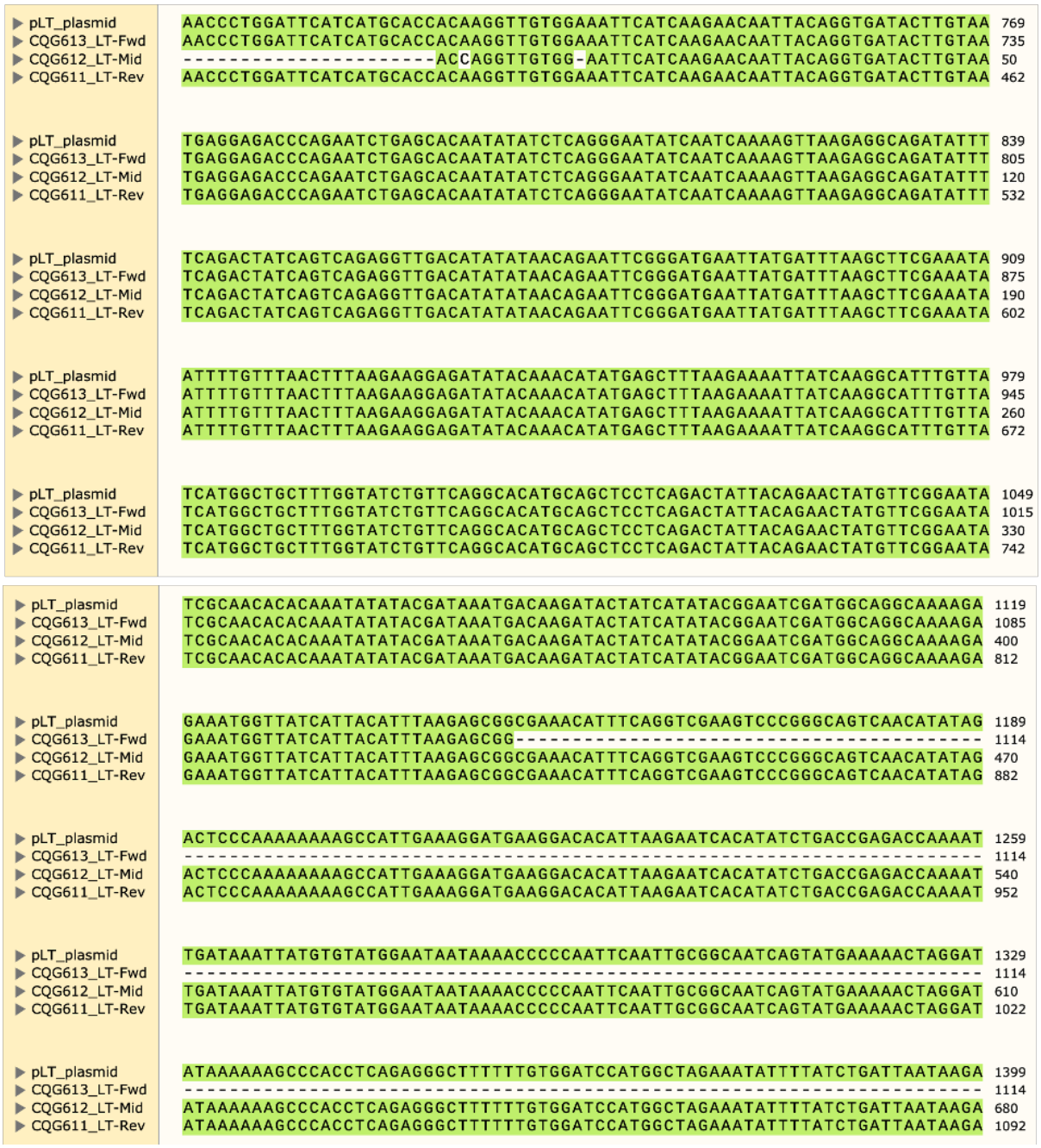
DNA oligo name	DNA sequence
NcoI_Fwd	CCGTTTTCACCTATGGGCAAATATTATAC
NcoI_Rev	GGGCGAAGAAGTTGTCCA

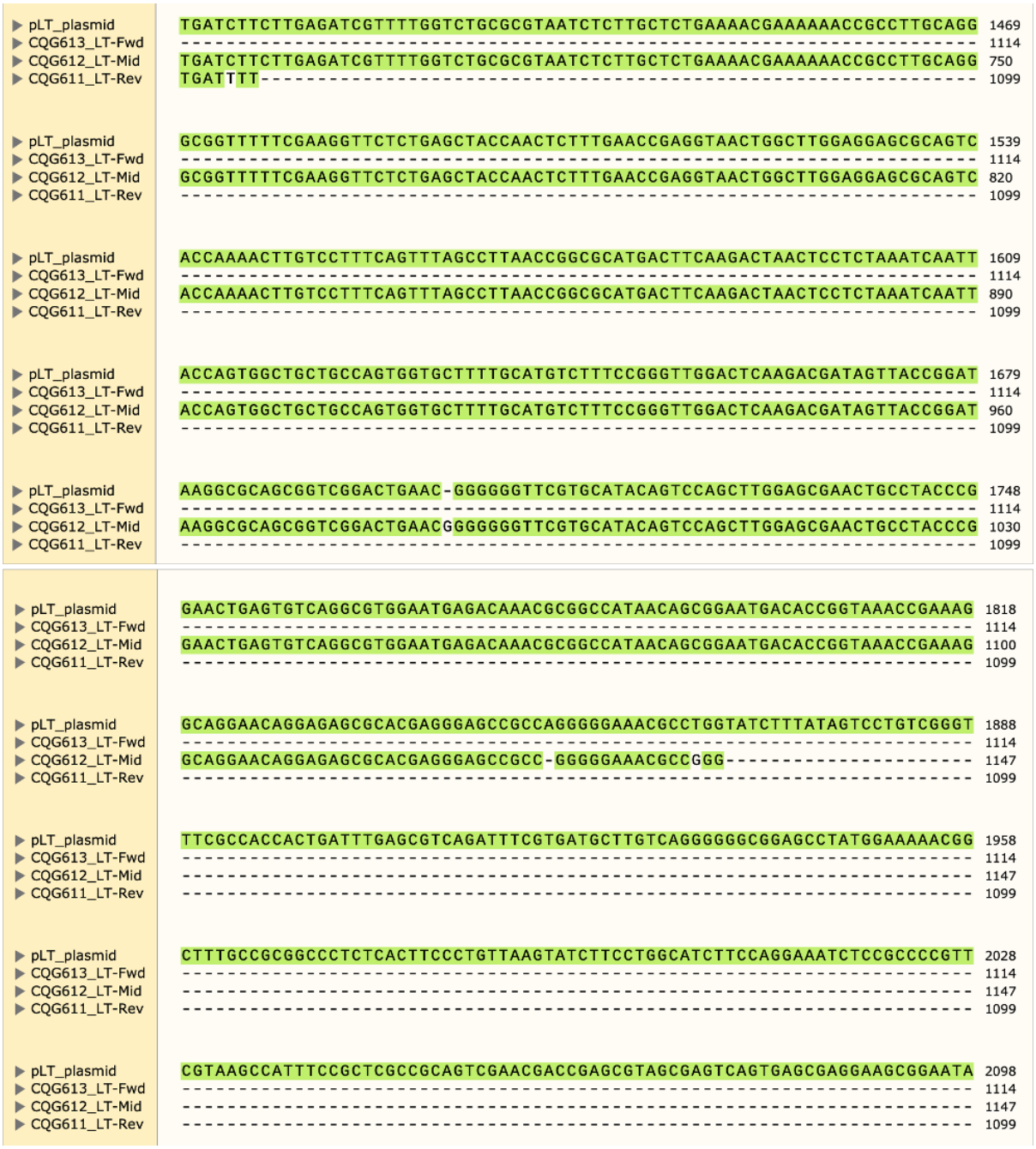
Seq 1 Jobling	GATCTTGGAGCATTCCCACA
Seq 2 CTAend	TTATAGCCACTGCACCCAACATG
Seq 3 CT_LT_Rev	CAAGAGATTACGCGCAGACC
Seq 4 LT_Mid	CTTGGAGAGAAGAACCCTGG
Seq 5 CT_LT_Fwd	TGCCGCGACTCTCTATAATTTC

