Structural basis of ubiquitin recognition and rational design of novel covalent inhibitors targeting Cdu1 from *Chlamydia Trachomatis*

Strukturelle Grundlage der Ubiquitin-Erkennung und rationales Design neuer kovalenter Inhibitoren gegen die Deubiquitinylase Cdu1 aus *Chlamydia trachomatis*

Doctoral thesis

for a doctoral degree at the Graduate School of Life Sciences
Julius-Maximilians-Universität Würzburg
Section Biomedicine

---

submitted by

Yesid A. Ramirez
Born in Pasto, Colombia

Würzburg 2019

Submitted on: ....

Members of the thesis committee

Chairperson: Prof. Dr. Ulrike Holzgrabe
Primary Supervisor: Prof. Dr. Caroline Kisker
Supervisor (Second):   Prof. Dr. Christoph Sotriffer
Supervisor (Third):   Prof. Dr. Thomas Rudel

Date of Public Defense: ....

Date of Receipt of Certificates: ....

Para mi tío Héctor

Mi trabajo siempre ha de ser para honrar su memoria, querido padrino

SUMMARY

The WHO-designated neglected-disease pathogen *Chlamydia trachomatis* (CT) is a gram-negative bacterium responsible for the most frequently diagnosed sexually transmitted infection worldwide. CT infections can lead to infertility, blindness and reactive arthritis, among others. CT acts as an infectious agent by its ability to evade the immune response of its host, which includes the impairment of the NF-κB mediated inflammatory response and the Mcl1 pro-apoptotic pathway through its deubiquitylating, deneddylating and transacetylating enzyme ChlaDUB1 (Cdu1). Expression of Cdu1 is also connected to host cell Golgi apparatus fragmentation, a key process in CT infections.

Cdu1 may this be an attractive drug target for the treatment of CT infections. However, a lead molecule for the development of novel potent inhibitors has been unknown so far. Sequence alignments and phylogenetic searches allocate Cdu1 in the CE clan of cysteine proteases. The adenovirus protease (adenain) also belongs to this clan and shares a high degree of structural similarity with Cdu1. Taking advantage of topological similarities between the active sites of Cdu1 and adenain, a target-hopping approach on a focused set of adenain inhibitors, developed at Novartis, has been pursued. The thereby identified cyano-pyrimidines represent the first active-site directed covalent reversible inhibitors for Cdu1. High-resolution crystal structures of Cdu1 in complex with the covalently bound cyano-pyrimidines as well as with its substrate ubiquitin have been elucidated. The structural data of this thesis, combined with enzymatic assays and covalent docking studies, provide valuable insights into Cdu1s activity, substrate recognition, active site pocket flexibility and potential hotspots for ligand interaction. Structure-informed drug design permitted the optimization of this cyano-pyrimidine based scaffold towards HJR108, the first molecule of its kind specifically designed to disrupt the function of Cdu1. The structures of potentially more potent and selective Cdu1 inhibitors are herein proposed.

This thesis provides important insights towards our understanding of the structural basis of ubiquitin recognition by Cdu1, and the basis to design highly specific Cdu1 covalent inhibitors.

ZUSAMMENFASSUNG

Der Krankheitserreger Chlamydia trachomatis (CT) - ein gramnegatives Bakterium - ist verantwortlich für die häufigste sexuell übertragene Infektionskrankheit weltweit, die CT basierte Chlamydiose. Sie wird von der Weltgesundheitsorganisation zu den vernachlässigten Krankheiten gezählt.

CT Infektionen können unter anderem zu Unfruchtbarkeit, Erblindung und reaktiver Arthritis führen. CT agiert als Krankheitserreger mittels seiner Fähigkeit, die Immunantwort des Wirts zu umgehen. Dies umfasst unter anderem die Schwächung und Störung der NF-κB vermittelten Entzündungsantwort und des Mcl1 pro-Apoptoseweges über ihr deubiquitinierendes, deneddylierendes und trans-acetylierendes Enzym ChlaDub1 (Cdu1). Die Expression von Cdu1 ist aber auch mit der Fragmentierung des Golgi-Apparates des Wirtes verknüpft, ein Schlüsselprozess bei Infektionen mit CT.

Cdu1 ist daher vermutlich ein attraktives Zielprotein für die Entwicklung von Wirkstoffen, um CT Infektionen zu behandeln. Eine Leitstrukturverbindung zur Entwicklung neuer wirksamer Inhibitoren war bislang jedoch noch nicht bekannt. Sequenzvergleiche und phylogenetische Untersuchungen verorten Cdu1 im CE Clan der Cysteinproteasen. Die Adenovirus-Protease (Adenain) gehört ebenfalls diesem Clan an und besitzt strukturelle Ähnlichkeit mit Cdu1. Unter Ausnutzung der topologischen Ähnlichkeiten der aktiven Zentren von Cdu1 und Adenain wurde ein Target-Hopping Ansatz mit einem klar definierten und fokussierten Satz von bei Novartis entwickelten Adenain-Inhibitoren verfolgt.

Die hierbei identifizierten Cyano-Pyrimidine stellen die ersten kovalenten Inhibitoren von Cdu1 dar, die an das aktive Zentrum von Cdu1 binden und es direkt adressieren. Hochauflösend wurden Kristallstrukturen sowohl von Komplexen von Cdu1 mit kovalent gebundenen Cyano-Pyrimidinen als auch mit Cdu1's natürlichem Substrat Ubiquitin bestimmt. Die Kristallstrukturdaten dieser Doktorarbeit in Kombination mit Enzymassays und kovalenten Docking-Studien liefern wertvolle Hinweise bezüglich der Aktivität des Enzyms, der molekularen Substraterkennung, der Flexibiliät der Proteintasche rund um das aktive Zentrum und potentielle Hotspots für die Wechselwirkung mit Liganden. Ein strukturbasiertes

Wirkstoffdesign erlaubte die Optimierung des Cyano-Pyrimidin-basierten Molekülgerüstes, die zu der Entwicklung der HJR108 Verbindung führte. Es ist das erste Molekül seiner Art, das speziell dazu entworfen wurde Cdu1 zu inhibieren. Strukturen potentiell noch wirksamerer und selektiver Cdu1 Inhibitoren werden in dieser Arbeit vorgeschlagen.

Diese Dissertationsschrift liefert somit wertvolle Beiträge zum Verständnis der strukturellen Grundlagen der molekularen Erkennung von Ubiquitin durch Cdu1 und Hinweise, die die Entwicklung hoch-spezifischer kovalenter Cdu1 Inhibitoren erlauben sollten.

CONTENTS

1 INTRODUCTION

1.1 The *Chlamydiales* order

1.1.1 Taxonomy

The members of the *Chlamydiales* order are obligate-intracellular and gram-negative bacteria. Based on 16S and 23S ribosomal RNA sequence analysis, the *Chlamydiales* order can be divided into four families comprising *Chlamydiaceae, Parachlamydiaceae, Waddliaceae* and *Simkaniaceae.* The *Chlamydiaceae* family consists of two genera, *Chlamydia* and *Chlamydophila* (1)*. Chlamydia trachomatis* (2) and *Chlamydia pneumoniae* (3) are two *Chlamydi*a species which are representatives of each genus and are considered as relevant human pathogens (Figure 1.1).
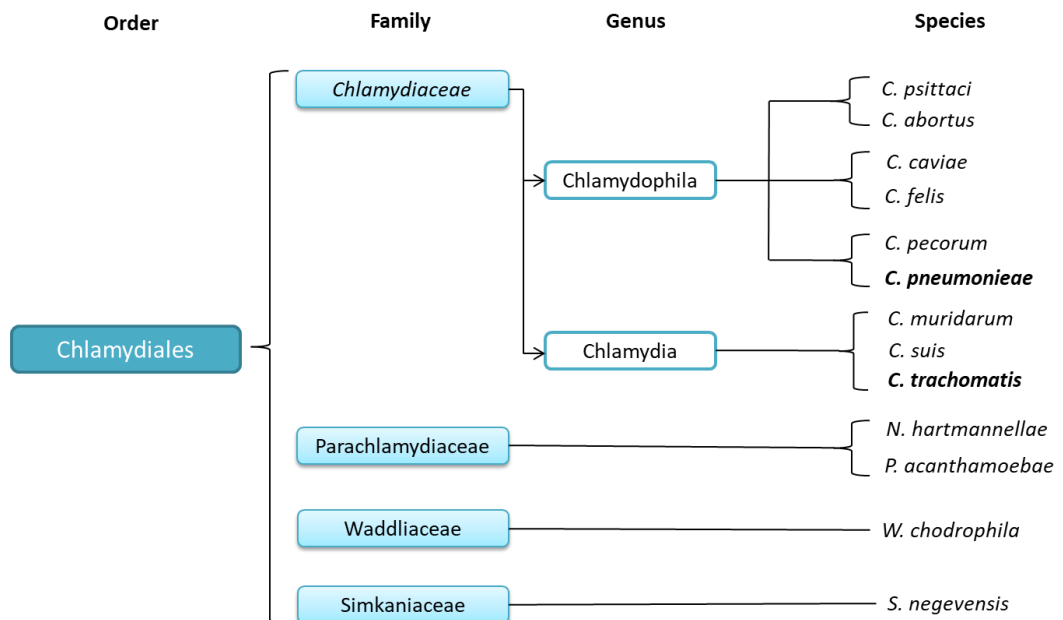
Figure 1.1 Taxonomic classification of the *Chlamydiaes* order according to 16S and 23S rRNA sequence analysis. (data adapted from Everett and Bush, 1999)

1.1.2 Developmental cycle

Chlamydia are characterized by a two-stage developmental cycle. Infective elementary bodies (EBs) attach to the surface of the host cells on the epithelium of mucous membranes (conjunctiva, urethra, cervical canal, endometrium and uterine tubes) and enter the cells via endocytosis (Figure 1.2). Within 6 to 8 hours of entering the cells, EBs differentiate into metabolically active reticulate bodies (RBs) inside a pathogen-containing vacuole known as *Chlamydia* inclusion. RBs continue to grow for 8 to 16 hours. From 16 to 48 hours post infection (hpi), RBs begin to differentiate into second generation EBs, which can lead to complete cell lysis up to 72 hours after infection and thereby to several hundreds of new infective units in the extracellular space. In some cases, only a part of the inclusion is budded off the cell while the original cell does not die. This is known as extrusion. Under certain stress conditions such as antibiotics and starvation, RBs can turn into non infective aberrant forms, which are viable, metabolically active and resistant to many antibacterial drugs, and may reverse into infective forms even months after infection (4). It has been shown that aberrant forms residing on mobile leukocytes may be carried to joint tissues to induce inflammation and interestingly chlamydial DNA is prevalent in up to 80 percent of patients with reactive arthritis. (5).

1.1.3 Medical relevance

Causing up to 10% of community-acquired pneumonia cases, *Chlamydia pneumoniae* (CP) is a major infectious agent of the human respiratory track (6) and can even lead to chronic diseases such as lung cancer (7) and asthma (8). It is estimated that 90% of CP infections are asymptomatic (9), and that some non-respiratory pathologies such as Alzheimer's disease (10) and atherosclerosis (11) are associated with CP infections.
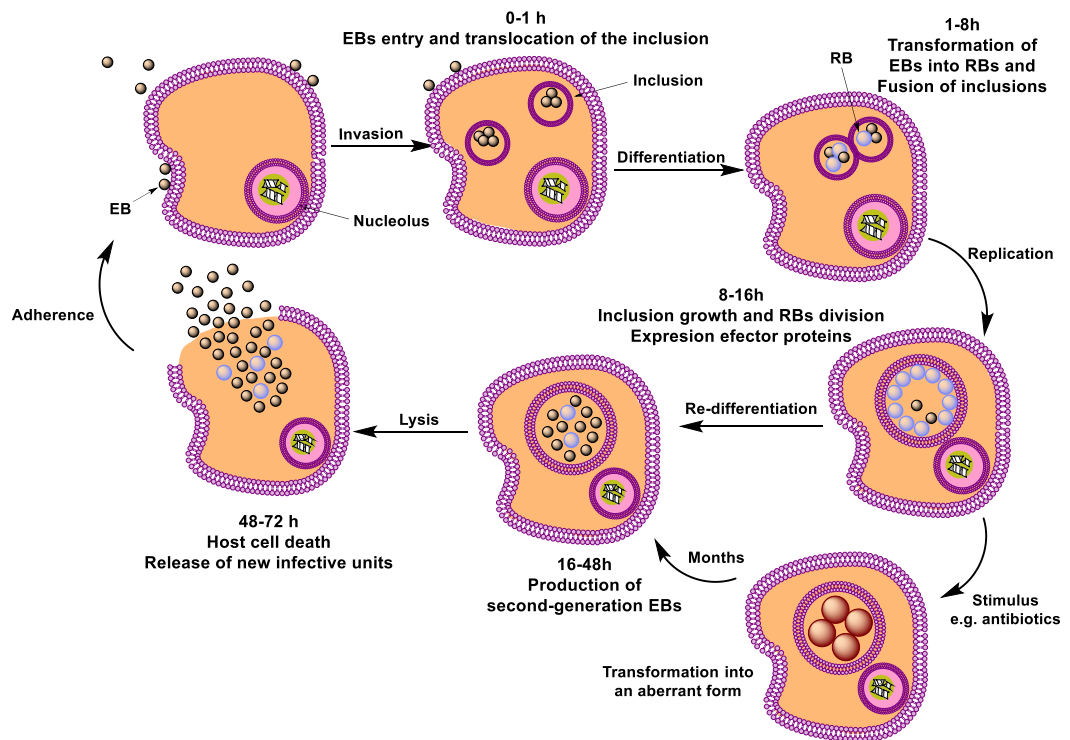
Figure 1.2 Life cycle of *Chlamydiae*. EBs refer to elementary body and RBs to reticular body (adapted from Karyagina *et al*. 2009)

Chlamydia trachomatis (CT) is responsible for major eye and genital diseases in humans. With more than 130 million new cases per year, it is the most commonly diagnosed sexually transmitted infection in the world (12). In women, untreated infections can cause severe complications such as Pelvic Inflammatory Disease (PID), leading to sequelae including ectopic pregnancy, chronic pelvic pain and infertility, the later affecting men as well (13). Sexually transmitted chlamydiosis can be a facilitating cofactor for HIV infections (14) and is also associated with several adverse outcomes during pregnancy, threatening the life of both mother and new born. CT may also be transmitted to the infant during labor, putting the new born at risk of developing conjunctivitis and nasopharyngeal infections which may cause life-long sequalae in the respiratory tracts (15). In the United States alone, the treatment cost of sexually transmitted chlamydiosis was estimated to have accumulated to $516.7 million dollars in 2008 (16). If the pathogen reaches the eye, the WHO-designated neglected disease

trachoma may be developed. Currently about 2.2 million people are affected by trachoma, of whom 1.2 million are irreversibly blind. It is estimated that 232 million people are at risk of blindness worldwide due to hyperendemic trachoma, mainly in territories belonging to developing countries in sub-Saharan Africa, the Middle East, South America and Asia (17).

1.2   Anti-chlamydial antibiotics

1.2.1   Frontline drugs against *Chlamydia* infections

Antibacterial drugs act by blocking essential biological processes in pathogens responsible for human infections. The antibacterial spectrum mostly includes inhibiting ribosomal protein biosynthesis, DNA replication and transcription, nucleotide biosynthesis, cell wall synthesis and disrupting the cell membrane integrity (Table 1.1) (18).

Targeted cellular process	Antibiotic class
cell wall biosynthesis	β-lactams, fosfomycin, glycopeptides
ribosomal protein biosynthesis	**macrolides**, **tetracyclines**, aminoglycosides, chloramphenicol, streptogramins, lincosamides, oxazolidinones
DNA biosynthesis	sulfonamides
DNA replication and transcription	**fluoroquinolones**, rifamycins
cell membrane integrity	lipopeptides, polymyxins

Table 1.1 Main antibiotic classes and their molecular targets. WHO-designated frontline drugs against CT infections are highlighted in bold (adapted from Kapoor *et al*. 2017)

For uncomplicated genital chlamydia infections, the World Health organization (WHO) recommends the use of either 1 g azithromycin orally as a single dose, or 100 mg doxycycline 100 mg orally twice a day for seven days. The form of treatment should be decided by the physicians depending on individual factors. For instance, if high value is to be put in reducing costs then doxycycline might be the best choice. If high value is placed on control and convenience, azithromycin as a single dose might preferentially be formulated, as compliance

for seven days is not guaranteed. Other alternatives include tetracycline (500 mg orally four times a day for 7 days), erythromycin (500 mg orally four times a day for 7 days) and ofloxacin (200-400 mg orally twice a day for 7 days). It should be noted that the use of doxycycline, tetracycline and ofloxacin is contraindicated in pregnant women (19).