Section E: DNA sequencing alignments

Alignment of sequencing results with the pLT gene in the plasmid used for production of the protein.


Start pLT gene
pLT_plasmid
CQG613_LT-Fwd
CQG612_LT-Mid
CQG611_LT-Rev
TTTTTGCCGCGACTCTCTATAATTTCTCCATACCTGTTTTTCTGGATGGAGTAAGACCAT-GGCGAAATA 69
GGTTTTCTGGATGGAGTAAGACCATGGGCGAAATA 35
ATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGCTTTAAGAAAATTATCAAGGCATTTGTTATCAT 139
ATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGCTTTAAGAAAATTATCAAGGCATTTGTTATCAT 105
GGCTGCTTTGGTATCTGTTCAGGCGCATGCAAATGGCGACCGTTTATACCGTGCTGACTCTAGACCCCCA 209
GGCTGCTTTGGTATCTGTTCAGGCGCATGCAAATGGCGACCGTTTATACCGTGCTGACTCTAGACCCCCA 175
GATGAAATAAAACGTTCCGGAGGTCTTATGCCCAGAGGGCATAATGAGTACTTCGATAGAGGAACTCAAA 279
GATGAAATAAAACGTTCCGGAGGTCTTATGCCCAGAGGGCATAATGAGTACTTCGATAGAGGAACTCAAA 245
TGAATATTAATCTTTATGATCACGCGAGAGGAACACAAACCGGCTTTGTCAGATATGATGACGGATATGT 349
TGAATATTAATCTTTATGATCACGCGAGAGGAACACAAACCGGCTTTGTCAGATATGATGACGGATATGT 315
GAGAGGAACACAAACCGGCTTTGTCAGATA-GATGACGGATATGT 44
TTCCACTTCTCTTAGTTTGAGAAGTGCTCACTTAGCAGGACAGTCTATATTATCAGGATATTCCACTTAC 419
TTCCACTTCTCTTAGTTTGAGAAGTGCTCACTTAGCAGGACAGTCTATATTATCAGGATATTCCACTTAC 385
TTCCA-TTCTCTTAGTTTGAGAAGTGCTCACTTAGCAGGACAGTCTATATTATCAGGATATTCCACTTAC 113
TATATATATGTTATAGCGACAGCACCAAATATGTTTAATGTTAATGATGTATTAGGCGTATACAGCCCTC 489
TATATATATGTTATAGCGACAGCACCAAATATGTTTAATGTTAATGATGTATTAGGCGTATACAGCCCTC 455
TATATATATG-TATAGCGACAGCACCAAATATGTTTAATGTTAATGATGTATTAGGCGTATACAGCCCTC 182
ACCCATATGAACAGGAGGTTTCTGCGTTAGGTGGAATACCATATTCTCAGATATATGGATGGTATCGTGT 559
ACCCATATGAACAGGAGGTTTCTGCGTTAGGTGGAATACCATATTCTCAGATATATGGATGGTATCGTGT 525
ACCCATATGAACAGGAGGTTTCTGCGTTAGGTGGAATACCATATTCTCAGATATATGGATGGTATCGTGT 252
TAATTTTGGTGTGATTGATGAGCGATTACATCGTAACAGGGAATATAGAGACCGGTATTACAGAAATCTG 629
TAATTTTGGTGTGATTGATGAGCGATTACATCGTAACAGGGAATATAGAGACCGGTATTACAGAAATCTG 595
TAATTTTGGTGTGATTGATGAGCGATTACATCGTAACAGGGAATATAGAGACCGGTATTACAGAAATCTG 322
AATATAGCTCCGGCAGAGGATGGTTACAGATTAGCAGGTTTCCCACCGGATCACCAAGCTTGGAGAGAAG 699
AATATAGCTCCGGCAGAGGATGGTTACAGATTAGCAGGTTTCCCACCGGATCACCAAGCTTGGAGAGAAG 665
CCG 3
AATATAGCTCCGGCAGAGGATGGTTACAGATTAGCAGGTTTCCCACCGGATCACCAAGCTTGGAGAGAAG 392


pLT_plasmid AACCCTGGATTCATCATGCACCACAAGGTTGTGGAAATTCATCAAGAACAATTACAGGTGATACTTGTAA 769
CQG613_LT-Fwd AACCCTGGATTCATCATGCACCACAAGGTTGTGGAAATTCATCAAGAACAATTACAGGTGATACTTGTAA 735
CQG612_LT-Mid ----------ACCAGGTTGTGG-AATTCATCAAGAACAATTACAGGTGATACTTGTAA 50
CQG611_LT-Rev AACCCTGGATTCATCATGCACCACAAGGTTGTGGAAATTCATCAAGAACAATTACAGGTGATACTTGTAA 462
pLT_plasmid TGAGGAGACCCAGAATCTGAGCACAATATATCTCAGGGAATATCAATCAAAAGTTAAGAGGCAGATATTT 839
CQG613_LT-Fwd TGAGGAGACCCAGAATCTGAGCACAATATATCTCAGGGAATATCAATCAAAAGTTAAGAGGCAGATATTT 805
CQG612_LT-Mid TGAGGAGACCCAGAATCTGAGCACAATATATCTCAGGGAATATCAATCAAAAGTTAAGAGGCAGATATTT 120
CQG611_LT-Rev TGAGGAGACCCAGAATCTGAGCACAATATATCTCAGGGAATATCAATCAAAAGTTAAGAGGCAGATATTT 532
pLT_plasmid TCAGACTATCAGTCAGAGGTTGACATATATAACAGAATTCGGGATGAATTATGATTTAAGCTTCGAAATA 909
CQG613_LT-Fwd TCAGACTATCAGTCAGAGGTTGACATATATAACAGAATTCGGGATGAATTATGATTTAAGCTTCGAAATA 875
CQG612_LT-Mid TCAGACTATCAGTCAGAGGTTGACATATATAACAGAATTCGGGATGAATTATGATTTAAGCTTCGAAATA 190
CQG611_LT-Rev TCAGACTATCAGTCAGAGGTTGACATATATAACAGAATTCGGGATGAATTATGATTTAAGCTTCGAAATA 602
pLT_plasmid ATTTTGTTTAACTTTAAGAAGGAGATATACAAACATATGAGCTTTAAGAAAATTATCAAGGCATTTGTTA 979
CQG613_LT-Fwd ATTTTGTTTAACTTTAAGAAGGAGATATACAAACATATGAGCTTTAAGAAAATTATCAAGGCATTTGTTA 945
CQG612_LT-Mid ATTTTGTTTAACTTTAAGAAGGAGATATACAAACATATGAGCTTTAAGAAAATTATCAAGGCATTTGTTA 260
CQG611_LT-Rev ATTTTGTTTAACTTTAAGAAGGAGATATACAAACATATGAGCTTTAAGAAAATTATCAAGGCATTTGTTA 672
pLT_plasmid TCATGGCTGCTTTGGTATCTGTTCAGGCACATGCAGCTCCTCAGACTATTACAGAACTATGTTCGGAATA 1049
CQG613_LT-Fwd TCATGGCTGCTTTGGTATCTGTTCAGGCACATGCAGCTCCTCAGACTATTACAGAACTATGTTCGGAATA 1015
CQG612_LT-Mid TCATGGCTGCTTTGGTATCTGTTCAGGCACATGCAGCTCCTCAGACTATTACAGAACTATGTTCGGAATA 330
CQG611_LT-Rev TCATGGCTGCTTTGGTATCTGTTCAGGCACATGCAGCTCCTCAGACTATTACAGAACTATGTTCGGAATA 742
pLT_plasmid TCGCAACACACAAATATATACGATAAATGACAAGATACTATCATATACGGAATCGATGGCAGGCAAAAGA 1119
CQG613_LT-Fwd TCGCAACACACAAATATATACGATAAATGACAAGATACTATCATATACGGAATCGATGGCAGGCAAAAGA 1085
CQG612_LT-Mid TCGCAACACACAAATATATACGATAAATGACAAGATACTATCATATACGGAATCGATGGCAGGCAAAAGA 400
CQG611_LT-Rev TCGCAACACACAAATATATACGATAAATGACAAGATACTATCATATACGGAATCGATGGCAGGCAAAAGA 812
pLT_plasmid GAAATGGTTATCATTACATTTAAGAGCGGCGAAACATTTCAGGTCGAAGTCCCGGGCAGTCAACATATAG 1189
CQG613_LT-Fwd GAAATGGTTATCATTACATTTAAGAGCGG---------- 1114
CQG612_LT-Mid GAAATGGTTATCATTACATTTAAGAGCGGCGAAACATTTCAGGTCGAAGTCCCGGGCAGTCAACATATAG 470
CQG611_LT-Rev GAAATGGTTATCATTACATTTAAGAGCGGCGAAACATTTCAGGTCGAAGTCCCGGGCAGTCAACATATAG 882
pLT_plasmid ACTCCCAAAAAAAAGCCATTGAAAGGATGAAGGACACATTAAGAATCACATATCTGACCGAGACCAAAAT 1259
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid ACTCCCAAAAAAAAGCCATTGAAAGGATGAAGGACACATTAAGAATCACATATCTGACCGAGACCAAAAT 540
CQG611_LT-Rev ACTCCCAAAAAAAAGCCATTGAAAGGATGAAGGACACATTAAGAATCACATATCTGACCGAGACCAAAAT 952
pLT_plasmid TGATAAATTATGTGTATGGAATAATAAAACCCCCAATTCAATTGCGGCAATCAGTATGAAAAACTAGGAT 1329
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid TGATAAATTATGTGTATGGAATAATAAAACCCCCAATTCAATTGCGGCAATCAGTATGAAAAACTAGGAT 610
CQG611_LT-Rev TGATAAATTATGTGTATGGAATAATAAAACCCCCAATTCAATTGCGGCAATCAGTATGAAAAACTAGGAT 1022
pLT_plasmid ATAAAAAAGCCCACCTCAGAGGGCTTTTTTGTGGATCCATGGCTAGAAATATTTTATCTGATTAATAAGA 1399
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid ATAAAAAAGCCCACCTCAGAGGGCTTTTTTGTGGATCCATGGCTAGAAATATTTTATCTGATTAATAAGA 680
CQG611_LT-Rev ATAAAAAAGCCCACCTCAGAGGGCTTTTTTGTGGATCCATGGCTAGAAATATTTTATCTGATTAATAAGA 1092


pLT_plasmid TGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCGCCTTGCAGG 1469
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid TGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCGCCTTGCAGG 750
CQG611_LT-Rev TGATTTT---------- 1099
pLT_plasmid GCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAGGTAACTGGCTTGGAGGAGCGCAGTC 1539
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid GCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAGGTAACTGGCTTGGAGGAGCGCAGTC 820
CQG611_LT-Rev ---------- 1099
pLT_plasmid ACCAAAACTTGTCCTTTCAGTTTAGCCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATT 1609
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid ACCAAAACTTGTCCTTTCAGTTTAGCCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATT 890
CQG611_LT-Rev ---------- 1099
pLT_plasmid ACCAGTGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTTGGACTCAAGACGATAGTTACCGGAT 1679
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid ACCAGTGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTTGGACTCAAGACGATAGTTACCGGAT 960
CQG611_LT-Rev ---------- 1099
pLT_plasmid AAGGCGCAGCGGTCGGACTGAAC-GGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTGCCTACCCG 1748
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid AAGGCGCAGCGGTCGGACTGAACGGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTGCCTACCCG 1030
CQG611_LT-Rev ---------- 1099
pLT_plasmid GAACTGAGTGTCAGGCGTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGAAAG 1818
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid GAACTGAGTGTCAGGCGTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGAAAG 1100
CQG611_LT-Rev ---------- 1099
pLT_plasmid GCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGT 1888
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid GCAGGAACAGGAGAGCGCACGAGGGAGCCGCC-GGGGGAAACGCCGGG---------- 1147
CQG611_LT-Rev ---------- 1099
pLT_plasmid TTCGCCACCACTGATTTGAGCGTCAGATTTCGTGATGCTTGTCAGGGGGGCGGAGCCTATGGAAAAACGG 1958
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid ---------- 1147
CQG611_LT-Rev ---------- 1099
pLT_plasmid CTTTGCCGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCCTGGCATCTTCCAGGAAATCTCCGCCCCGTT 2028
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid ---------- 1147
CQG611_LT-Rev ---------- 1099
pLT_plasmid CGTAAGCCATTTCCGCTCGCCGCAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATA 2098
CQG613_LT-Fwd ---------- 1114
CQG612_LT-Mid ---------- 1147
CQG611_LT-Rev ---------- 1099