Azithromycin (AZM) and erythromycin belong to the macrolide class of antibiotics. They target the conserved sequences of the peptidyl transferase center of the 23S rRNA of the 50S ribosomal subunit to produce a premature detachment of incomplete peptide chains. Doxycycline and tetracycline belong to the tetracycline class and act upon the conserved sequences of the 16S r-RNA of the 30S ribosomal subunit to prevent the binding of t-RNAs. The use of macrolides or tetracyclines (TETs) leads to protein synthesis inhibition, an essential process, for the pathogen's cells proliferation. Ofloxacin is a fluoroquinolone which inhibits DNA replication by blocking the activity of the bacterial DNA gyrase. This results in excessive positive supercoiling of the DNA strands and impairs subsequent transcription. (20, 21).

1.2.2 Antibiotic resistance in *Chlamydia*

As a Gram-negative bacterium, CT possess an additional membrane layer outside the peptidoglycan layer, which makes it intrinsically resistant against many antibacterial drugs (22). Additionally, it has been proposed that genes encoding for efflux pumps, ribosomal protections proteins or antibiotic-inactivating enzymes (23) may be acquired by accumulation of point mutations or obtained from other pathogens by horizontal gene transfer. (24).

Although CT infections can be treated with antichlamydial antibiotics such as TETs and AZM (Section 1.2.1) there are indications that the bacterium can readily develop resistance to frontline drugs when exposed to sub-inhibitory

antimicrobial concentrations (24). In the mid-1990s, a tetracycline resistant *Chlamydia suis* (CS) strain was isolated from diseased pigs (25). Genetic analysis allowed the identification of genes encoding for a TET efflux pump (*tet*[C]), a regulatory repressor (*tetR*) and a unique insertion (IScs605) apparently involved in plasmid replication and mobilization (26). Remarkably, the *tet*[C] and IScs605 sequences are at least 99% identical to TET resistance genes from the aquatic Gram-negative pathogens *Aeromonas salmonicida* (26) and *Laribacter hongkongensis* (27), respectively. This leads to the question whether uncontrolled practices in the pig meat industry which rely on the prophylactic use of TETs and use fish as a feed source, may have facilitated the incorporation of resistance genes into the CS genome. Interestingly, this also represents the first identification of antibiotic resistance acquired through horizontal transfer in any obligate-intracellular pathogen. Alarmingly, the ability of other members of the *Chlamydia* genus to acquire TET resistance from CS has already been demonstrated *in vitro* with the successful generation of a TET resistant CT strain (28). Nevertheless, stable antibiotic resistant CT is still to be documented in clinical settings (24).

 The recent spread of a *Neisseria gonorrhoeae* strain resistant to sulfonamides, penicillins, early-generation cephalosporins, tetracyclines, macrolides, fluoroquinolones (29) and even last-line extended-spectrum cephalosporins (ESCs), calls for the development of new effective drug regulations, prescription policies and treatment options (30). Considering that 15-60% of individuals with genital gonorrhea are also infected with CT (31, 32), and given the ability of the *Chlamydia* genus to acquire resistance by horizontal gene transfer, it may be only a matter of time before CT develops front-line drug resistance.

1.3 Exploring the ubiquitin system for drug development

1.3.1 The ubiquitin system as therapeutic target

Ubiquitylation is a post-translational modification where one or several lysine residues on a target protein are covalently modified with a small protein known as ubiquitin. This modification is made possible by the intimate collaboration between E1 activating, E2 conjugating and E3 ligating enzymes. Ubiquitylation may be propagated by transferring additional ubiquitin molecules to one of seven lysine residues that ubiquitin itself possesses, or to its N-terminal amino group. Therefore, eight homogeneous, or multiple heterotypic ubiquitin conjugates are possible. (33). The topology of the ubiquitin chains acts as a molecular marker that leads to different biological outcomes (34). For instance, K11 and K48 chains are related to cell division and proteasomal degradation (35, 36), whereas K63 and linear ubiquitin chains play prominent roles in protein-complex assembly, DNA-repair and inflammation (37-39). Branched chain ubiquitin chains are also important. For instance, K11/K48 mixed conjugates are known to be involved in mitotic regulation and quality control of nascent proteins (40).

As ubiquitylation is a central process in the cell that regulates key cellular functions such as protein degradation, cell cycle progression, transcriptional regulation, receptor internalization and signal transduction, it is not surprising that serious diseases may develop if a component of the system is mutated or deleted (41). For instance, mutations in the E3 ligase PARKIN are known to be a cause of Parkinson's disease (34). Accordingly, many components of the ubiquitin system are attractive therapeutic targets (42).

Proteasomal degradation may also be targeted for therapy. The FDA approved the proteasome inhibitor, bortezomib, which is a peptide boronate that forms tetrahedral adducts with the catalytically active N-terminal threonine (Thr1) of the

proteasome's β5 subunits. This may result in a dramatic disease regression in multiple myeloma patients, possibly by stabilization of IκBα, the suppressor of NF-κB signaling (43, 44). Carfilzomib, a tetrapeptide epoxyketone, is the second FDA approved proteasome inhibitor which surpasses bortezomib in potency by simultaneously binding to the hydroxyl group and the α-amino group of Thr1, forming a morpholino adduct (45, 46). Ixazomib is the first orally-available proteasome inhibitor with FDA approval that, as its predecessors, triggers apoptosis in multiple myeloma cells (44).

The enzymatic machinery of the ubiquitin system (particularly E3 ligases) may be harnessed for the development of therapeutic alternatives other than the classic occupancy-driven pharmacology which relies on inhibitor-enzyme interactions. Protein-targeting chimeric molecules (PROTACs) are bivalent ligands that consists of an E3 ligase-moiety linked to a target-protein recruiting moiety. The hijacked E3 ligase ubiquitinylates the target protein to induce its proteasomal mediated degradation (47). This approach has been successfully used to induce cereblon E3 ligase mediated degradation of BRD4, a protein critical for cancer cell growth and survival to yield delayed leukemia progression in mice (48). While PROTAC development is still at its infancy, induced protein degradation may have several advantages over protein inhibition including the ability to target the undruggable proteome, to overcome the accumulation of the drug target and prevent its evolution, and the possibility to overcome pharmacokinetic problems (plasma half-life, permeability) due to the catalytic nature of PROTACs (49).

1.3.2   Targeting deubiquitylase enzymes

Like other post-translational modifications, ubiquitylation is a reversible process mediated by ∼100 human deubiquitylase enzymes (DUBs) (50). DUBs are proteases that can be classified into seven families according to their sequence

and domain conservation. These include UCHs (carboxyl-terminal hydrolases, USPs (ubiquitin specific proteases), OTUs (ovarian tumor like proteases), MJDs (Machado-Josephin domain-containing proteases), MINDYs (motif-interacting with ubiquitin-containing novel DUB family), JAMMs (JAB1, MPN, MOV34 family) and the recently discovered ZUFSP family (Zinc finger with UFM1-specific peptidase domain protein) (51). With the exception of JAMMs which are zinc-metalloproteases, all DUBs are cysteine proteases (52). As ubiquitylation and related processes control a myriad aspects of human cell biology, deregulation of DUBs is implicated in various health disorders including cancer. Notable examples include:

- The ubiquitin-specific protease 9X (USP9X), which plays a pivotal role both as oncogene and tumor suppressor by promoting the activity of antiapoptotic factors such as Mcl1 (53). Mutations in USP9X are associated with intellectual disability and abnormal neuronal cell migration and growth (54).

- The UCH family member BRCA1-Associated Protein 1 (BAP1), mutations in which have been reported to confer tumor development susceptibility in melanoma and mesothelioma (55).

- The ubiquitin-specific protease 7 (USP7), which induces apoptosis in various multiple myeloma cell lines when inhibited, including those resistant to the proteasome inhibitor bortezomib (56). USP7 is mutated in a human neurodevelopmental disorder (57)

- The ubiquitin-specific protease 8 (USP8), which mutations are known to cause excess cortisol production (Cushing's disease) (58).

- The ubiquitin-specific protease 15 (USP15), which promotes oncogenesis through the activation of TGF-beta signaling in glioblastoma (59).

- The ubiquitin-specific protease 37 (USP37), which mediates the G1/S transition in the cell cycle by stabilizing cyclin A, and prevents the degradation of the c-MYC oncoprotein through its deubiquitylating activity (60, 61).

As USP37, many other DUBs are known to stabilize oncoproteins. Therefore, these DUBs represent an attractive drug target for cancer therapy (42). The clinical success of active-site directed proteasome inhibitors for the treatment of multiple myeloma (section 1.3.1), prompted the exploration of other components of the proteasome as potentially more specific drug targets. As proteasomal DUBs facilitate the degradation of target proteins by removing ubiquitin moieties which would otherwise impede entry to the proteasome's core particle, it has been proposed that inhibiting these DUBs may pose a therapeutic advantage (62). For instance, the proteasome associated DUBs USP14 and UCHL5 are overexpressed in several forms of cancer (63, 64) and may play a substantial role in disease progression. VLX1570, the very first DUB inhibitor in phase I trials (now suspended) for the treatment of multiple myeloma and solid tumors, has been reported to target these DUBs (65).
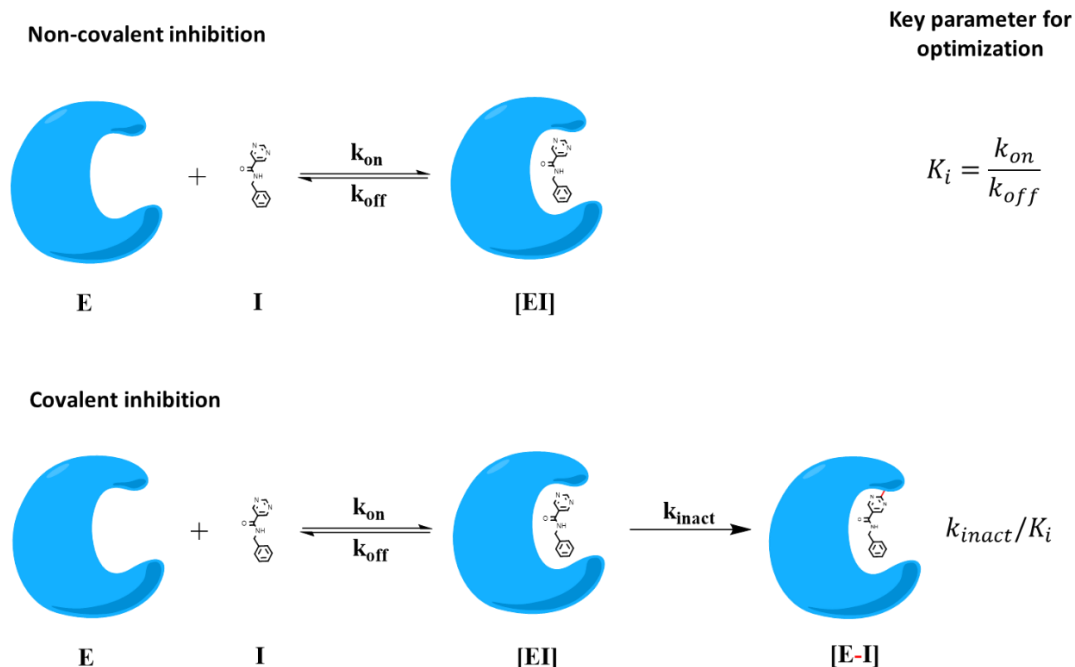
As ubiquitylation is important for the activation of various innate immune signaling pathways in the eukaryotic cell, perhaps is not surprising that many pathogens possess genes encoding for DUBs which are used to bypass the host cell defense mechanisms upon infection. For example, the highly pathogenic human coronavirus (HCoV) encodes for a papain-like protease termed SARS-CoV PLpro. This protease mediates the removal of the Ubl ISG15 from host cell factors, hindering the host cells' antiviral response (66). Human bacterial pathogens including *Salmonella typhimurium*, *Rickettsia bellii*, *Legionella pneumophila,* *Escherichia coli, Shigella flexneri and Chlamydia spp*, rely on secreted effector DUBs to manipulate host signaling pathways in diverse ways, with strong preference for K63 ubiquitin chains (67). Therefore, these proteases may represent interesting drug targets

Parasitic pathogens, apart from possessing native eukaryotic Ubl pathways, may also express DUBs that target host cell functions, similar to viruses and bacteria. For instance, the PfUCH54 protease encoded by *Plasmodium falciparum* has been shown to possess both deubiquitylating and deNeddylating activity and to be

essential for parasite survival (68, 69). A similar protein is expressed by *Toxoplasma gondii*, the dual specificity of which appears to be conserved in evolution from an ancestral eukaryotic homolog, known as UCHL3 (70). However, distinct structural features of these parasitic DUBs compared to their human counterparts, may constitute an opportunity to harness DUBs as therapeutic targets against these pathogens (69).

1.3.3 Covalent inhibition: a rediscovered paradigm?

A common focus of modern small-molecule design is to optimize desired drug-protein interactions which are known produce a therapeutic effect. The non-covalent molecular interactions determine the ratio between drug-protein complex formation (k_{on}) to unbound drug and free protein dissociation (k_{off}) under equilibrium binding conditions. The k_{on} / k_{off} ratio, K_i, measures the ability of a non-covalent inhibitor to interact with its target and serves as an indicator of the drug's potency (71). In covalent inhibition, the ligand undergoes a bond-forming event with the target protein that complements traditional reversible interactions to yield an overall more potent ligand. The covalent bond may be irreversible within the half-life of the target protein, forming a drug-protein complex that does not follow classic equilibrium kinetics (Figure 1.3) (71). In contrast to non-covalent inhibitors, the reaction between a covalent drug and a protein does not proceed to equilibrium, but rather to completion. Therefore, the rate constant for covalent modification, k_{inact}, must be considered and the k_{inact} / K_i ratio may be used as a reliable descriptor to rank the potency of covalent ligands (Figure 1.3).

Non-covalent inhibition

$$K_i = \frac{k_{on}}{k_{off}}$$

Key parameter for optimization

E + I ⇌ [EI]

Covalent inhibition

E + I ⇌ [EI] → [E-I]

$$k_{inact}/K_i$$

Figure 1.3 Comparison of the interactions modes of an inhibitor (I) with an enzyme (E) under non-covalent (top) and covalent (bottom) conditions. E = Enzyme, I = Inhibitor EI = Enzyme-inhibitor complex, E-I = Enzyme-inhibitor covalent complex). Adapted from C. Noe and Gilbert 2012.

Historically, the pharmaceutical industry has excluded reactive chemical moieties that form covalent adducts with their targets from their drug discovery campaigns (e.g. electrophiles), fearing that their inherent reactivity could lead to off-target side effects and toxicity (72). However, this widespread belief contrasts with the overwhelming clinical success of drugs that rely on covalent interactions to address their targets. This includes nonsteroidal anti-inflammatory drugs (NSAIDs) which irreversibly inhibit cyclooxygenase (COX)-1 and 2, such as aspirin (73), and antimicrobial drugs from the β-lactam and Fosfomycin classes, which covalently label essential proteins for bacterial cell wall biosynthesis (74) (Figure 1.4). In fact, as the underlying mechanisms of covalent inhibition become better understood, observations highlighting the potential of this type of inhibitors are becoming more abundant in the literature (75). This has prompted a shift towards the rediscovery of the covalent inhibition paradigm as a valuable tool for modern therapeutics. This

ligand class may have several advantages compared to their non-covalent counterparts:

- The generation of a physical bond between the enzyme and the ligand effectively shifts the equilibrium towards an enzyme-inhibitor covalent complex (E-I), which reduces the dissociation rate (k_{off}, figure 1.3). This leads to an increased period for which the ligand is bound to its target (residence time) and to a longer duration of action. Covalent drugs may thus be administered at lower doses and with a reduced frequency (76, 77).

- Because they operate under non-equilibrium binding kinetics, competition with endogenous ligands is mitigated (78).

- The theoretical limit in the strength of non-covalent interactions, sets a boundary to the affinity of such ligands at around 10 pM. Some covalent ligands may bypass this limit as they bind the targets permanently. Covalent inhibitors are common outliers in the measurement of drug potencies as a function of heavy atom-count, which reflects to a better biochemical efficiency (79, 80)

- Because a covalent drug is tightly bound to its target, its pharmacodynamic effect can persist long after the pharmacokinetic clearance of the drug from plasma has taken place (81).

- Proteins with small or narrow 'undruggable' binding sites may be readily addressed by a covalent warhead (82).