Section F: Protein sequences and biophysical characteristics (Molecular weight, theoretical pI and extinction coefficient)

Amino acid sequence of cCT (without signal sequences)

>CTA
NDDKLYRADSRPPDEIKQSGGLMPRGQSEYFDRGTQMNINLYDHARGTQTGFVRHDDGYVSTSISLRS
AHLVGQTILSGHSTYYIYVIATAPNMFNVNDVLGAYSPHPDEQEVSALGGIPYSQIYGWYRVHFGVLDEQ
LHRNRGYRDRYYSNLDIAPAADGYGLAGFPPEHRAWREEPWIHHAPPGCGNAPRSSMSNTCDEKTQSL
GVKFLDEYQSKVKRQIFSGYQSDIDTHNRIKDEL

Proteolytic cleavage into CTA1 and CTA2 is highlighted in yellow.

Molecular mass: 21,194 Da
Theoretical pI: 5.98

>CTB
TPQNITDLCAEYHNTQIHTLNDKIFSYTESLAGKREMAIITFKNGATFQVEVPGSQHIDSQKKAIERMKDT
LRIAYLTEAKVEKLCVWNNKTPHAIAAISMAN

Molecular mass: 11,607 Da
Theoretical pI: 7.75

CT holotoxin
Molecular mass: 85,140 Da
Theoretical pI: 7.24
Extinction coefficient (ε): 92,430 (assuming all pairs of Cys residues form cysteins)

Amino acid sequence of pLT (without signal sequence)

> pLTA
NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYVSTSLSLRS
AHLAGQSILSGYSTYYIYVIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERL
HRNREYRDRYYRNLNIAPAEDGYRLAGFPPDHQAWREEPWIHHAPQGCGNSSRTITGDTCNEETQNLS
TIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL

>pLTB
APQTITELCSEYRNTQIYTINDKILSYTESMAGKREMVIITFKSGETFQVEVPGSQHIDSQKKAIERMKDTL
RITYLTETKIDKLCVWNNKTPNSIAAISMKN

pLT holotoxin
Molecular mass: 86,690

Theoretical pI: 8.44
Extinction coefficient (ε):107,330 (assuming all pairs of Cys residues form cysteins)

Amino acid sequence of PDI (without signal sequence)

>N-term His-tag-PDI
HHHHHHDAPEEEDHVLVLRKSNFAEALAAHKYLLVEFYAPWCGHCKALAPEYAKAAGKLKAEGSEIRLAKVDATEESDLAQQY
GVRGYPTIKFFRNGDTASPKEYTAGREADDIVNWLKKRTGPAATTLPDGAAAESLVESSEVAVIGFFKDVESDSAKQFLQAAEAI
DDIPFGITSNSDVFSKYQLDKDGVVLFKKFDEGRNNFEGEVTKENLLDFIKHNQLPLVIEFTEQTAPKIFGGEIKTHILLFLPKSVSD
YDGKLSNFKTAAESFKGKILFIFIDSDHTDNQRILEFFGLKKEECPAVRLITLEEEMTKYKPESEELTAERITEFCHRFLEGKIKPHLM
SQELPEDWDKQPVKVLVGKNFEDVAFDEKKNVFVEFYAPWCGHCKQLAPIWDKLGETYKDHENIVIAKMDSTANEVEAVKV
HSFPTLKFFPASADRTVIDYNGERTLDGFKKFLESGGQDGAGDDDDLEDLEEAEEPDMEEDDDQKAVKDEL

Molecular mass: 56,166 Da
Theoretical pI: 4.82
Extinction coefficient Oxidized PDI (ε): 45,755 (assuming all pairs of Cys residues form cysteines)
Extinction coefficient Reduced PDI (ε): 45,380 (assuming all Cys residues are reduced)

Section G: CT crystallization screen

Table S38. CT Crystallization screen. Wells highlighted in blue corresponded to conditions where crystals were formed.

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.1M Na3 Cit + 24%w/v PEG 3350 +300 mM Galctose in Buffer G	0.1M Mg Acet + 18%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Na3 Cit + 24%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.2M Mg Acet + 26%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Mg Acet + 16%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Na3 Cit + 20%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.175M Mg Acet + 22%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Mg Acet + 18%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.1M Na3 Cit + 18%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Na3 Cit + 18%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Na3 Cit + 22%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.175M Mg Acet + 20%w/v PEG 3350 + 300 mM Galctose in Buffer G
B	0.1M Mg Acet + 26%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.2M Mg Acet + 20%w/v PEG 3350 +300 mM Galctose in Buffer G	0.15M Mg Acet + 18%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Mg Acet + 24%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Na3 Cit + 24%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Mg Acet + 26%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.2M Mg Acet + 22%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.1M Mg Acet + 22%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.175M Mg Acet + 26%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Mg Acet + 22%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Na3 Cit + 26%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.1M Na3 Cit + 26%w/v PEG 3350 + 300 mM Galctose in Buffer G
C	0.1M Mg Acet + 24%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Mg Acet + 16%w/v PEG 3350 +300 mM Galctose in Buffer G	0.2M Mg Acet + 24%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Na3 Cit + 18%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.175M Mg Acet + 24%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Mg Acet + 22%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Na3 Cit + 22%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.1M Na3 Cit + 20%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Na3 Cit + 20%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.175M Mg Acet + 18%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.2M Mg Acet + 16%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Na3 Cit + 26%w/v PEG 3350 + 300 mM Galctose in Buffer G
D	0.125M Mg Acet + 20%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.1M Na3 Cit + 22%w/v PEG 3350 +300 mM Galctose in Buffer G	0.15M Mg Acet + 20%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.1M Mg Acet + 20%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Mg Acet + 26%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.2M Mg Acet + 18%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Mg Acet + 24%w/v PEG 3350 +300 mM Galctose in Buffer G	0.175M Mg Acet + 16%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.1M Mg Acet + 16%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.1M Na3 Cit + 16%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.15M Na3 Cit + 16%w/v PEG 3350 + 300 mM Galctose in Buffer G	0.125M Na3 Cit + 16%w/v PEG 3350 + 300 mM Galctose in Buffer G

Section H: Additional protein purification results

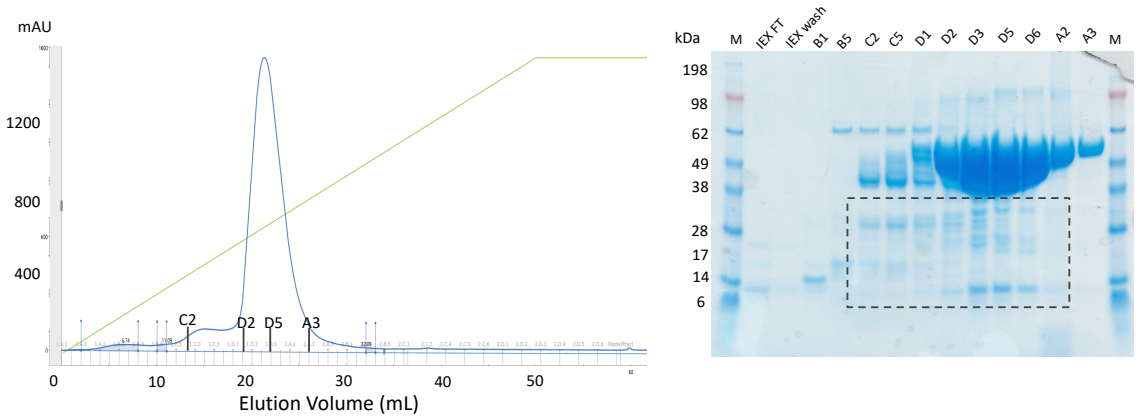


Figure S1. Two overlapping peaks are obtained during IEX elution of PDI with a linear gradient of NaCl. Although the gel was heavily overloaded, several contaminants co-purified with PDI (SDS-PAGE, right panel).

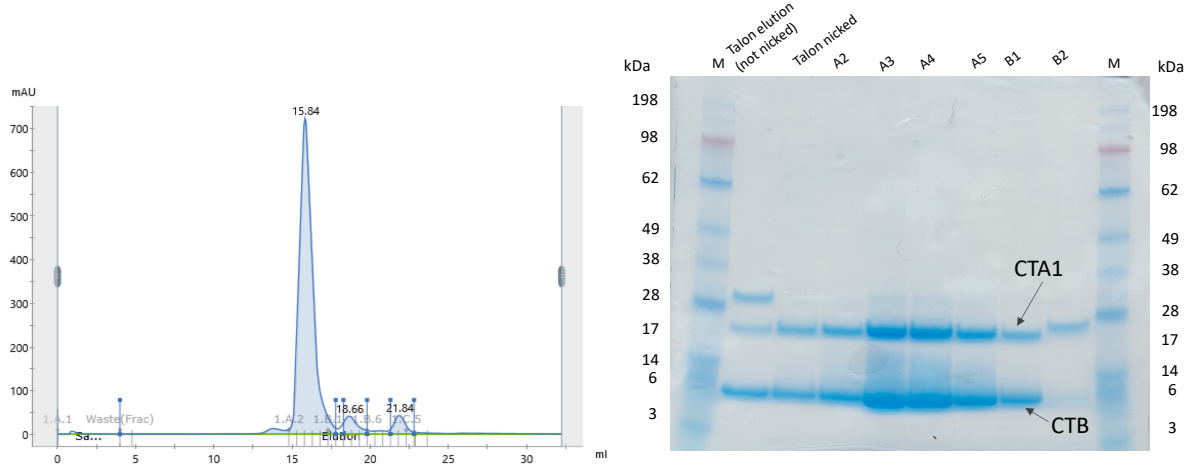


Figure S2. CT trypsin cleavage. Purification of cleaved toxin by SEC (left panel) and SDS-PAGE of collected fractions (right panel). Two clear bands consistent with CTA1 and CTB molecular weights are obtained, indicating correct cleavage.

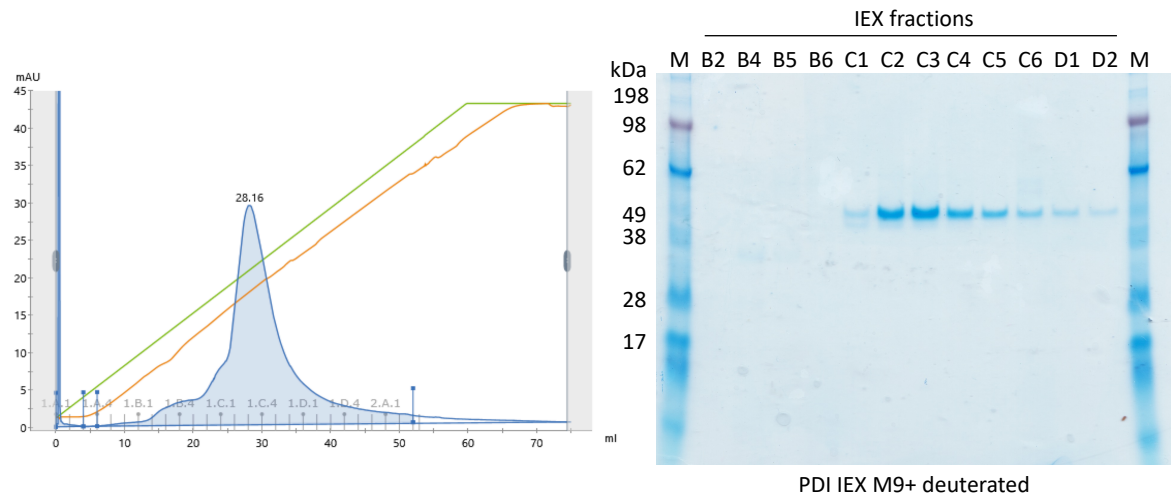
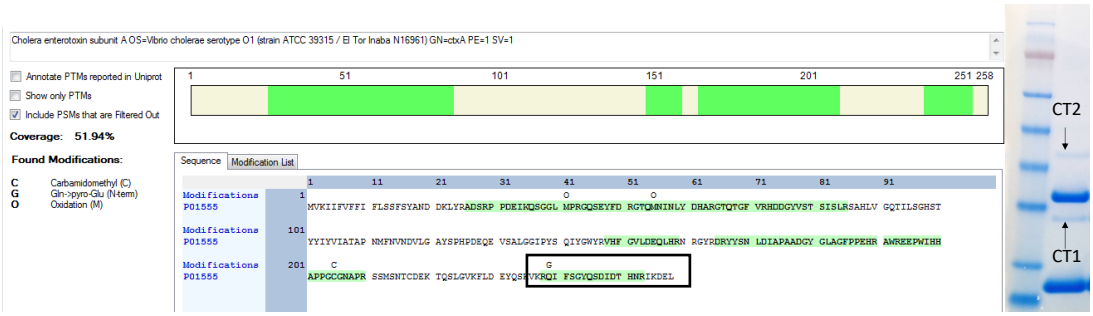


Figure S3. IEX purification of deuterated PDI. Chromatogram (left panel) and SDS-PAGE with fractions (right panel).