- Covalent ligands which target the active site of a protein might be more resistant to therapy-induced resistance, as mutations in the active site would compromise the function of the target protein (83). The same is true for non-covalent active site inhibitors.
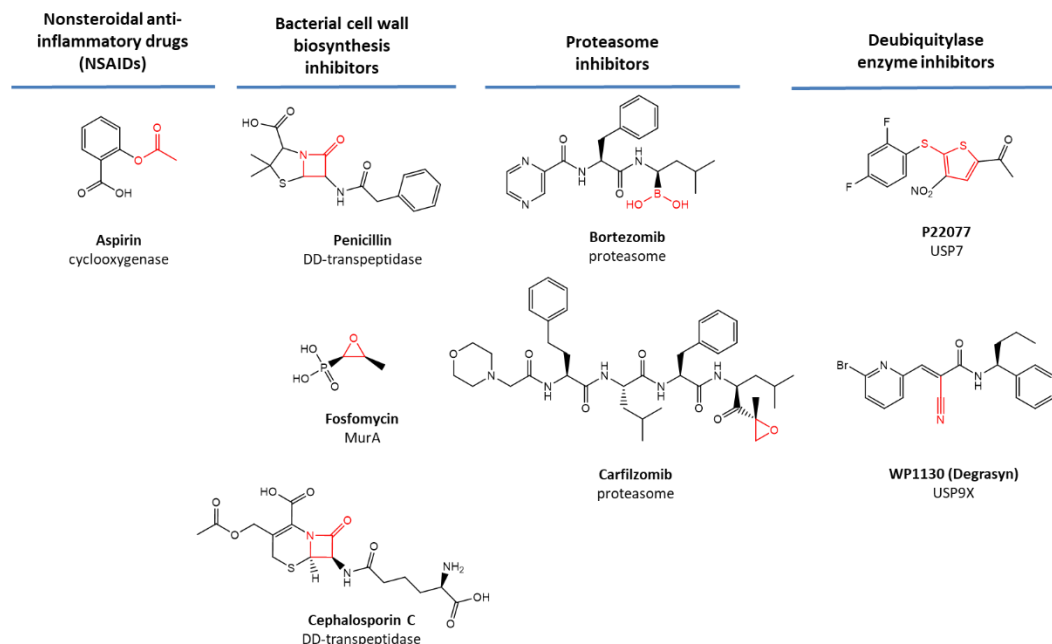
Figure 1.4. Examples of covalent drugs and their cellular targets. Electrophile moieties are highlighted in red. Except for the DUB inhibitors (right column) all ligands are FDA approved. Most examples are taken from Bauer et al, 2015.

Although covalent inhibitors may provide unique benefits in drug development, some liabilities must be considered as well. Drugs with highly reactive war heads such as Michael acceptors and epoxides, may cause rare or unpredictable drug reactions (idiosyncratic drug-related toxicity) by modification of off-targets. This phenomenon could lead to the haptenization of proteins, which may produce an undesired and potentially dangerous host immune response against covalently modified proteins. The use of less-reactive or reversible electrophiles, like acrylamide and nitrile moieties, has been proposed as an alternative to bypass this issue. Reversible covalent inhibitors (i.e. covalent inhibitors that dissociate from their targets after a given amount of time) are thought to have a lower risk of toxicity, since the drug-protein adducts they produce are transitory and may not accumulate to a level that leads to a host immune reaction (75).

The hepatotoxic properties of some drugs that undergo first-pass metabolism to produce highly reactive intermediates that react with circulating nucleophiles and hepatic proteins, must also be considered. For instance, acetaminophen undergoes hepatic metabolism to yield N-acetyl-p-benzoquinone imine (NAPBQI). Given the acute toxicity of NAPBQI, acetaminophen is unfortunately one of the most frequently used drugs in intentional overdoses worldwide (84). Finally, it must be highlighted that the use of covalent inhibitors may be ineffective, or even toxic, in systems that require a short residence time or in protein targets with rapid turnover (75).

Remarkably, molecules targeting the ubiquitin system can be found among clinically successful covalent drugs: the proteasomal inhibitors Bortezomib and Carfilzomib (Section 1.3.1 and Figure 1.4). On the other hand, DUBs contain a well-defined active site mostly harboring a catalytic cysteine residue (65) making them prone to inhibition by covalent modification of the active-site residue. Although the development of DUB inhibitors is still in its infancy, potent covalent active-site inhibitors have been successfully developed for some DUBs such as USP7 and USP9X (85, 86) (Section 1.3.2 and Figure 1.4). Therefore, DUB covalent inhibitors may be expected to enter clinical development in the years to come.

1.4 The *Chlamydia trachomatis* deubiquitylase 1 (Cdu1)

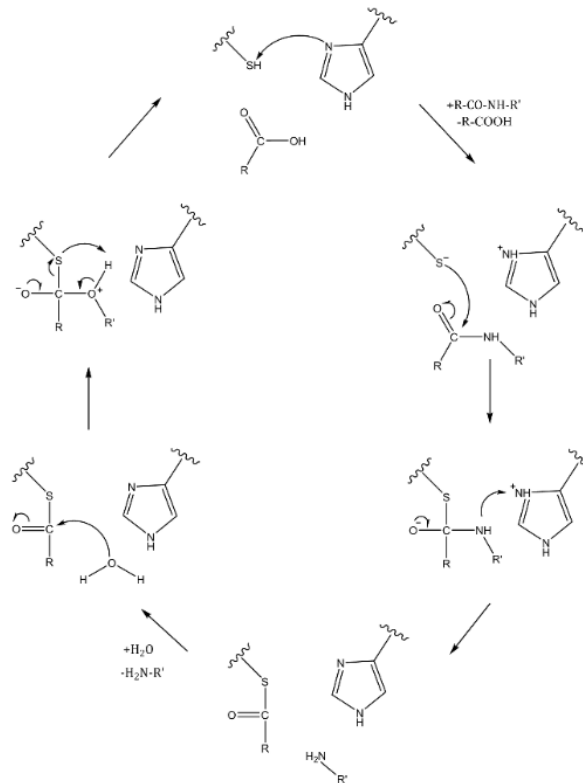### 1.4.1 Cdu1, a cysteine protease

Cysteine proteases are proteins that cleave other proteins, characterized for sharing a common proteolytic mechanism that depends on the nucleophilicity of a sulfur atom at a catalytic cysteine (Figure 1.5). An adjacent basic amino acid, usually a histidine, assists in the deprotonation of the thiol group, constituting a catalytic dyad (87). The catalytic triad can be completed by an aspartic or glutamic acid which stabilizes the resulting positive charge of the protonated histidine and

is located adjacent to the catalytic dyad. The resulting anionic sulfur on the cysteine is then able to perform a nucleophilic attack on the carbonyl group of an isopeptide bond. Subsequently, a cleaved N-terminal fragment of the substrate is released, and the catalytic histidine returns to its deprotonated form. The resulting thioester is then hydrolyzed, releasing a fragment of the substrate with a carboxylic terminus and reconstituting the enzyme (Figure 1.6).
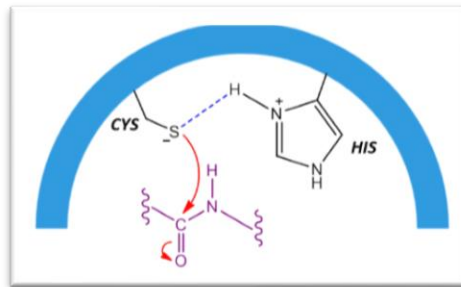
Figure 1.5. The catalytic cysteine in cysteine proteases performs a nucleophilic attack on an isopeptide bond.


Figure 1.6. Catalytic cycle of cysteine proteases, showing the role of the cysteine and histidine in the catalytic center

1.4.2 Cdu1 as a drug target

Sections in between " " are taken from Ramirez et al, 2018 with permission

As intracellular gram-negative bacteria, CT relies on a needle-like apparatus called type 3 secretion system (T3SS) or injectosome (88) to secrete several effector proteins into the inclusion lumen, the inclusion membrane or the host cell cytoplasm (89). Two of these effector proteins are the Chlamydial deubiquitylases 1 and 2 (Cdu1 and Cdu2). They are synthesized inside RBs and are translocated into the inclusion membrane at early stages of an infection, where they can be detected until late developmental stages of the pathogen (90, 91). Cdu2 is concomitantly localized at the inclusion membrane along with Cdu1, but also localizes at the host cell plasma membrane at late stages of infection (Figure 1.7) (91). Cdu1 is a highly versatile protein with deubiquitylating, deneddylating and acetyltransferase activities (92, 93). ''As ubiquitylation and deubiquitylation regulate many key cellular processes (Section 1.3), it has been proposed that these enzymes play a role in the infection mechanism and pathogenicity through the alteration of the ubiquitin-proteasome pathway (92). The Cdu1 cysteine protease permits the bacterium to bypass the human host cell's inflammatory response regulated by NF-κB through deubiquitylation of its inhibitory subunit, IκBα (94)''.

Cdu1 also enables CT to interfere with the programmed cell death response of infected host cells by impairing the proteasome-mediated degradation of the apoptosis inhibitor, Mcl1 (90), and it plays a role in stabilizing the eukaryotic cell glucose transporter 1 (GLU1), a host protein sequestered in CT infections to ensure a steady carbon source for *Chlamydia* development at the inclusion (95). Additionally, both Cdu1 and Cdu2 have been linked to fragmentation of the host Golgi apparatus, a key process in chlamydia infections (93). One may thus hypothesize that inhibiting Cdu1 could pose a disadvantage for the pathogen and may therefore represent an attractive drug target.
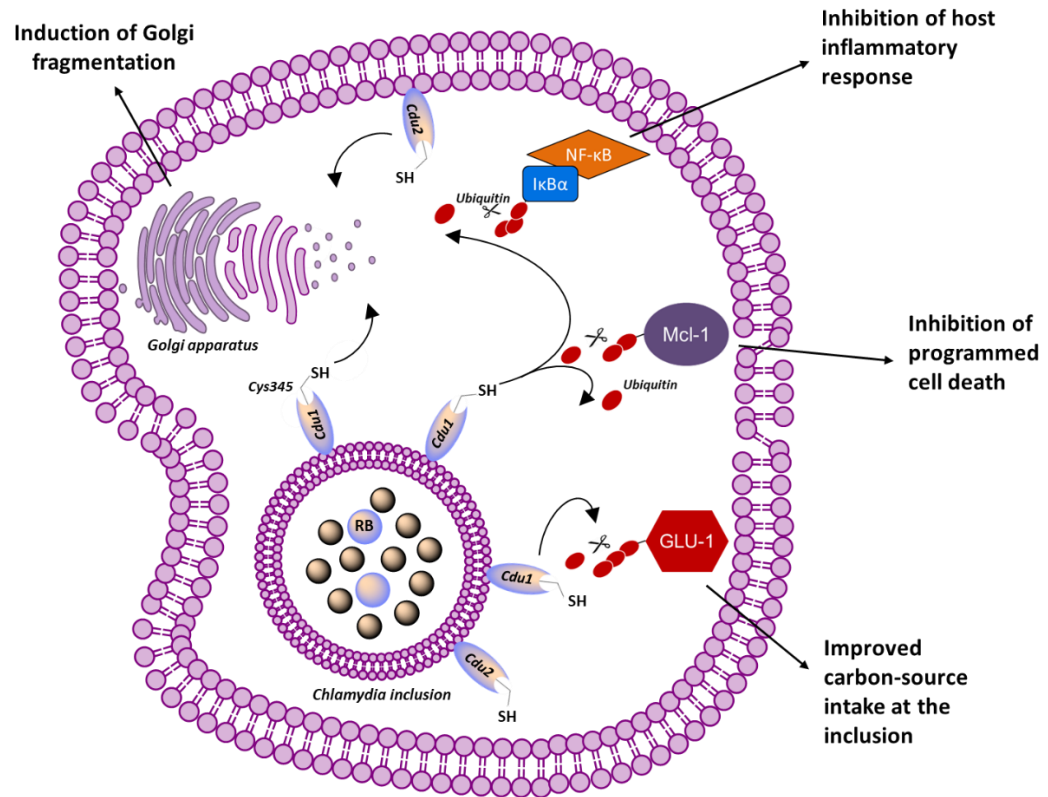
Figure 1.7. Cellular localization of Cdu1 and Cdu2 on infected cells, and their known molecular targets illustrating how these DUBs pose an advantage for CT in terms of infectivity.

1.4.3 The CE clan of cysteine proteases
Sections in between " " are taken from Ramirez et al, 2018 with permission

Cysteine-dependent proteinases are often referred to as 'the cysteine protease family'. However, they have at least ten different evolutionary origins, each of which has produced a group of proteins with unique structures and properties (Figure 1.8). Cysteine proteases are thus classified into clans (proteins which are evolutionarily related) and may be further divided into distinct families according to their differences in the catalytic domains (96). ''Cdu1 belongs to a group of evolutionarily related proteins known as the CE clan of cysteine proteases. The CE clan features DUBs that can hydrolyze conjugates of ubiquitin as well as Ubl-specific proteases (ULPs) which catalyze the cleavage of ubiquitin-

18

like proteins (Ubl) such as Nedd8 and SUMO (97). In eukaryotes, the CE fold is present in peptidases responsible for maintaining free ubiquitin levels, rescuing proteins from ubiquitin-mediated degradation, and controlling the dynamics of ubiquitin-mediated signaling by cleaving Ubls from diverse substrates (98). In contrast, bacteria and viruses encode CE effector proteins to bypass the eukaryotic host cell defense mechanisms (99). These enzymes include not only deSUMOylases (100) and deubiquitinylases (101), but also acetyltransferases (102)''.
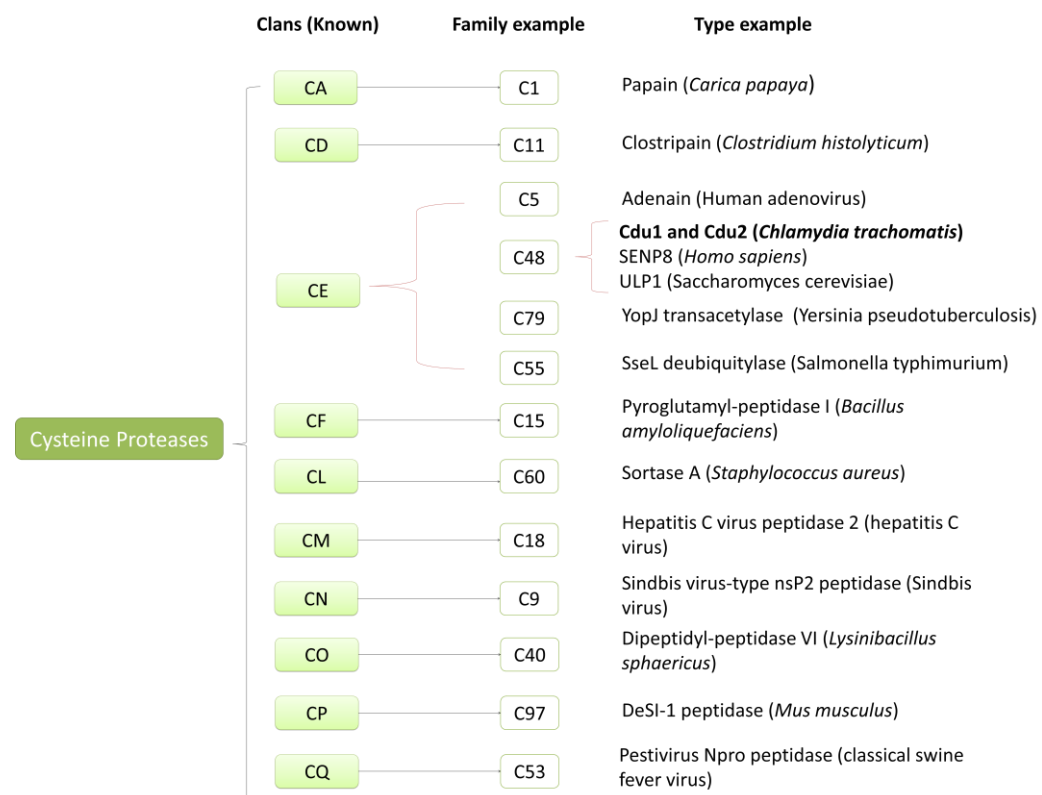
Figure 1.8: Evolutionary line of cysteine proteases. The ten known clans can contain up to 24 families. An example for a family member of every clan is shown on the right. Cdu1 and Cdu2 belong to the C48 family, clan CE. Adapted from Barrett and Rawlings et al. 2001.