Section I: Mass Spectrometry analysis

CT1



CT2

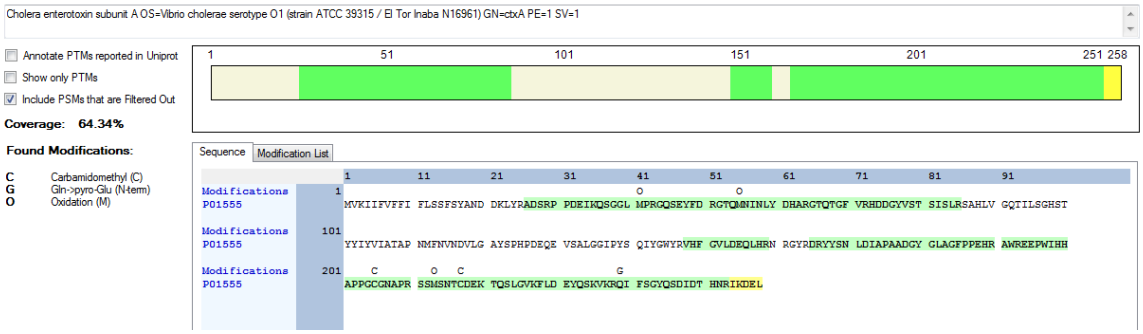


Figure S4. MS analysis of the two additional bands observed in CT samples after SEC purification. The presence of a matching peptide corresponding to residues of the CTA2 peptide (indicated by the square), would suggest that the band labeled 'CT1' is not the cleaved CTA1. The CT2 sample was also identified as CTA.

pLT

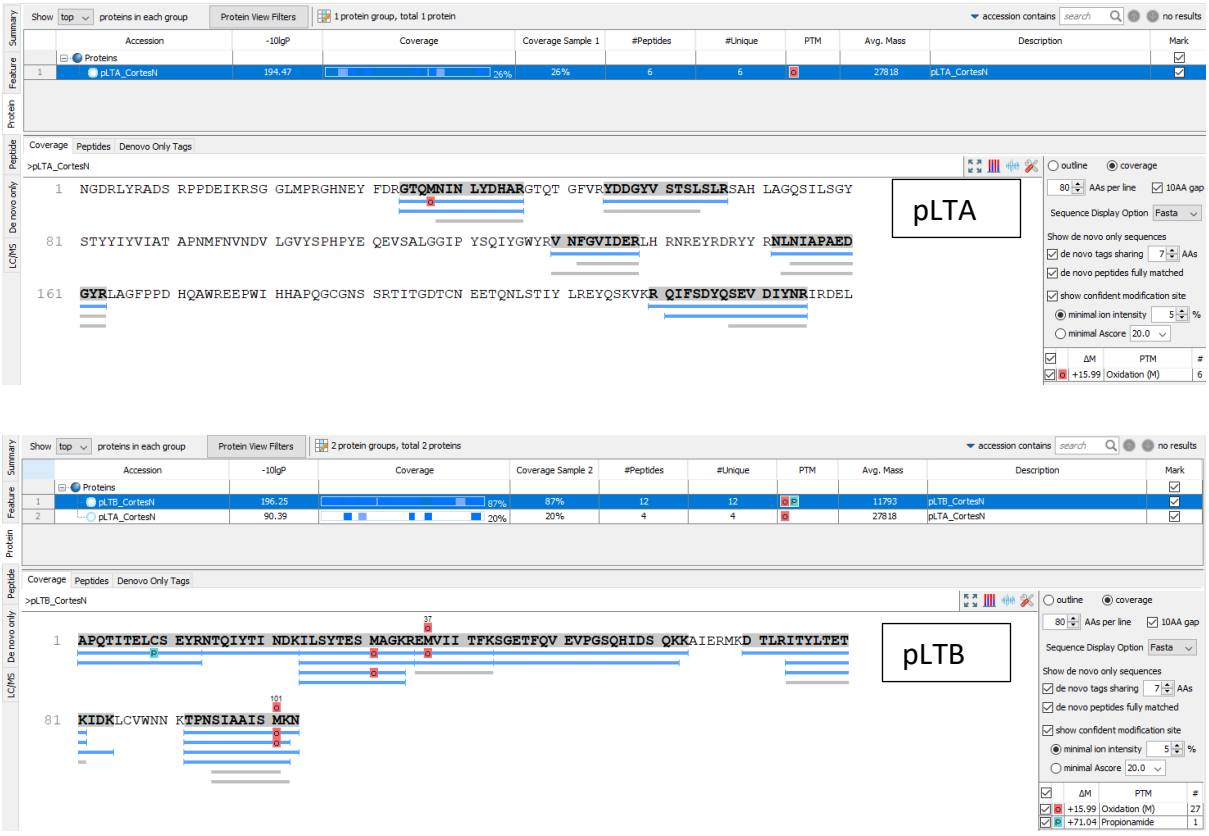
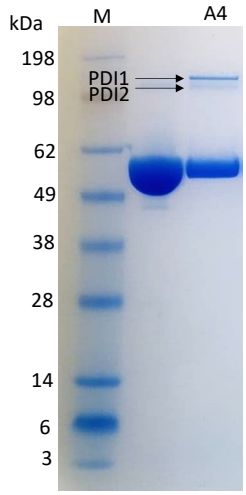


Figure S5. MS analysis confirms production of pLT in Vmax



PDI1 Sample

Accession	Description	Coverage	# Peptides	# PSMs	# Unique Pept	# Protein Grou	# AAs	MW [kDa]
P07237	Protein disulfide-isomerase OS=Homo sapiens GN=P4HB PE=1 SV=3	91,7322835	78	942	78	1	508	57,081
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	34,3167702	20	38	19	1	644	65,999

PDI2 Sample

Accession	Description	Coverage	# Peptides	# PSMs	# Unique Pept	# Protein Grou	# AAs	MW [kDa]
P07237	Protein disulfide-isomerase OS=Homo sapiens GN=P4HB PE=1 SV=3	89,5669291	51	302	51	1	508	57,081
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	47,2049689	32	114	29	1	644	65,999

Figure S6. MS analysis of gel bands obtained during PDI purification. Both bands correspond to PDI, indicating that the protein forms covalently linked oligomers.