Cdu1 can be allocated into the family C48 of the CE clan of the cysteine proteases (97). The peptidases of this family share a common architecture of their catalytically active C-terminal domain, while the N-is weakly conserved (103). C48

19

Ubl-hydrolases of known structure include the Ulp1 protease, from *Saccharomyces cerevisiae* (104), the Sentrin-Specific protease 8 known as SENP8 from *Homo sapiens* (105) and XopD from *Xanthomonas campestris* (Figure 1.9 A). ''The versatility of the CE fold can be attributed to the presence of a constant region (CR) near the conserved α-helix containing the active-site cysteine, which is surrounded by 3 variable regions (VR-1, 2 and 3) (Figure 1.9 B)''.
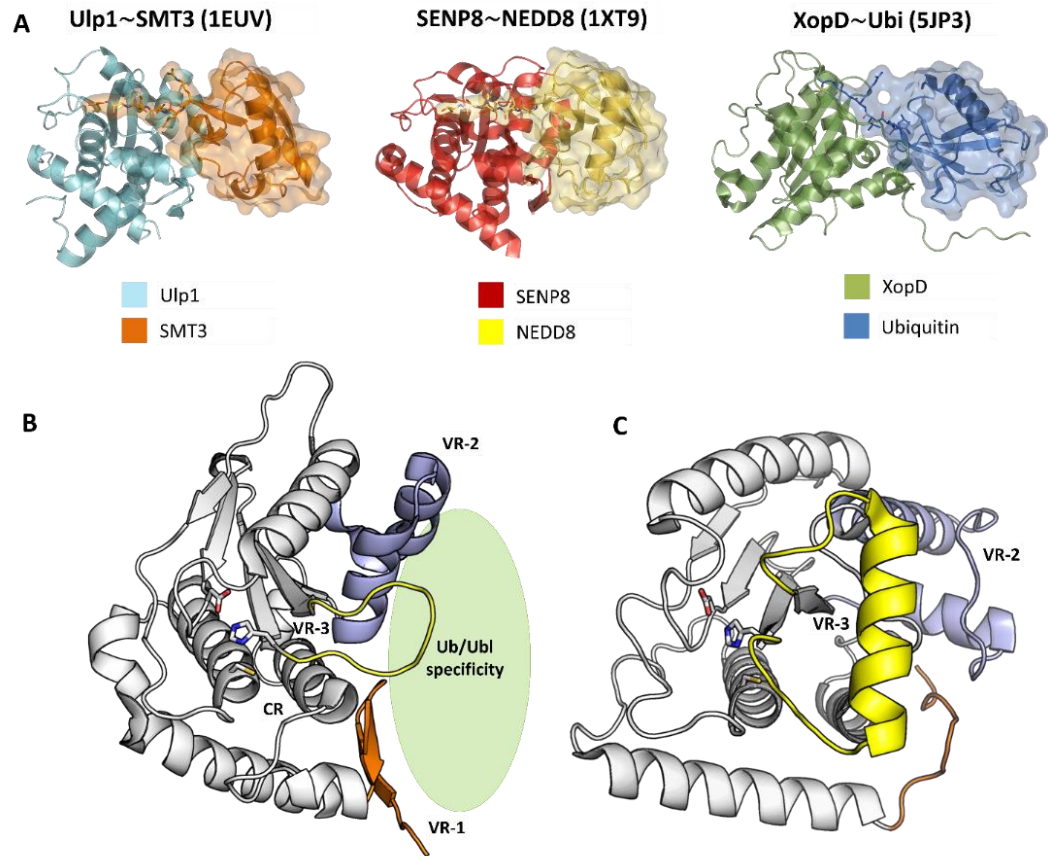

Figure 1.9. (A) Structures of the yeast Ulp1~SMT3 complex (PDB: 1EUV, left), the human SENP8~NEDD8 complex (PDB: 1XT9, center) and the Xanthomonas XopD-Ubi complex (PDB: 5JP3, right) illustrating a common Ubl binding mode among eukaryotic and bacterial C48 proteases. **(B)** Basic architecture of the CE proteases based on the structure of SENP8 (PDB: 2BKR). The conserved CE architecture contrasts with VRs 1, 2 and 3, which are hypothesized to be involved in substrate recognition. **(C)** Crystal structure of Cdu1 (PDB: 5HAG). VR-3 is a unique Chlamydial α-helical insertion, while VR-1 is absent. Partially adapted from Pruneda, et al, 2016.

These variable regions are proposed to be responsible for substrate specificity, with VR-2 and VR-3 appearing to be diverse among CE-fold DUBs. In Cdu1, VR-1 is absent and is instead occupied by the protein´s C-terminus (Figure 1.9 C). It was shown that the removal of VR-3 in Cdu1, which is a unique α-helical feature of chlamydial DUBs, impairs binding to ubiquitin (67).

1.4.4   The moonlighting function of Cdu1

In addition to its deubiquitylating and deneddylating activities (section 1.4.2), the recently described acetyltransferase activity of Cdu1 (AcT), makes it a remarkable example of protein moonlighting (93). The CE clan acetyltransferase YopJ (*Yersinia pseudotuberculosis*) is known to acetylate and inhibit kinase activation by blocking phosphorylation (102). Although cellular acetylation targets of Cdu1 are still to be described, one may speculate that Cdu1 uses acetylation to block newly deubiquitylated and deneddylated sites as a mechanism to inhibit re-ubiquitylation, analogous to YopJ. Crystal structures of Cdu1 in complex with ubiquitin and coenzyme A (CoA), demonstrate how a single catalytic center is accessible to different substrates from two different sites (Figure 1.10).

The unique chlamydial insertion α-helix D (VR-3) plays a pivotal role on this dual activity. A positively charged region above Cdu1 VR-3 stabilizes binding to CoA (Lys268 and His249). The active site of Cdu1 may also be reached from a channel below VR-3, where binding to ubiquitin is stabilized by hydrophobic interactions with Ile267 and Ile225. An active-site directed inhibitor against Cdu1 thus has two possible ways to address Cys345, Cdu1's catalytic cysteine residue Importantly, the coenzyme A coordinating residues are not present in Cdu2, making it a dedicated DUB (67).
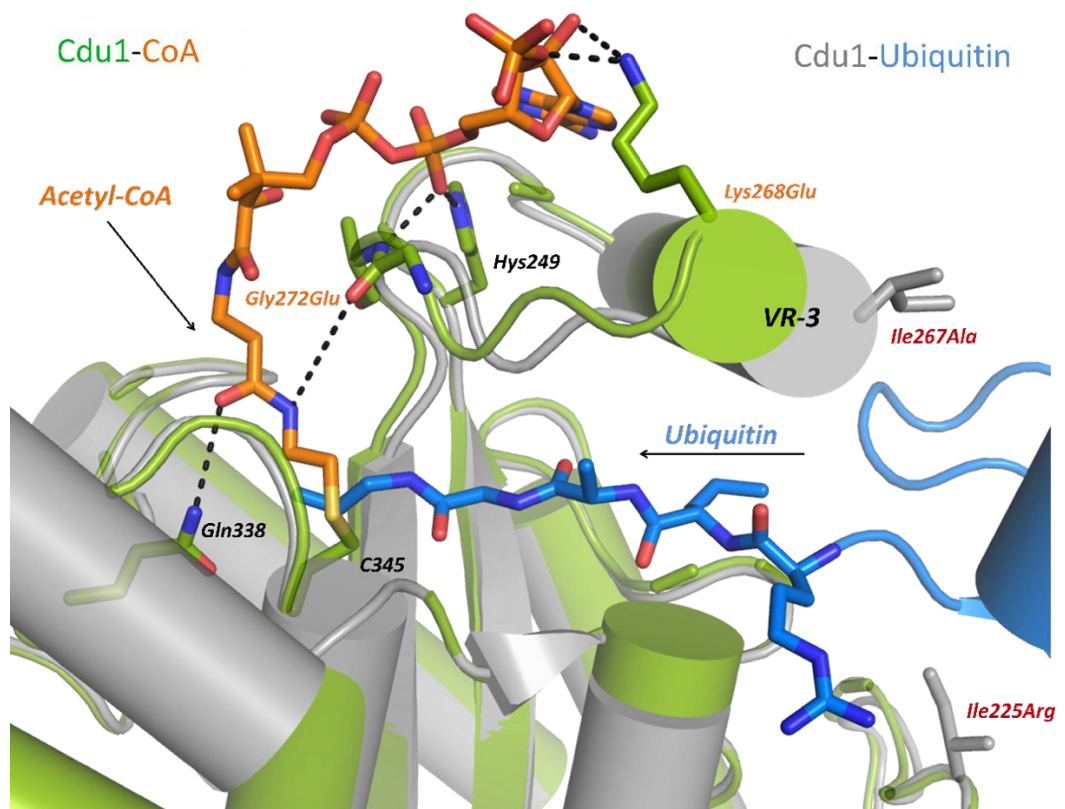
Figure 1.10. The two possible paths for an active-site directed inhibitor to reach the active site cysteine of Cdu1 (black arrows). Mutations on residues which impair binding to CoA and Ubi are displayed in orange and red respectively. Adapted from Pruneda et al 2018 (PDB: 6GZT and 6GZS).

Although the inhibition of NF-κB signaling and cell death by Cdu1 have been attributed to the DUB activity of Cdu1, one cannot discard a yet to be characterized contribution of the AcT activity on these processes.

1.4.5 Covalent inhibition in the CE clan

Cdu1 shares a high degree of similarity with the evolutionarily related adenovirus cysteine protease (adenain), from the C5 family of the CE protease clan (Figure 1.8). Adenain plays a crucial role in the adenovirus lifecycle by facilitating the entry of viral genetic material into the human host cell (106). Interestingly, this protease has also been reported to possess DUB activity (107).

Adenain has therefore been a research topic at the Novartis Institute of Biomedical Research in Basel, Switzerland, where a set of potent and specific inhibitors have been developed (Figure 1.11) (108, 109). These inhibitors are able to address the active site cleft of adenain by forming a covalent bond with the active site cysteine, thereby blocking adenain's proteolytic activity. To date, adenain is the only CE protease which has been the subject of a structure-informed drug design campaign, making its covalent inhibitors an attractive starting point for the development of ligands for other CE proteases.
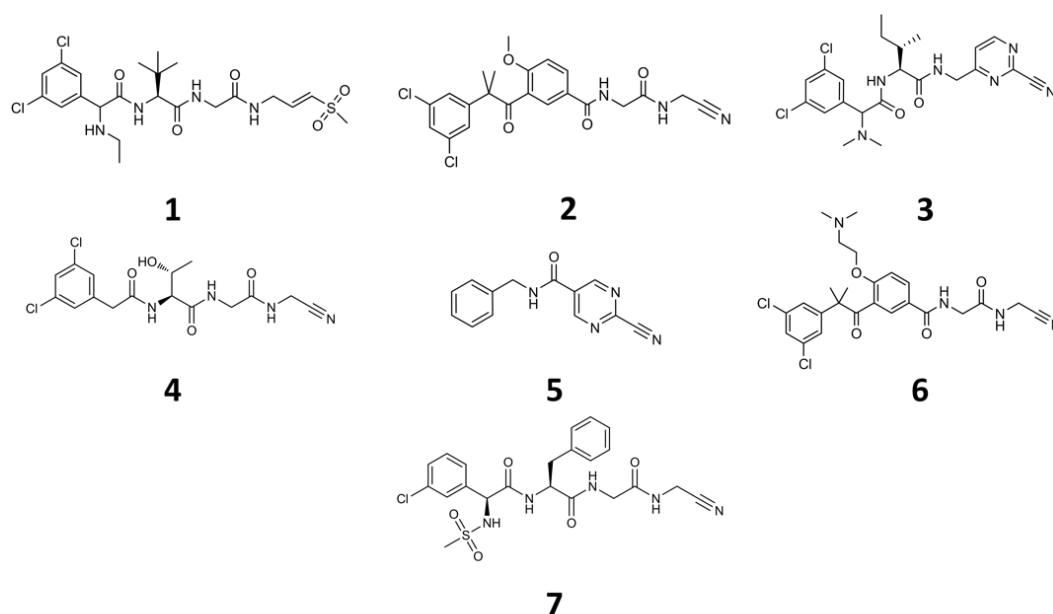
Figure 1.11. Structures of some adenain covalent inhibitors developed at Novartis. Adapted from Mac Sweeney et al. 2014 and Grosche et al. 2015

1.5   Aim of the thesis

In the decades following the so-called Golden age of antibiotics (1950-1960) when one-half of the drugs commonly used today were discovered (110), humanity got the false impression that infectious diseases had been conquered. However, due to the rapid development of antibiotic resistance by many pathogens, mankind is currently in danger to slide back into the dark days of the

pre-antibiotic era. The constant development of alternative ways to treat infectious diseases, including new antibiotics, may serve as additional lines of defense against these pathogens.

Because of their severe sequalae, high health care costs and potential to develop front-line antibiotic resistance, CT infections are considered a significant threat to modern public health. In recent years, the *Chlamydia trachomatis* deubiquitylase 1 (ChlaDub1 or Cdu1), has garnered significant attention as an attractive target for the treatment of CT infections. However, at the beginning of this work, most Cdu1 studies focused on its cellular function, while none provided structural information. To the best of our knowledge, no inhibitors of Cdu1 have been described up to date.

In a joint effort between structural biology and chemoinformatics, this thesis aims to provide insight into the molecular basis of ubiquitin recognition and covalent inhibition of Cdu1. In collaboration with Novartis, we aimed to harness the structural similarities of the active site of Cdu1 with the evolutionarily related protease adenain to obtain the first active-site covalent inhibitors targeting Cdu1. The thereby obtained data, should be used to establish criteria for the structure-informed medicinal chemistry optimization of potential Cdu1 specific inhibitors, and a first lead-molecule backbone against Cdu1.

2 MATERIALS AND METHODS
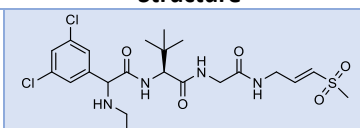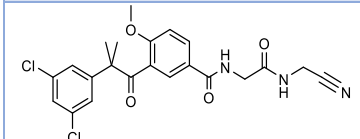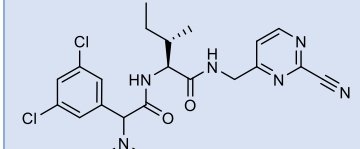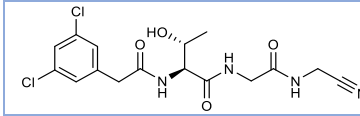
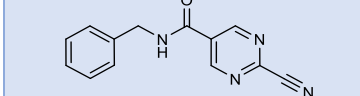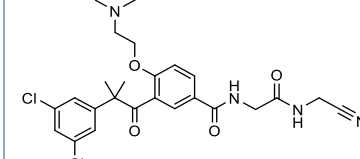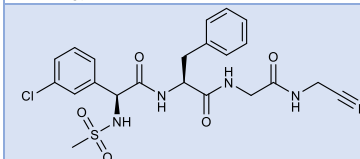## 2.1 Materials

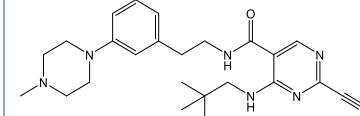### 2.1.1 Consumables

Chemicals

All buffers and solutions were prepared using analytical grade chemicals and ultrapure water from a TKA GenPure system.