Section J: Synchrotron SAXS samples

Table S39. Summary of samples measured at different conditions with SAXS

Sample	Temperature	Buffer	Reducing agent	Additional treatment	Protein Concentration (mg/mL)
Nicked CT	10 ℃	20 mM Bis-Tris pH 6.5, 200 mM NaCl	20 mM TCEP	-	0.5 2
PDI					0.5 1 2
Nicked CT+PDI (1:1)				Minimal Incubation time (< 3 min)	CT= 2 ; PDI=1.32
Nicked CT+PDI (1:2)					CT= 1.57; PDI=2
Un-nicked CT			None	-	0.5 2
PDI	37 ℃	PBS pH 7.4	None	-	0.33 0.66 1.32
EDC-treated PDI					0.33 0.66 1.32
Un-nicked CT					0.5 1 2
Q3R-treated PDI		PBS pH 7.4, 2% DMSO			0.33 0.66
Un-nicked CT					0.5 1
PDI	37 ℃	PBS pH 7.4	2 mM GSH	-	0.33 0.66 1.32
Un-nicked CT					0.33 0.66 1.32
Un-nicked CT+PDI (1:1)				Incubation time > 30 min	0.5 1 2
Q3R-treated PDI		PBS pH 7.4, 2% DMSO		-	0.33 0.66
Un-nicked CT					0.5 1
Un-nicked CT+Q3R-PDI (1:1)				Incubation time > 30 min	CT= 1 ; PDI=0.66 CT= 0.5 ; PDI=0.33
PDI	37 ℃	PBS pH 7.4	20 mM GSH	-	0.33 0.66 1.32
Un-nicked CT					0.33 0.66 1.32
EDC-treated PDI					0.5 1 2
Un-nicked pLT					0.4
Un-nicked CT+PDI (1:1)				Incubation time > 30 min	CT= 2 ; PDI=1.32 CT= 1 ; PDI=0.66 CT= 0.5 ; PDI=0.33
Un-nicked CT+EDC-PDI (1:1)					CT= 2 ; PDI=1.32 CT= 1 ; PDI=0.66 CT= 0.5 ; PDI=0.33
Un-nicked pLT+PDI (1:1)					pLT 0.4; PDI 0.26
Un-nicked pLT+EDC-PDI (1:1)					pLT 0.4; PDI 0.26

Section K: Additional SAXS results

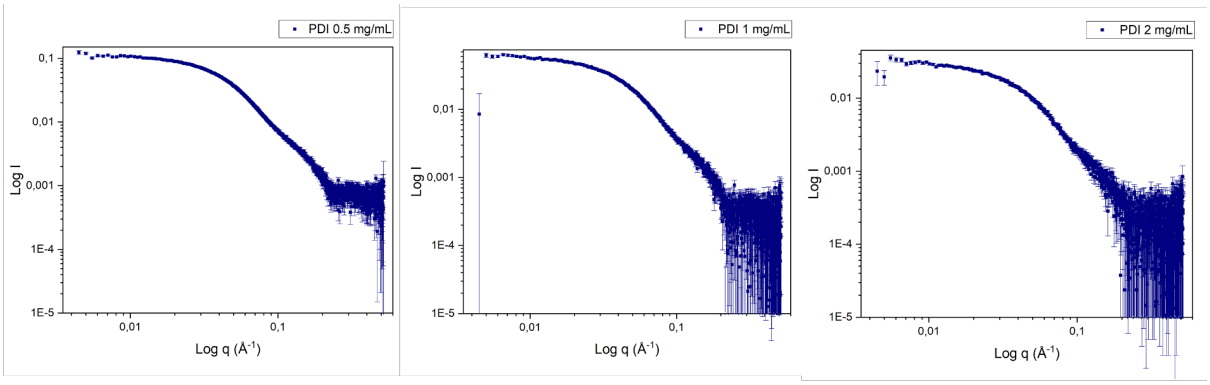


Figure S7. SAXS curves of PDI at 10 ℃ in a Bis-Tris buffer pH 6.5 containing 20 mM TCEP.

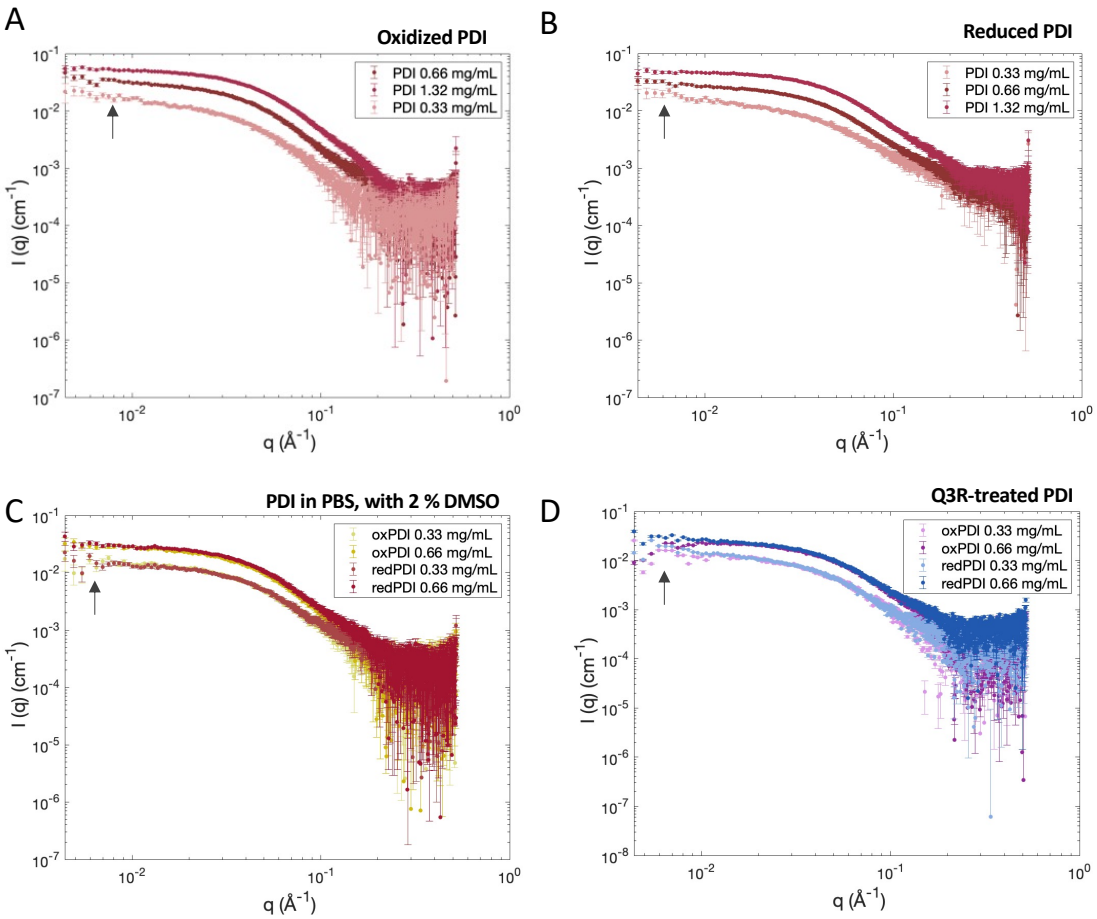


Figure S8. SAXS analysis of PDI samples at 37 ℃, in PBS buffer containing 2 mM GSH. An upturn in the low-q region of the plots is more prominent at lower protein concentrations.

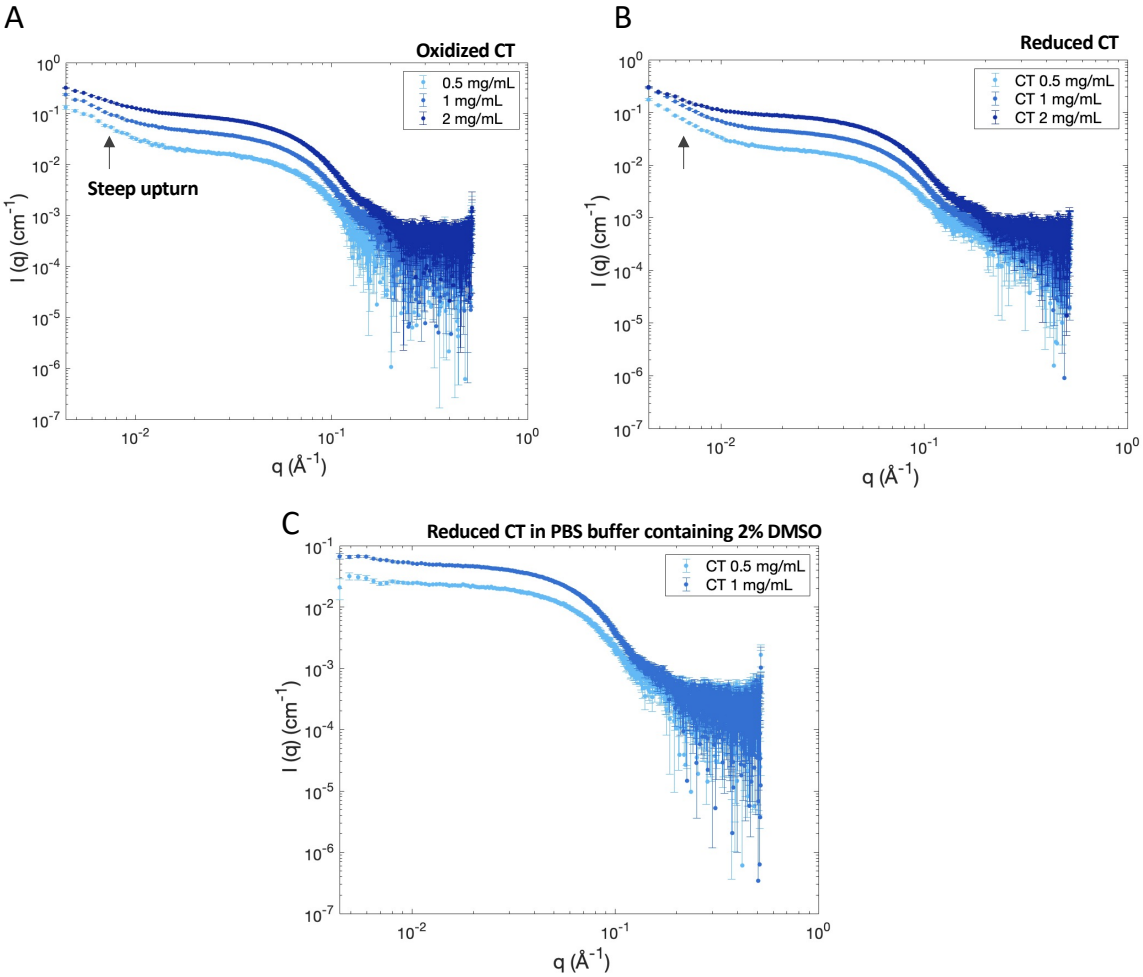


Figure S9. SAXS analysis of un-nicked CT samples at 37 ℃, in PBS buffer containing 2 mM GSH. An upturn in the low-q region of the plots is more prominent at lower protein concentrations.

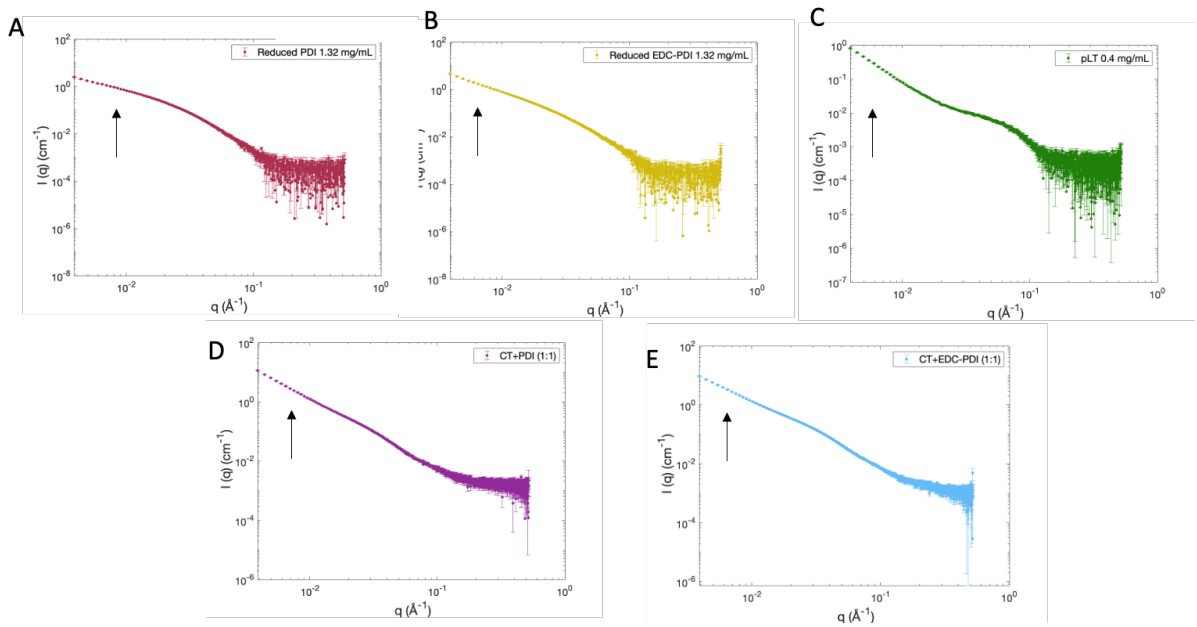


Figure S10. SAXS analysis of different samples at 37 ℃, in PBS buffer containing 20 mM GSH. A steep upturn in the low-q region is observed in all the plots, making the data impossible to interpret.

Section L: CT-PDI Docking with HDOCK

SAXS data: The file was uploaded by users.

```
RG by the SAXS data:    29.654 A
MW by the SAXS data:    41.073 kDa

WARNING: The SAXS data may not be compatible with the structures!!

WARNING: The MW determined by the SAXS data is too small!!

WARNING: MWsaxs is too smaller than MWcomplex!
Disorder/flexibility check.............OK!
```

Figure S11. Error obtained when generating CT-PDI docking models with HDOCK. The R_g and molecular weight calculated from the SAXS data is not compatible with a complex the size of CT-PDI.