Substance	CAS Nr.	Supplier
1,4 Dioxane	123-91-1	Carl Roth
2'-Deoxyadenosine 5'-triphosphate (dATP), sodium salt solution	1927-31-7	New England Biolabs
2'-Deoxycytidine 5'-triphosphate (dCTP), sodium salt solution	102783-51-7	New England Biolabs
2'-Deoxyguanosine 5'-triphosphate (dGTP), sodium salt solution	93919-41-6	New England Biolabs
2'-Deoxythymidine 5'-triphosphate (dTTP), sodium salt solution	18423-43-3	New England Biolabs
2-Mercaptoethanol	60-24-2	Applichem
2-(N-morpholino)ethanesulfonic acid (MES)	145224-94-8	Sigma-Aldrich
4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES)	7365-45-9	Carl Roth
Agarose NEEO ultra quality	9012-36-6	Carl Roth
Ammonium Chloride	12125-02-9	Sigma-Aldrich
Ammonium persulfate (APS)	7727-54-0	Carl Roth
Bicine	150-25-4	Sigma-Aldrich
Bromphenol blue	115-39-9	Carl Roth
Calcium chloride	10043-52-4	Carl Roth
Coomassie Brilliant Blue G-250	6104-58-1	Carl Roth
Coomassie Brilliant Blue R-250	6104-59-2	Carl Roth
Dimethyl sulfoxide (DMSO)	67-68-5	Carl Roth
Disodium phosphate	7558-79-4	Carl Roth
Dithiothreitol (DTT)	3483-12-3	Carl Roth
Ethylenediaminetetraacetic acid (EDTA)	60-00-4	Carl Roth
Ethanol	64-17-5	Carl Roth
Glucose	2280-44-6	Carl Roth
Glycerol	56-81-5	Carl Roth
Hydrochloric acid (HCl)	7647-01-0	Carl Roth
Imidazole	288-32-4	Carl Roth
Iron(II) sulfate	13463-43-9	Fluka
Isopropyl-β-D-thiogalactopyranoside (IPTG)	367-93-1	Carl Roth
Kanamycin sulfate	25389-94-0	Carl Roth
L-Isoleucine	73-32-5	Fluka
L-Leucine	61-90-5	Fluka
L- Lysine	56-87-1	Fluka
L-Phenylalanine	150-30-1	Fluka
L-Threonine	80-68-2	Fluka
L-Valine	516-06-3	Fluka
Magnesium sulfate	7487-88-9	Sigma-Aldrich
Monopotassium phosphate	7778-77-0	Fluka

PEG 400	25322-68-3	Sigma-Aldrich
PEG 20000	25322-68-3	Sigma-Aldrich
Phenylmethylsulfonyl fluoride (PMSF)	329-98-6	Carl Roth
Selenomethionine (Se-Met)	1464-42-2	Acros
Sodium chloride (NaCl)	7647-14-5	Carl Roth
Sodium dodecyl sulfate (SDS)	151-21-3	Carl Roth
Sodium chloride (NaCl)	7647-14-5	Carl Roth
Sodium hydroxide (NaOH)	1310-73-2	Carl Roth
SYBR® Green	163795-75-3	Sigma-Aldrich
Tetramethylethylenediamin (TEMED)	110-18-9	Carl Roth
Tris-(2-carboxyethyl)-phosphine (TCEP)	51805-45-9	Carl Roth
Tris[hydroxymethyl]-aminomethane (Tris)	77-86-1	Carl Roth
Xylene cyanol	2650-17-1	Sigma-Aldrich

Inhibitors

#	Structure	Synthesis	Compound ID
1		Altmann lab, Novartis, Switzerland	AF-85-KR98
2		Altmann lab, Novartis, Switzerland	AA-60-KI75
3		Altmann lab, Novartis, Switzerland	ZD-82-KL49
4		Altmann lab, Novartis, Switzerland	EB-69-PM55
5		Altmann lab, Novartis, Switzerland	ZB-87-UL61
6		Altmann lab, Novartis, Switzerland	JB-75-AM20
7		Altmann lab, Novartis, Switzerland	NB-29-0Y23
93		Altmann lab, Novartis, Switzerland	XE-80-ED00

26

138		Altmann lab, Novartis, Switzerland	**VD-60-IK37**
139		Altmann lab, Novartis, Switzerland	**LD-10-XP35**
140		Altmann lab, Novartis, Switzerland	**DF-10-DX20**
141		Altmann lab, Novartis, Switzerland	**TB-76-AB20**
230		Altmann lab, Novartis, Switzerland	**PF-52-MJ20**
231		Altmann lab, Novartis, Switzerland	**ZD-00-PG32**
232		Altmann lab, Novartis, Switzerland	**ZA-59-OW77**
HJR108		Schirmeister lab, Mainz University, Germany	**KSSK14**

Manufactured composites

Name	Type	Supplier
Bayer silicon grease medium viscosity	Silicon grease	Jena Biosciences
GC buffer	DNA polymerase reaction buffer	New England Biolabs
GeneRuler™ 1 kb DNA Ladder	DNA molecular-weight size marker	Thermo Fisher Scientific
GeneRuler™ 100 bp DNA Ladder	DNA molecular-weight size marker	Thermo Fisher Scientific
HF buffer	DNA polymerase reaction buffer	New England Biolabs
Lysogeny broth (LB) medium (Lennox)	Standard *E. coli* culture medium	Carl Roth
NEBuffer™ 2	Endonuclease reaction buffer	New England Biolabs
PageRuler™ Plus Prestained Protein ladder	Protein molecular-weight marker	Thermo Fisher Scientific
PageRuler™ Prestained Protein Ladder	Protein molecular-weight marker	Thermo Fisher Scientific

peqGOLD PCR Mastermix S	DNA polymerase reaction mastermix	peqlab
Rotiphorese® Gel 30 (37.5:1)	30 % acrylamide/bisacrylamide	Carl Roth
Rotiphorese® Gel 40 (29:1)	DNA ligase reaction buffer	New England Biolabs
T4 DNA Ligase Reaction Buffer	DNA ligase reaction buffer	New England Biolabs

Screens

Name	Type	Supplier
Cdu1 Ultimate Screen	Crystallization screen, optimization of primary conditions	*Self-designed**
Crystal Screen I + II	Crystallization screen, primary sparse matrix screen	Hampton Research
Index	Crystallization screen, primary sparse matrix screen	Hampton Research
Optimix™ 3	Crystallization screen, primary sparse matrix screen	Hampton Research
Optimix™ PEG	Crystallization screen, primary sparse matrix screen	Fluidigm
OptiMix™-PEG	Crystallization screen, primary sparse matrix screen	Fluidigm
PEGs Suite	Crystallization screen, primary sparse matrix screen	Qiagen
Protein Complex Suite	Crystallization screen, primary sparse matrix screen	Qiagen
Wizard Screen I +II	Crystallization screen, primary sparse matrix screen	Emerald BioSystems

* Compositions are given in the Appendix, page 130.

Disposable hardware

Name	*Type*	*Supplier*
24-well hanging-drop crystallization plate	Crystalgen SuperClear™ Plate	Jena Bioscience
96-well sitting-drop crystallization plate	Crystalquick™ 1 square well, flat bottom, low profile	Greiner Bio-One
Centrifugal concentrator	Amicon Ultra-4 and -15	Millipore
Cover slides	22 mm circular cover slides - siliconized	Jena Biosciences
Cuvette	Rotilabo®-single-use cells, 1.6 ml	Carl Roth
Eppendorf tube	SafeSeal tube, 0.5 ml	SARSTEDT AG & Co.
Eppendorf tube	SafeSeal tube, 1.5 ml	SARSTEDT AG & Co.
Eppendorf tube	SafeSeal tube, 2.0 ml	SARSTEDT AG & Co.
Falcon tube	CELLSTAR® 15 ml tubes	Greiner bio-one
Falcon tube	CELLSTAR® 50 ml tubes	Greiner bio-one
Optical quality sealing foil	VIEWseal™	Greiner Bio-One
PCR tube	Multiply® - Pro cup 0.2 ml	SARSTEDT AG & Co.
PCR tube (8 units)	Multiply® - µStrip 0.2 ml chain	SARSTEDT AG & Co.

Kits

Name	Use	Supplier
JBS Floppy-Choppy	*In situ* proteolysis for protein crystallization and structure determination	Jena Biosciences
NucleoSpin® Gel and PCR Clean-up kit	Extraction of DNA fragment from gels and purification of PCR products	Macherey-Nagel
NucleoBond® Xtra Midi	Medium-scale isolation of plasmid DNA from *E. coli* cells	Macherey-Nagel

Enzymes

Name	Type	Supplier
Bovine Serum Albumin (BSA)	Carrier protein (used in the enzymatic reaction buffer)	New England Biolabs
DNase I	Deoxyribonuclease	Invitrogen
DpnI	Restriction endonuclease	Stratagene
HRV-3C Protease	Human Rhinovirus 3C Protease	In-house production
peqGOLD Taq	DNA polymerase	peqlab
Phusion®	High-Fidelity DNA polymerase	New England Biolabs
RecA	Single-stranded DNA binding protein	New England Biolabs
T4 DNA Polymerase	DNA polymerase	New England Biolabs

Plasmids

Name	Description	Resistance*	Origin
pETM14	*E. coli* expression vector providing the T7 promoter/ terminator system, N-terminal 6xHis sequence and 3C protease recognition site.	Kanamycin	Florian Sauer**
pCDF	*E. coli* expression vector providing the T7 promoter/ terminator system, thioredoxin solubility tag, N-terminal 6xHis sequence and 3C protease recognition site.	Streptomycin	Florian Sauer**
pETM14 (SMT3 version)	*E. coli* expression vector providing the T7 promoter/ terminator system, N-terminal 6xHis sequence, SMT3 solubility tag, and 3C protease recognition site.	Kanamycin	Florian Sauer**
pET28a	*E. coli* expression vector providing T7 promoter/ terminator system, N-terminal 6xHis sequence and thrombin cleavage site.	Kanamycin	Annette Fischer***
pUC57	Custom vector encoding for the Cdu1_CT construct, where non-conserved cysteines have been mutated while the active site cysteine has been left intact (see section 3.5.1)	Ampicillin	GeneScript

*Concentration of antibiotics used for selection: 50 μg/ml

**AG Kisker, Rudolf Virchow Center for Experimental Biomedicine (RVZ), University of Würzburg.

***AG Rudel, Biocenter, University of Würzburg.

Ubiquitin and derivatives

Name	Description	Origin
Ubiquitin	Wild type human ubiquitin	Florian Sauer and Theresa Klemm**
Ubiquitin propargylamide	Ubiquitin variant where the C-terminal glycine has been replaced with a propargylamide moiety (PA)	Florian Sauer and Theresa Klemm**
Ubiquitin-Rhodamine 110	Ubiquitin variant derivatized with rhodamine at the C-terminus. Used for enzymatic assays.	UbiQ

**AG Kisker, RVZ, University of Würzburg.

E-coli strains

Name	Use	Genotype	Supplier
DH5α	Plasmid amplification	F- φ80 lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17(rK-,mK+) phoA supE44 λ- thi-1 gyrA96 relA1	Invitrogen
BL21 Star (DE3)	Protein expression	F⁻$ompT$ hsdS$_B$ (r$_B^-$, m$_B^-$) $galdcmrne$131 (DE3)	Invitrogen

Primers

5'-3' sequence	Direction	Variant
SLIC cloning		
CTGGAAGTTCTGTTCCAGGGGCCCCAAACACACCCTTCTACTAAGGAGC	fwd	Cdu1_L (65-401)
CGACGGAGCTCGAATTCGGATCCGGTTATGCTTCAGGCCAAGAAAGCTCT	rev	Cdu1_L (65-401)
CTGGAAGTTCTGTTCCAGGGGCCCAAGGCTCCCAAACCAGTTAAAACGC	fwd	Cdu1_S (125-401)
CGACGGAGCTCGAATTCGGATCCGGTTATGCTTCAGGCCAAGAAAGCTCT	rev	Cdu1_S (125-401)
CTGGAAGTTCTGTTCCAGGGGCCCCGACGACAGACCATCGAAGC	fwd	Cdu1_LCMS (155-401)
CGACGGAGCTCGAATTCGGATCCGGTTATGCTTCAGGCCAAGAAAGCTCT	rev	Cdu1_LCMS (155-401)
AAGTGTCTATGCTATTTCCACACCCTTTACCCAGGT	fwd	Cdu1_S1 (173-401)
AGGGTGTGGAAATAGCATAGACACTTGGGCCCCTGGAACAGAACTTCCAG	rev	Cdu1_S1 (173-401)
AAAATCATCTCGATTTTAGAAGACAAAAAAGCTGTTTTACGAGG	fwd	Cdu1_S2 (202-401)
TTGTCTTCTAAAATCGAGATGATTTTGGGCCCCTGGAACAGAACTTCCAG	rev	Cdu1_S2 (202-401)
AAAGAGAATTACCGCCGCCATTTACAAAGAACA	fwd	Cdu1_KENY (229-401)
TGTAAATGGCGGCGGTAATTCTCTTTGGGCCCCTGGAACAGAACTTCCAG	rev	Cdu1_KENY (229-401)

I

5'-3' sequence	Direction	Fragment	Mutation
Mutagenesis			
GGGATTCTGACATTATTTTCAAGGCGCTATGCTATTTCCACACCCTTTA	fwd	1	C174A
TAAAGGGTGTGGAAATAGCATAGCGCCTTGAAAATAATGTCAGAATCCC	rev	Linear plasmid	C174A
TCAAAGGCCCTCTGCCTATCAGCTGTTCGAAAGAGAATTAC	fwd	2	C226S
GTAATTCTCTTTCGAACAGCTGATAGGCAGAGGGCCTTTGA	rev	1	C226S
CCGGATCCAGCGCGGGCGCTTGGTG	fwd	3	C345A
CACCAAGCGCCCGCGCTGGATCCGG	rev	2	C345A
TCATTTGTCCAGGCTGCGGAAGCGGCTGTTCAG	fwd	Linear plasmid	C368A
CTGAACAGCCGCTTCCGCAGCCTGGACAAATGA	rev	3	C368A

2.1.2 Equipment

Instruments

Name	Model	Supplier
-80 °C Fridge	HERA-Freeze	Thermo Scientific
-20 °C Fridge	Comfort	LIEBHERR
4 °C Fridge	Profi-Line	LIEBHERR

Agarose gel electrophoresis system (DNA)	Mini-Sub® Cell GT System	Bio-Rad Laboratories
Agarose gel electrophoresis system (Protein)	Mini-PROTEAN® Cell	Bio-Rad Laboratories
Autoclave	Systec V-150	Systec
Balance, analytical	XS 105 DR	Mettler-Toledo
Balance	XS 6002S DR	Mettler-Toledo
Cell disruption system	M-110P	Microfluidics
Centrifuge	Avanti J-26 XP	Beckman Coulter
Centrifuge	Avanti J-HC	Beckman Coulter
Centrifuge	Centrifuge 5415 D	Eppendorf
Centrifuge	Centrifuge 5415 R	Eppendorf
Crystallography: cryo-loop	CryoLoop	Hampton Research
Crystallography: sample holder	CrystalCap™ Magnetic	Hampton Research
Crystallography: sample vial	CryoVial	Hampton Research
Crystallography: handling tool	CrystalWand™ Magnetic	Hampton Research
Crystallography: storage pucks	SPINE Puck	Jena Bioscience
FPLC system	ÄKTA avant 25	GE Healthcare
FPLC system	ÄKTA pure 25	GE Healthcare
FPLC system	ÄKTA purifier	GE Healthcare
Heating block	Rotilabo®-Block-Heater H250	Carl Roth
Heating + mixing plate	VS-C7	VWR
Ice Machine	094774	Ziegra
Incubator shaker	ISF1-X	Kühner
Incubator shaker	LT-X	Kühner
Infrared imager	Odyssey	LI-COR
Liquid handling robot	Honeybee 963	Zinsser
Liquid handling robot	LISSY	Heidolph Instruments
MALS detector	DAWN® 8+ HELEOS® II	Wyatt Technology
pH-meter	pH-meter	SCHOTT
Power supply	PowerPack™ Basic	Bio-Rad Laboratories
PCR-cycler	Mastercycler® EPgradient S	Eppendorf
Robotic sealing unit for microplates	RoboSeal	HJ-BIOANALYTIC
Rocker	Duomax 1030	Heidolph Instruments
Rocker	Gyro Mini	Labnet
Rotor assembly (4 x 2250 ml)	JS-5.0	Beckman Coulter
Rotor assembly (8 x 50 ml)	JA-25.50	Beckman Coulter
Spectrophotometer	BioPhotometer	Eppendorf
Spectrophotometer	NanoDrop ND 1000	Peqlab
Thermomixer	Thermomixer comfort	Eppendorf
Ultrapure water system	TKA GenPure	Thermo Scientific
Ultrasonic bath sonicator	Sonorex RK 255 H	BANDELIN electronic
UV-illumination table	Electronic UV Transilluminator	Ultra Lum
UV imaging system	Gel Doc™ XR System	Bio-Rad Laboratories
Vortex mixer	Vortex-Genie 2	Scientific Industries
X-ray cryosystem	X-Stream™ 2000	Rigaku
X-ray detector	R-AXIS HTC	Rigaku
X-ray generator	MicroMax™-007 HF	Rigaku
X-ray optics	VariMax™	Rigaku

Chromatography Columns and Media

Name	Description	Supplier
Protino® Ni-TED	Immobilized metal-ion affinity chromatography resin	Macherey-Nagel
Superdex™ 75 10/300 GL	Preparative SEC FPLC column	GE Healthcare
Superdex™ 75 16/60 GL	Analytical SEC FPLC column	GE Healthcare

Work station components

Type	CPU	Cores	RAM [GB]	HD	Video
FS Celsius R550	2 Xeon (Quad Core)	8	8	2*750 GB (RAID 1)	GeForce Quadro FX 3700

2.1.3   Software

Computer Programs

Name	Description	Supplier / Reference
AIMLESS	Scaling and merging of diffraction data	(111)
AxioVision	Recording of microscopy images	ZEISS
BLAST	Database search program for homologous biosequences	(112)
CCP4	X-ray crystallography; software suite for determination of macromolecular structures	(113)
CCP4i	Graphical interface to *CCP4*	(114)
ChemDraw	Chemical structure-drawing program	PerkinElmer
Coot	Model-building software, X-ray crystallography	(115)
CLARIOstar®	Microplate reader control interface	BMG LABTECH
CrystalClear	X-ray data collection and basic processing	Rigaku
DALI	Protein 3D conservation mapping	(116)
DSSP	Secondary structure assignment	(117)
ExPASy ProtParam tool	Computation of physical and chemical properties of proteins	(118)
GENtle	Plasmid map database management	University of Cologne (http://gentle.magnusmanske.de/)
ExPASy Translate tool	Translation tool of nucleotide sequences to protein sequences	(118)
MARS	Data analysis software	BMG LABTECH
Microsoft Excel	Spreadsheet software	Microsoft Corporation
ODYSSEY	Infrared imaging software	LI-COR
PHASER	Phasing software	(119)
Phenix	Software suite for macromolecular X-ray structure determination	(120)
Phyre2	Biosequence analysis; protein 3D-structure prediction	(121)
POINTLESS	Space-group determination	(122)

Prism	Scientific graphing	GraphPad Software
PyMOL	3-dimensiona molecular visualization and graphical illustration software	Schrödinger
Quantity One®	UV imaging system control; UV image recording and analysis	BioRad
REFMAC	Macromolecular structure refinement	(123)
UNICORN	FPLC instrument control; recording, analysis and management of chromatograms	GE Healthcare
XDS	Processing and scaling of diffraction images	(124)

Chemoinformatics and docking

Name	Description	Version	Supplier / Reference
Canvas	Cheminformatics computing environment	3.5.011	Schrödinger
Drugscore	Knowledge-based re-scoring of ligand poses	0.89	(125)
HERMES	Graphical user interface for GOLD	1.8.1	CCDC Software Ltd
GOLD	Genetic algorithm program for docking flexible ligands into protein binding sites	5.4.1	(126)
MOE	Drug discovery software platform for visualization and modeling	2018.01	(127)

Databases

Name	Content	Web address	Reference
PDB	Structural data of biological macromolecules	www.rcsb.org	(128)
UniProt	Protein sequences	www.uniprot.org	(129)
ZINC15	Commercially-available compounds for virtual screening	zinc15.docking.org	(130)

2.2 Methods description

2.2.1 Molecular biology

Molecular Cloning

The process of molecular cloning implies the introduction of a recombinant deoxyribonucleic acid (DNA) encoding for a target protein into a suitable vector for its subsequent expression and purification. The assembly of DNA fragments was achieved using homologous recombination and single-strand annealing. This technique is known as sequence and ligation-independent cloning (SLIC) (131). Appropriate primers for the amplification of the different Cdu1 (see section 3.2.1) constructs were designed using GENtle and ordered (Sigma). Two versions of the pETM14 plasmid containing a Hexa-histidine (His6) and an His6-SMT3 tag upstream to the precision protease (3C) cleavage site, were linearized (See appendix, Figure 6.1).

Component	Amount per 50 µL reaction	Final concentration
Ultra-pure water	37.9 µL	n.a
5X Phusion HF buffer	10 µL	1X
10 mM dNTPs	1 µL	200 µM
Forward primer	0.1 µL	0.5 µM
Reverse primer	0.1 µL	0.5 µM
DNA template	10 ng	n.a
Phusion DNA polymerase	0.5 µL	0.02 U·µl^{-1}

Table 2.1. Phusion® polymerase reaction mixture

Gene fragments were amplified using the Phusion high fidelity polymerase, according to the manufacturer's guidelines (Table 2.1 and appendix, Figure 6.2). The thermocycler program suggested by the manufacturer was used for all gene amplifications (Table 2.2). The annealing temperature and extension time were adjusted in each case according to the primer's melting temperature and length of the construct, respectively.

Cycle step	Temperature	Time	Cycles
Initial denaturation	98 °C	2:00 min	1
Denaturation	98 °C	10 s	30
Annealing	50-80 °C	30 s	30
Extension	72°C	2:00 -5:00 min	30
Final Extension	72°C	10:00 min	1
Hold	4 °C	∞	n.a.

Table 2.2. Thermocycler program used for the amplification of DNA fragments

The PCR products were analyzed by agarose gel electrophoresis (see next paragraph) using a Mini-Sub® Cell GT System, followed by incubation in the presence of 0.1-0.2 U·µl^{-1} DpnI for 3 h at 37 °C to digest parental DNA. The samples were purified using a NucleoSpin® Gel and PCR Clean-up kit**.** Single-stranded 3'-5' overhangs were generated by incubating the DNA with 0.3 U·µl^{-1} of T4-DNA polymerase in NEBuffer 2 and 0.2 µg·µl^{-1} BSA for 45 min at room temperature. The digestion was stopped by adding 1 mM dCTP. Subsequently, the fragments were assembled by mixing the insert and the vector at a 4-fold molar excess of insert and heated to 75°C for 5 min followed by 10 min at 25°C. All constructs were cloned into the pETM14 vector, except for the Cdu1_L fragment which was additionally cloned into the pCDF vector with an N-terminal Thioredoxin (TRX) solubility tag. The reaction products were transformed into competent DH5α *E. coli* cells, as described in Page 36. Successful cloning was verified by gene sequencing (Eurofins Genomics, Ebersberg, Germany).

Agarose gel electrophoresis

Separation of deoxyribonucleic acids (DNA) was achieved by agarose gel electrophoresis. Gels were prepared by dissolving 1.0% (w/v) NEEO ultra quality agarose in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA pH 8.0) and approximately 0.4 µg/mL SYBR® Green was added as nucleic acid stain. A Mini-Sub cell GT system was used to cast the gels. PCR reaction products and DNA samples were diluted six-fold to reach the working concentration of the 6X DNA sample

buffer (10 mM Tris, 0.3% bromphenol blue, 0.3% xylene cyanol, 60% glycerol, 60 mM EDTA, pH 7.6), and loaded into the gel chambers submerged in TAE buffer, along with an appropriate molecular weight marker. Electrophoresis was performed at 120 V for 30 to 40 minutes at room temperature. All gels were exposed to UV light in a Gel DocTM XR transillumination system for detection of the DNA. Pictures of the gels were recorded using the built-in camera of the device.

Transformation of competent *E. coli* cells

For the transformation of plasmids into chemically competent DH5α *E. coli* cells, 50-100 ng of vector were preincubated on ice with a 100 µL aliquot of the competent cells for 20 minutes (30 minutes for SLIC). Heat shock transformation was induced by exposing the cells to 42 °C for 60 seconds. After a 5-minute incubation on ice, 0.5 ml of LB medium were added and the cells were incubated at 37 °C for 30 minutes (60 minutes for SLIC) in a thermomixer. Afterwards, 200 µL of this mixture were plated on an LB agar plate with the appropriate antibiotic and incubated overnight at 37 °C.

Colony PCR

To verify the presence of the specific DNA inserts in the *E. coli* clones, individual colonies were picked and transferred into a PCR reaction tube containing the PCR reaction mix supplemented with the appropriate primers (peqGOLD PCR Mastermix S, 2.5 mM $MgCl_2$, 1 µM forward primer, 1 µM reverse primer). The PCR-reaction was carried out according to the protocol depicted in Table 2.2. The clones were analyzed by agarose gel electrophoresis.

Amplification of plasmid DNA

To amplify large amounts of plasmid DNA for long term storage, vectors were transformed into DH5α *E. coli* cells as described above. A single colony was transformed into 200 ml of sterile LB medium containing the appropriate antibiotics and incubated over night at 37 °C under constant shaking at 250 rpm. After separating the cells from the medium by centrifuging at 4000 x g for 20 minutes at 4 °C, plasmid DNA was isolated according to the manufacturer's instructions using the NucleoBond® Xtra Midi kit. The DNA was aliquoted, frozen and stored at x C.

Mutagenesis

Mutations of the non-conserved cysteines were achieved by generation of 4 DNA fragments including the desired mutations (see section 2.1.1, mutagenesis and Figure 6.4 A the appendix) which contained overhangs complementary to each other. The PCR products after each amplification were combined and amplified in a thermocycler using the following program:

Cycle step	Temperature	Time	Cycles
Initial denaturation	98 °C	2:00 min	1
Denaturation	98 °C	10 s	10
Annealing	50-80 °C	30 s	10
Extension	72°C	2:00 -5:00 min	10
Final Extension	72°C	10:00 min	1
Hold	4 °C	∞	n.a.

Table 2.3. Thermocycler program used for the amplification mutagenesis.

Once all fragments were combined (appendix, Figure 6.4 B and C), the resulting insert was inserted into a linearized pETM14 vector (appendix, Figure 6.4 D) using SLIC, as described on page 34.

Expression test

To evaluate protein expression from the cloned constructs, small-scale expression tests were performed. 50 ml of LB media (supplemented with the appropriate antibiotics) were inoculated with 2 ml from a pre-culture of transformed BL21 Star (DE3) *E. coli* cells which grew in the presence of 50 µg/ml kanamycin. Cells were incubated at 37 °C under constant shaking at 250 rpm until an optical density of 1.0 at λ = 600 nm (OD_{600}) was reached. At that moment, protein expression was induced by adding 1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), while the temperature was reduced to either 15 or 30°C for overnight incubation. An identical control, which was not induced with IPTG, was prepared simultaneously.

A 1.5 ml aliquot from each condition was transferred to an Eppendorf tube and the cells were separated from the media by centrifugation for 10 minutes at 4 °C and 3220 x g. The pellet was resuspended in lysis buffer (page 42), and the cytosolic content of the cell was extracted by submitting the tubes to 3 freeze-thaw cycles with liquid nitrogen. Subsequently, the soluble and insoluble fractions were separated by centrifugation for 15 min at 20,000 x g and 4 °C. Fractions were then analyzed by SDS-polyacrylamide gel electrophoresis as described on page 39.

Protein expression and lysis

Large scale expression of recombinant Cdu1 was conducted similar as for the expression test but at a volume of 2 l of sterile LB medium supplemented with 50 µg/ml kanamycin and inoculated with 20 ml of a pre-culture of transformed BL21 Star (DE3) *E. coli* cells. After incubation at 37 °C under constant shaking, protein expression was induced with the addition of 1 mM IPTG at an OD_{600} of 1.0 and the cells were grown overnight at 15 °C. At the end of the expression phase, the cells were harvested by centrifugation at 4000 x g at 37 °C for 15 minutes. The cell pellets were either aliquoted and frozen at -80 °C or resuspended in lysis buffer in preparation for affinity chromatography (page 42).

Limited proteolysis and sample preparation for LCMS analysis.

Dilutions from the proteases in the JBS Floppy-Choppy® kit were prepared according to the manufacturer's recommendations (10^{-1} to 10^{-3}). Approximately 100 µg of Cdu1_S were transferred into 4 different vials and the proteolytic reactions were started by adding 10 µl of each diluted protease and incubated for either 30 min, 1 h, 1:45 h or 2:15 h. The reactions were stopped by diluting each sample with SDS buffer (1:5) and incubated at 95 °C for 5 min. Iodoacetamide (IAA) was added to a final concentration of 14 mM to alkylate the cysteines. The samples were incubated for 30 min at room temperature and in the dark. The proteolytic products were analyzed by SDS-PAGE, as described in section 2.2.2. The bands of interest were cut out of the gel and submitted for LCMS analysis in the laboratory of Prof. Dr Andreas Schlosser.

Generation of the Cdu1~ubiquitin covalent adduct

For the generation of the Cdu1-Ubi covalent adduct, 15 µM of the Cdu1_CT variant was incubated overnight at 4 °C with 30 µM ubiquitin propargylamide (Ub-PA, kindly provided by Dr. Florian Sauer) Complex formation was confirmed through SDS-PAGE and the fusion product was purified by size exclusion chromatography utilizing a Superdex 75 10/300 column (GE Healthcare) and the same buffers as for wild type Cdu1.

2.2.2 Biochemical and biophysical characterization

SDS-PAGE

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was used to assess protein purity based on the molecular weight of the individual proteins (132). Protein samples were mixed 5:1 with sample buffer and incubated for 5 minutes at 95 °C. Through this process, the proteins lose their 3D structures and

their primary chains can interact with the dodecyl chain of SDS. Polyacrylamide gels with a 15% acrylamide/bisacrylamide resolving gel and 5% acrylamide/bisacrylamide stacking gel were installed in a Mini Protean II system. 10 µl of each sample were loaded into the wells of the polyacrylamide gels next to 3 µl of a protein ladder sample to estimate the size of the protein bands. Electrophoresis was performed at 250 V for 30-40 minutes at room temperature, with the negatively charged proteins migrating towards the anode. The gels were stained by immersing them in G-250 staining solution and heating them for 1 min at 800 W using a microwave. The gels were destained by repeating the same procedure using distilled water until only the protein bands, and not the dye, were visible.

Separating gel	10%- 18% acrylamide/bisacrylamide mix, 167 mM Tris-HCl pH 8.8, 0.1% SDS, freshly added: 0.1% APS and 0.04% TEMED
Stacking gel	5% acrylamide/bisacrylamide mix, 125 mM Tris pH 6.8, 0.1% SDS, freshly added 0.25% APS and 1% TEMED
SDS-PAGE running buffer	192 mM glycine, 25 mM Tris, 0.1% SDS
SDS-PAGE sample buffer	250 mM Tris pH 6.8, 0.5 M DTT, 0.5% (w/v) bromphenol blue, 10% (w/v) SDS, 50% glycerol
G-250 staining solution	80 mg Coomassie, 3 ml HCL to 1 l ddH2O

Molecular weight determination through multi-angle light scattering (MALS)

To determine the molecular mass of Cdu1 in solution, SEC-MALS experiments were performed at room temperature. 100 µl protein solution at a concentration of 30 µM were injected onto a Superdex 75 10/300 GL analytical size exclusion chromatography column (GE Healthcare) pre-equilibrated with reducing or non-reducing buffer (25 mM HEPES pH 7.5, 250 mM NaCl with or without 2 mM DTT) which was coupled to a MALS detector (DAWN HELEOS II, Wyatt Technology) and an RI detector (Optilab t-rEX, Wyatt Technology). Mass analysis was achieved with the Astra software (Wyatt technology).

Determination of protein concentrations using UV spectrophotometry

To determine the concentration of the purified Cdu1 constructs, the UV absorbance of the protein was measured using a NanoDrop ND 1000 UV/Vis spectrophotometer. A UV spectrum corrected by the absorption of the buffer and ranging from 220 nm − 350 nm was recorded. The *ExPASy ProtParam* tool was used to estimate the molar extinction coefficient from the protein's primary sequence. The protein concentration was then calculated from the absorption (A) at λ = 280 nm and using the Lambert-Beer equation:

$$A = \varepsilon bc$$

Where ε is a wavelength-dependent absorptivity coefficient, b is the path length, and c is the analyte concentration (133).

2.2.3   Protein purification

High purity and homogeneity of the protein samples are essential factors which facilitate crystallization. Therefore, all Cdu1 constructs utilized in this work were purified with two chromatographic steps: immobilized affinity chromatography (IMAC), which isolates proteins decorated with a hexa-histidine tag; and size-exclusion chromatography (SEC), which separates the target protein from other protein aggregates and other contaminants based on their size differences.

Cell lysis

Cell pellets were thawed on ice overnight and resuspended in lysis buffer (50 ml for every 10 g of cells), supplemented with a spatula tip of phenylmethane sulfonyl fluoride (PMSF) and 1 U/ml DNaseI. To separate the cytosolic content of the cells from the membranes, the cells were pumped through a cell disruption system 3

times applying a pressure of 1.5 kbar, followed by centrifugation at 25000 x g and 4 °C for 1 h (JA 25.50 rotor). The supernatant was used for Ni^{2+} affinity chromatography.

Lysis buffer	25 mM HEPES (pH 7.5), 100 mM NaCl, 5 mM imidazole and 5 mM β-mercaptoethanol

Affinity Chromatography and cleavage of the affinity tag with the 3C protease

All Cdu1 constructs were purified by Immobilized Metal Affinity Chromatography (IMAC), which is based on the affinity between Ni^{2+} ions coordinated with a chelating resin, and a Hexa-histidine motif attached to the recombinant proteins to achieve their separation (134). The cell lysate from 10 to 15 g of cells was mixed with 1.0 to 1.5 g of Protino® Ni-TED beads and incubated under mild agitation for 1.5 h at 4°C. The supernatant was carefully removed with an electric pipet. The nickel beads were washed with 3 x 50 ml lysis buffer and loaded into an empty column. The protein was subsequently eluted in 10 fractions of 10 ml each using elution buffer supplemented with 250 mM imidazole, which releases the protein from the nickel by competing for its coordination sites. Fractions containing Cdu1 were dialyzed overnight against dialysis buffer, while the His-tag was cleaved off by the addition of 3C protease.

Lysis buffer	25 mM HEPES (pH 7.5), 100 mM NaCl, 250 mM imidazole and 5 mM β-mercaptoethanol
Dialysis buffer	25 mM HEPES (pH 7.5), 100 mM NaCl and 1 mM DTT

Size-exclusion Chromatography (SEC)

Size-exclusion chromatography allows the separation of proteins according to their molecular size and shape. The matrix of a gel filtration column (stationary phase) contains a matrix of porous beads made of agarose and dextran, which

have different sizes. When a protein-containing sample is applied, smaller macromolecules have a greater retention time than bigger ones, as they penetrate deeper into the porous matrix. Bigger proteins or aggregates have a shorter retention time and exit the column first as they cannot enter the beads and must run past them (135).

All Cdu1 constructs were purified utilizing a Superdex 75 16/60 column (GE Healthcare) in SEC buffer. After IMAC and cleavage of the affinity tag, Cdu1 samples were concentrated in an Amicon centrifugal concentrator with a MW cutoff of 10.000 Da until a volume of 5 ml was achieved. Injection into either an ÄKTA avant 25, ÄKTA pure 25 or ÄKTA purifier FPLC system (GE Health Care) was performed after equilibrating the column with 1.2 column volumes (CV) at a flow rate of 1.5 ml/min. Protein fractions were eluted with 1.2 CV, and fractionation was started at 0.3 CV. The elution was monitored by measuring the UV absorption at 260 and 280 nm. Peak fractions were analyzed by SDS-PAGE (as described on page 39). All Cdu1 samples were concentrated in an Amicon centrifugal concentrator with a MW cutoff of 10.000 Da until a concentration of 10 mg/ml was reached, except for the Cdu1_Mut construct which was concentrated to 20 mg/ml. Subsequently the protein was aliquoted into 25 µL fractions, frozen in liquid nitrogen and stored at – 80 C.

SEC Buffer	25 mM HEPES (pH 7.5), 100 mM NaCl and 1 mM DTT.

2.2.4 Enzymatic assays
Sections in between " " are adapted from Ramirez et al, 2018 with permission

Enzyme kinetics is the study of chemical reactions which are catalyzed by enzymes (136). The measurement of the kinetic parameters of enzymatic reactions provides valuable information of biochemical events such as complex formation and catalysis. In drug discovery, the interaction between a macromolecule and small inhibitor-like molecules may be characterized by enzyme kinetics to lead

drug development campaigns. Enzyme-catalyzed reactions may be described by the following scheme:

$$E+S \xrightleftharpoons[k_{-1}]{k_1} [ES] \xrightarrow{k_2} E+P$$

Where the concentration of the Enzyme-Substrate complex [ES] is dictated by the ratio of its dissociation into either enzyme and free substrate and enzyme and product ($k_{-1} + k_2$) and its formation from enzyme and free substrate (k_1).

K_M determination

If the substrate is in large excess over the enzyme, the *Steady state* approximation states that the formation of the ES binary complex is exactly matched by its decay into free enzyme and products and the temporal variation of [ES] equals 0.

$$\frac{d[ES]}{dt} = 0$$

Under these conditions, the *Michaelis—Menten equation* describes the rate (V_o) achieved by the system under substrate saturating conditions and is defined as:

$$V_o = \frac{V_{Max}[S]}{K_M + [S]}$$

Where V_{Max} is the maximum velocity and the *Michaelis-Menten constant, K_M*, is the substrate concentration at which half of the enzyme active sites in the sample are saturated (136).

"The K_M for Cdu1 was determined using the fluorogenic substrate Ub-Rh110Gly (137) (UbiQ Bio BV). The kinetic parameters were determined by monitoring the progression of substrate cleavage using a CLARIOstar plate reader (BMG LABTECH) with fluorescence excitation at 487 nm and emission detection at 535 nm in the presence of 20 mM NaCl, 20 mM HEPES (pH 7.5), 1 mM TCEP, and 50 µg/mL BSA at 28 °C. A substrate concentration response curve was plotted by titrating 0.3 nM

Cdu1_S (125-401) with increasing concentrations of Ub-Rh110Gly (0.25 to 5.0 µM).'' The data for the initial velocities and substrate concentration were fitted into the *Michaelis—Menten equation* using GraphPad®.

Determination of the percentage of inhibition

To evaluate the ability of a compound library to inhibit an enzyme at a given concentration, an appropriate metric needs to be defined. This parameter should be simple enough to be performed at a cost-efficient manner at the laboratory for a large number of molecules, while remaining robust enough to reliably identify hits. In this work, all molecules which inhibited more than 50% of Cdu1's enzymatic activity after incubation for 1 hour at an inhibitor concentration of 100 µM were defined as hits. The percentage of inhibition (138) is therefore defined as:

$$\% \, of \, inhibiton = 100 \left(1 - \frac{V_i - S_b}{V_0 - S_b}\right)$$

Where V_i and V_0 are the reaction velocity in the presence and absence of inhibitor at 100 µM, respectively, and S_b is the background signal.

''The inhibition of Cdu1 activity by a set of 7 adenain inhibitors (109, 139) (see section 1.4.5) and 180 vinyl-methyl-ester (VME) fragment-like covalent binders for cysteine proteases (140, 141) (Appendix Table 3) was evaluated in 20 mM NaCl, 20 mM HEPES (pH 7.5), 1 mM TCEP, and 50 µg/mL BSA at 28°C. 0.6 nM Cdu1 (125-401) were incubated for 1.5 h at room temperature in the presence of 100 µM of each inhibitor (2% final DMSO concentration). The reaction was started by adding 20 µL of 500 nM Ub-Rh110Gly (final volume of 40 µL per well). This resulted in a final enzyme and substrate concentration of 0.3 nM and 250 nM, respectively. A well containing all components of the reaction except for the enzyme and the inhibitor was used as a blank. For the determination of the initial velocities, the data points corresponding to the first 2 minutes of the reaction were blank-

corrected and plotted against time. The percentage of inhibition for each ligand was calculated as the ratio of the initial velocities for the inhibited reactions and a DMSO control''.

IC$_{50}$ determination

The IC$_{50}$ value represents the concentration of inhibitor needed to impair the enzymatic activity by half (142). Although it has been suggested that IC$_{50}$ values might not be the best way to describe covalent inhibitors as this parameter is known to depend on the incubation time (140), they are certainly useful to obtain an idea of the potency of a given covalent drug, especially when they have been shown to have a reversible inhibition mechanism (143).

''Determination of the IC$_{50}$ values for cyano-pyrimidines **3, 5** and **HJR108** was performed by titrating 0.3 nM Cdu1_S (125-401) with increasing inhibitor concentrations ranging from 3 to 3000 µM. After incubation for 1 h, the rate of Ub-Rh110Gly cleavage was measured in duplicate at a substrate concentration equivalent to the K$_M$ for three independent protein batches. The dose response semi-logarithmic curves were fitted using GraphPad and the logistic function:

$$y = A1 + \frac{(A2 - A1)}{1 + 10^{(x - log(IC_{50}))}}$$

where y is the % of inhibition at the inhibitor concentration x, A1 is the lowest inhibition value and A2 is the maximum inhibition value.''

Reversibility assay

The characterization of the underlying mechanisms involved in covalent inhibition are generally considered to be more complex than for their non-covalent counterparts (144). This is partially because they involve at least a two-step reaction mechanism which may be, in the simplest case, represented as follows:

$$E+I \xrightleftharpoons[k_2]{k_1} E\text{-}\text{-}I \xrightleftharpoons[k_4]{k_3} E\text{-}I$$
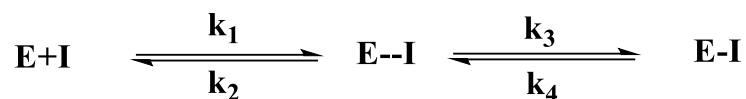
The first step is the formation of the non-covalent enzyme-inhibitor complex E--I, which is always reversible. If the covalent warhead is appropriately positioned in the vicinity of a reactive side chain of the protein (for instance, a cysteine), a second step takes places place which involves the formation of a covalent bond. The ratio between k_1 and k_2 determines if the non-covalent complex is stable enough for the covalent reaction to take place. For $k_4 \sim 0$, the reaction is considered irreversible and when $k_4 \gg 0$ is considered reversible. Assessing the reversibility or irreversibility of a given inhibitor is an important step to be taken into consideration for drug-design, as this parameter could be directly related with the drug's safety profile (see section 1.3.3). Experimentally, however, dissecting the individual kinetic contributions of the non-covalent association and the bond formation may be a laborious task (143). In this work, the reversibility of the Cdu1 covalent inhibitors was assessed through a rapid dilution assay (145). "30 nM Cdu1_S (125-401) was incubated with approximately 10-fold the IC_{50}-equivalent concentration of cyano-pyrimidines **3**, **5** and **HJR108**. After incubation for 1 h at room temperature, samples were diluted 100-fold with assay buffer containing substrate, and the time course of substrate cleavage was monitored. In the enzyme control, Cdu1_S was mixed with DMSO before dilution into buffer-containing substrate. In the substrate control, no enzyme was used."

2.2.5   Molecular docking

Pose reproduction by covalent docking

Understanding the principles underlying the recognition of small molecules or inhibitors by macromolecular biological receptors is of paramount importance. The successful reproduction of experimentally observed ligand poses through

simple and cost-efficient computational methods readily enables structure-informed drug design campaigns. The software GOLD (**G**enetic **O**ptimization for **Li**gand **D**ocking) uses a nature-inspired evolutionary algorithm which modifies and optimizes the position, orientation and conformation of a ligand within a protein's binding pocket. The ligand geometry parameters are mapped into chromosomes, which are modified by a set of genetic operators and evaluated by a scoring function. This process is repeated through several generations, until convergence into a satisfactory solution (146). In the present work, the consensus between the empirical ChemScore and the knowledge-based DrugScoreX scoring functions was used.

ChemScore, an empirical scoring function

ChemScore (147) is one of the scoring functions used to evaluate the different ligand poses generated by GOLD. It estimates the free energy of binding according to the expression:

$$\Delta G_{binding} = \Delta G_0 + \Delta G_{hbond}S_{hbond} + \Delta G_{metal}S_{metal} + \Delta G_{lipo}S_{lipo} + \Delta G_{rot}H_{rot}$$

where the ΔG terms are coefficients derived by multiple linear regression from exerimentally observed interactions in a training set of 82 protein-ligand complexes with crystal structures in the PDB and measured affinities. S_{hbond}, S_{metal} and S_{lipo} are the descriptors for the hydrogen-bond, metal and lipophilic interactions, respectively. The H_{rot} term accounts for the loss of conformational entropy of the ligand upon binding (148).

DrugScoreX, a knowledge-based scoring function

DrugScoreX (DSX) (125), is a knowledge-based scoring function that calculates the total score of a binding pose based on individual statistical potentials derived from a database of known protein-ligand complexes, in contrast to empirical scoring functions which decompose the total energy into discrete individual terms (such

as ChemScore). DSX defines the scores based on distant-dependent atom pair potentials ($score_{pair}$), torsion angle potentials ($score_{tors}$), and solvent accessible surface-dependent potentials (SAS) ($score_{SR}$) derived from comprehensive structural data from the Protein Data Bank (PDB) (128) and the Cambridge Structural Database (CSD) (149). The DSX score is calculated as follows:

$$Score_{total} = w_p score_{pair} + w_t score_{tors} + w_s score_{SR}$$

where $w_{p/t/s}$ are weighting coefficients of the individual potentials.

For pose-reproduction (see section 3.8.1), "the crystal structures of the protein-ligand complexes of cyano-pyrimidines **3** and **5** (PDB: 6FDU, 6FDQ) were prepared for docking using MOE (Chemical Computing Group, 2018) (127) (127) (127) (127) (127) (149) (149). After removal of ligands and water molecules MOE's protein structure preparation pipeline (150-155) was applied to automatically assign tautomer and protonation states and optimize the hydrogen-bond networks. Protonate3D (156) was internally invoked at pH 7.5. Cyano-pyrimidines **3** and **5** were additionally subjected to a MMFF94x (157, 158) geometry optimization in MOE (RMS gradient convergence criterion: $0.001 \cdot kcal \cdot mol^{-1} \cdot Å^{-1}$). Post-reacted species with a nucleophilic sulfur atom attached to the cyano group of cyano-pyrimidines **3** and **5** (required by GOLD) were manually prepared in MOE's Builder tool. GOLD (126, 146) (v5.4.1) was used to dock the ligands under the constraint of a covalent link to Cys345. Non-default GOLD (159) settings (see appendix, Table 6.1 and 6.2) regarding enhanced ligand confor-mational flexibility were used. ChemScore was found to robustly model key intermolecular interactions and, hence, to yield the best crystal pose reproduction among all of GOLD's scoring functions. Pose reproduction performance was assessed by RMSDs between docked poses and crystal ligand structures calculated in GOLD. The 50 best ChemScore poses were rescored using the DSX (v0.89) (with CSD and PDB potentials (version 05/11; in the case of CSD potentials, torsion, intramolecular clash and SAS potentials were also activated). For selection of the best poses, a

consensus ranking approach was devised by weighting the ranks of each of the 3 scoring functions equally: $1/3 \cdot (R_{ChemScore} + R_{DSX\_PDB} + R_{DSX\_CSD})$."

Virtual screening and sub-structure search

Commercially available cyano-pyrimidines were retrieved using the search tool of the ZINC15 database (130). Substructure searches and library preparation was performed using Canvas (Schrödinger). Preparation of the compound library and covalent docking analysis was performed using the same settings as for pose-reproduction as described in page 47.

Ligand design and validation

Compound design was performed based on the structural and docking data obtained for Cdu1 (see section 3.9) using MOE. Preparation and evaluation of the generated compound library was conducted using the same settings and criteria as in the previous section.

2.2.6   X-ray crystallography

Protein Crystallization

Only the purest protein fractions obtained after SEC (see section 2.2.3) were used for crystallization. Vapor diffusion crystallization experiments were carried out with a HoneyBee 963 robot in the sitting drop format. The drops consisted of 0.3 µl protein solution mixed with 0.3 µl mother liquor and a reservoir of 40 µl mother liquor in a 96-well crystallization plate and sealed with adhesive sealing film. The mother liquor and reservoir composition of each well consisted of the commercially available screens depicted in Table 2.4, which were set up by Nicole Bader. Each screen consists of 96 different buffer conditions.

Screen	Company	Origin
Crystal Screen I + II	Hampton	In house production
Index	Hampton	In house production
Nextal - PEG Suite	Qiagen	In house production
Nextal pH Clear	Qiagen	In house production
Nucleix Suite	Qiagen	In house production
Protein Complex Suite	Qiagen	In house production
Topaz OptiMix 3	Fluidigm	In house production
Topaz OptiMix PEG	Fluidigm	In house production
Wizard I + II	Emerald	In house production

Table 2.4. Commercially available crystallization screens used in the present work

Crystallization of Cdu1_Mut (155-401) was achieved overnight at 20°, at a protein concentration of 20 mg/ml with a reservoir solution containing 100 mM Bicine pH 9.0, 10% PEG 20000 and 2% 1,4-dioxane.

Crystal seeding

For crystallization of wild type Cdu1_LCMS (155-401), microseeds obtained from Cdu_Mut (155-401) crystals were required. Crystal seeding can be used to increase the number of positive hits in a crystallization screen, and by using the crystals of a mutant protein there is a chance to obtain native crystals and vice versa (160). Using a pipette, a 2 µl mother liquor drop containing Cdu1_Mut crystals was carefully transferred into a PCR tube containing 5 µl of mother liquor. The mix was then macerated, vortexed and sonicated to generate micro crystals. A serial dilution of this seed solution was prepared from 10^0 to 10^{-3}, using the mother liquor. A 10 mg/mL stock of the Cdu1_LCMS construct was supplemented with 6 mM DTT directly prior to the crystallization assay and the freshly set drops (1 µl protein + 1 µl mother liquor) were streaked with a cat whisker soaked in the different seed dilutions. Formation of the first crystal-like objects was observed immediately across the streak line; the plate was therefore left at 4 °C to decrease

the speed of crystal formation. This process was repeated 3 times in consecutive days using the best obtained crystals as seeds for the next round in this serial seeding process, until plate like crystals of a suitable size were obtained. This seeding approach produced no crystals for the Cdu1_L (65-401) and Cdu1_S (125-401) constructs. Crystals of the selenomethionine containing Cdu1 construct were obtained by using wild type Cdu1_LCMS (155-401) crystals as seeds and utilizing the same crystallization conditions as described above.

Soaking and co-crystallization

For the structural analysis of Cdu1 bound to covalent inhibitors, Cdu1_LCMS crystals were soaked in cryo protectant solution (100 mM Bicine pH 9.0, 10% PEG 20000, 2% 1,4-dioxane and 40% PEG 400) containing additionally 50 mM of cyano-pyrimidine **3** (racemic), or 5 mM cyano-pyrimidine **5** for 65 h and 30 min, respectively. The Cdu1~Ub complex was crystallized overnight at a protein concentration of 20 mg/ml in 0.1 M MES (pH 6.5) and 12% PEG 20000.

Data collection

All crystals were flash frozen in liquid nitrogen. Diffraction data were collected either at beamline 14-1 (BESSY II, HZB Berlin), P14 (PETRA III, EMBL-Hamburg), or the European Synchrotron Radiation Facility (ESRF, Grenoble), beam line ID23-1. The following parameters were applied for the collection of each data set:

Collection parameter	Apo-Cdu1	Cdu1-SeMet	Cdu1-3	Cdu1-5	Cdu1-Ubi
Beamline	BESSY BL 14.1	DESY BL P14	ESRF ID 23-1	ESRF ID 23-1	ESRF ID 23-1
Wavelength (Å)	0.9184	0.9802	0.9800	0.9762	0.9677
Detector	Dectris PILATUS 6M	EIGER 16M	Dectris PILATUS 6M	Dectris PILATUS 6M	Dectris PILATUS 6M
Detector distance (mm)	323.72	344.08	455.11	411.51	103.58
Number of images	1800	2400	1800	1800	3600
Oscillation (°)	0.10	0.15	0.15	0.10	0.10
Exposure time (s)	0.30	0.30	0.50	0.50	0.30

Structure Determination and Refinement

''Diffraction data of Cdu1_Mut (155-401) were integrated with XDS (124) and scaled with AIMLESS (161). The CRANK2 (162) pipeline was used to solve the structure by the SAD method exploiting the anomalous signal of the diffraction data collected at the Se-peak. Subsequently, the structure of wild type Cdu1_LCMS (155-401) was solved by molecular replacement using Phaser (119). Manual model building was performed in COOT (115) and model refinement was carried out with PHENIX (120). The datasets for the Cdu1~cyano-pyrimidine **3** and Cdu1~cyano-pyrimidine **5** complexes were scaled and integrated using the same tools as described above. The structures were solved by molecular replacement in Phaser (119) using the apo Cdu1 structure (PDB 5B5Q) as search model. The Cdu1~Ub complex was solved using the apo Cdu1 structure (PDB 5B5Q) and Ub (PDB 1UBQ) as search models. Manual model building was performed in COOT (115) and model refinement was carried out with REFMAC (123, 163)''

3 RESULTS AND DISCUSSION

3.1 Construct design

3.1.1 Secondary structure prediction

In order to identify suitable constructs for the structural characterization of Cdu1 a secondary structure prediction was carried out using the Phyre2 Protein Fold Recognition Server (121), which uses homology modeling to provide an approximation of the secondary structure of a given amino acid sequence. According to Phyre2, Cdu1 is composed mainly of alpha helices (49%) and a rather high amount of disordered area (36%). Amino acids 31 to 65 constitute a transmembrane domain (TMD). In addition, amino acids 66 to 125 appear to be disordered towards the N-terminus of the protein (Figure 3.1 A and B).

The first Cdu1 fragment selected for cloning was a 65 amino acid deletion towards the N-terminus which included the TMD, this construct was named Cdu1_L. Since transmembrane domains are rich in amino acids with hydrophobic side chains, they can cause solubility issues and promote precipitation. The second fragment was a 125 amino acid deletion, thus excluding the TMD and a 60 amino acids long unstructured area (Cdu1_S). Disordered regions may contain exposed hydrophobic side chains directed towards the solvent and may lead to similar solubility issues as described for the TMD´s and inhibit crystal packing due to the lack of a defined structure.

3.1.2 In situ proteolysis and LCMS analysis

In order to identify stable Cdu1 constructs, *in situ* proteolysis was performed. A stock of purified Cdu1_S (Section 3.2.2), was treated with four different proteases and evaluated at diverse time points. The formation of a
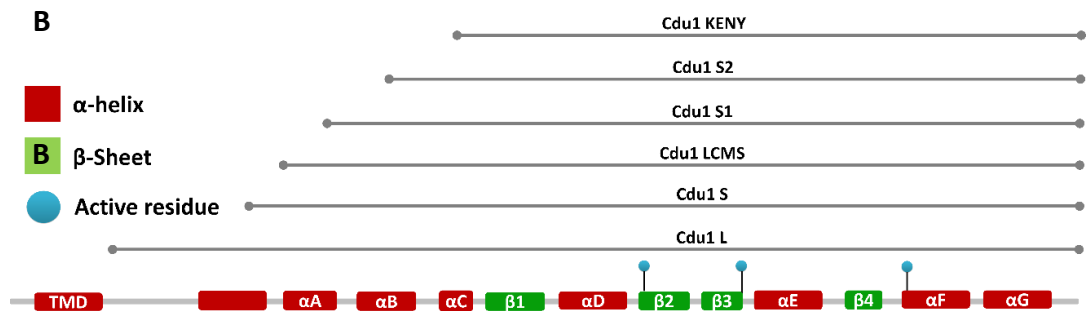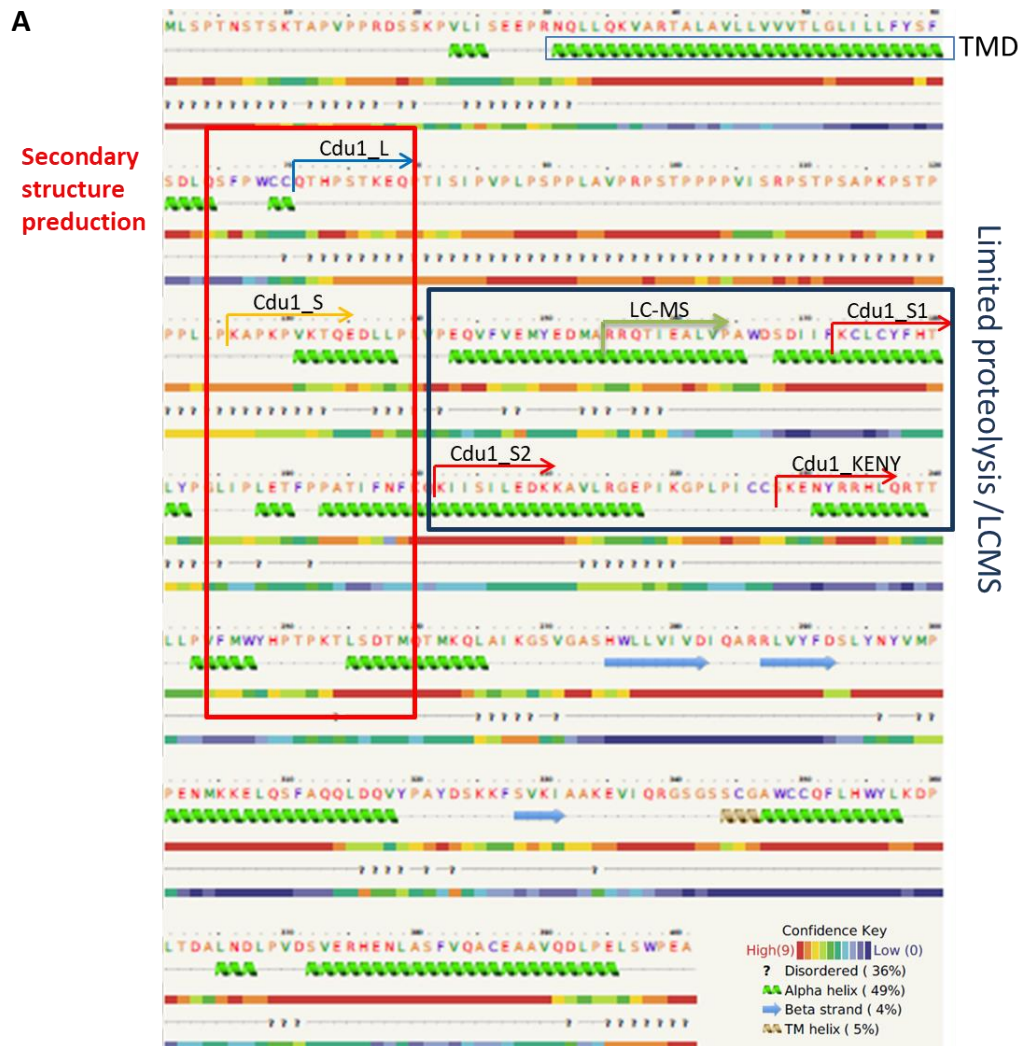
Figure 3.1. (A) Cdu1 fragments selected for cloning based on secondary structure prediction (Phyre2, red) and limited proteolysis/LCMS experiments (Blue). **(B)** Domain architecture of Cdu1 cloned constructs.

proteolytic product was already observed after 30 min. After two hours, a semi-stable band shortened by 20 to 50 amino acids was observed in the presence of trypsin and subtilisin. α-Chymotrypsin showed little proteolytic effect, in contrast to papain which completely digested the protein (Figure 3.2).
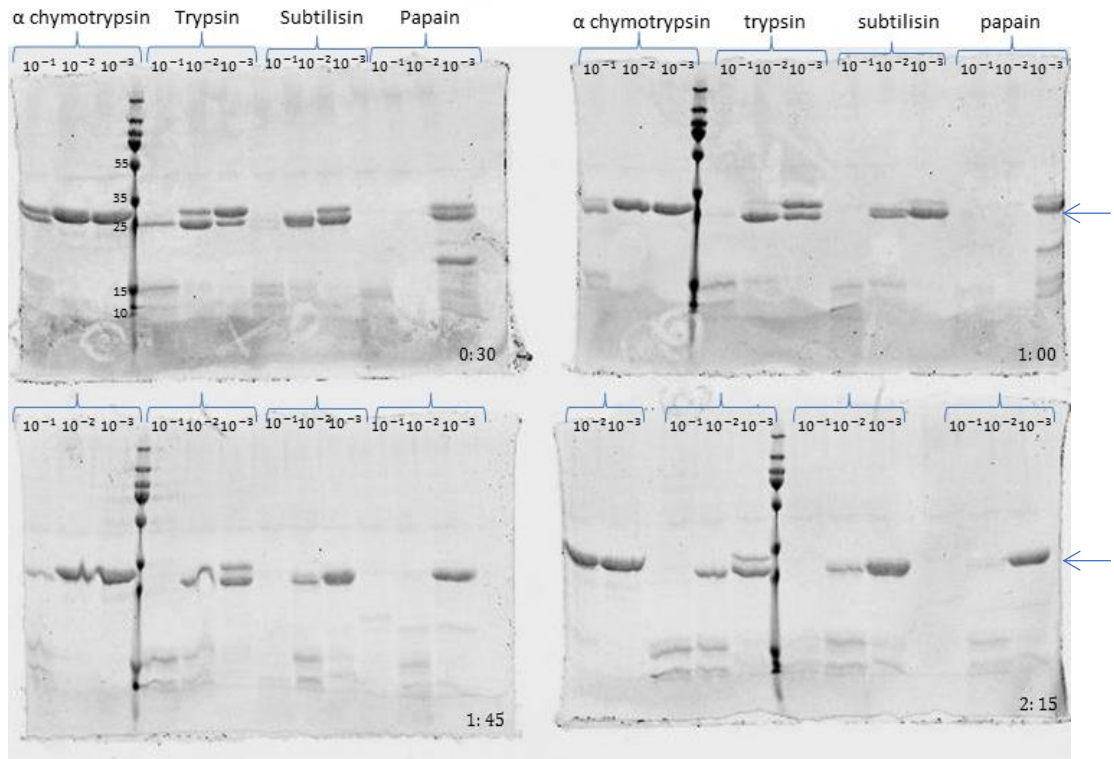
Figure 3.2 *In situ* proteolysis of the Cdu1_S fragment. Digestion was evaluated by SDS-PAGE after 30 min, 1 h, 1:45 h and 2:15 h. Protease dilutions (10-1 to 10-3) were made from a 1 mg /ml stock.

The band marked by a blue arrow in Figure 3.2 was cut out from the gel and sent for mass spectrometry analysis. In this experiment, the band was further digested with trypsin to generate peptides that could be detected by the instrument. The different peptides were identified based on their masses and the sequence of Cdu1. By comparing the peptides detected from the proteolytic product fragment

and an untreated control, it was possible to estimate the location of the tryptic cutting site on the proteolysis resistant fragment.

Figure 3.3 shows the summation of peptide intensities plotted against the amino acid position. The red line corresponds to the "full" non-treated Cdu1_S fragment, while the blue line is the proteolytic product. In this graph, it is evident that the population of peptides N-terminal to amino acid 51 is significantly lower than in the intact Cdu1_S fragment. Therefore, four tryptic sites within this region were selected as the N-terminus for new Cdu1 constructs to be tested for crystallization, namely Cdu1_LCMS, Cdu1 S1, Cdu1_S2 and Cdu1_KENY (Figure 3.1, blue box)
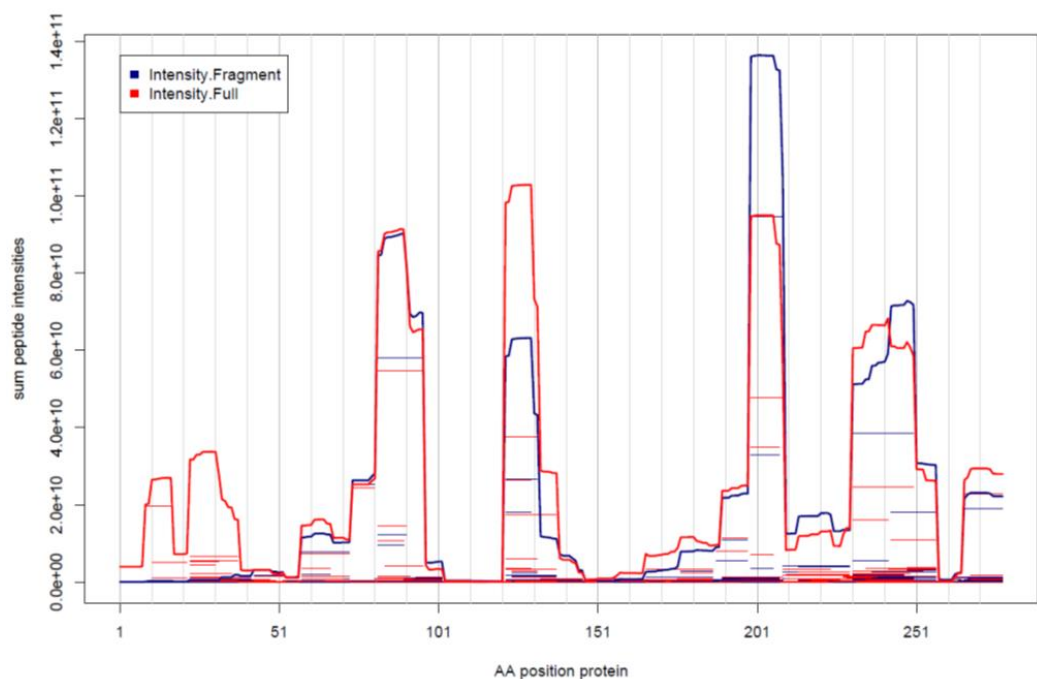
Figure 3.3. Summation of peptide intensities for the *in situ* proteolysis product (blue) and an untreated control (red), plotted against their amino acid position.

3.2 Cloning, expression and purification

3.2.1 Cloning of Cdu1 constructs and expression tests

All Cdu1 constructs (Figure 3.4A) were generated using SLIC cloning (see Materials and Methods, section 2.2.1). Positive clones were verified by colony PCR and sequencing (see appendix, Figure 6.3). Expression tests revealed prominent overexpression bands for Cdu1_L and Cdu1_S, especially for the Cdu1_S_Smt3_pETM14 constructs. Intense precipitation for the S1, S2 and KENY constructs was observed, with only weak overexpression in a soluble form for S1 and KENY. The LCMS construct showed a more prominent overexpression with less protein in the pellet (Figure 3.4B).

A Construct	Nr	Amino acid range	Average yield (mg/l culture)
Hys6 · 3C · Cdu1_L · BamHI · pETM14	1	65-401	0.1
Hys6 · 3C · Cdu1_S · BamHI · pETM14	2	125-401	4.0
TRX · Hys6 · 3C · Cdu1_L · BamHI · pCDF	3	65-401	0.1
Hys6 · SMT3 · 3C · Cdu1_S · BamHI · pETM14	4	125-401	4.0
Hys6 · 3C · Cdu1_S1 · BamHI · pETM14	5	173-401	0.1
Hys6 · 3C · Cdu1_S2 · BamHI · pETM14	6	202-401	Not soluble
Hys6 · 3C · Cdu1_KENY · BamHI · pETM14	7	229-401	0.1
Hys6 · 3C · Cdu1_LCMS · BamHI · pETM14	8	155-401	5.0
Hys6 · 3C · Cdu1_Mut · BamHI · pETM14	9	155-401	10.0

➔ See legend on next page
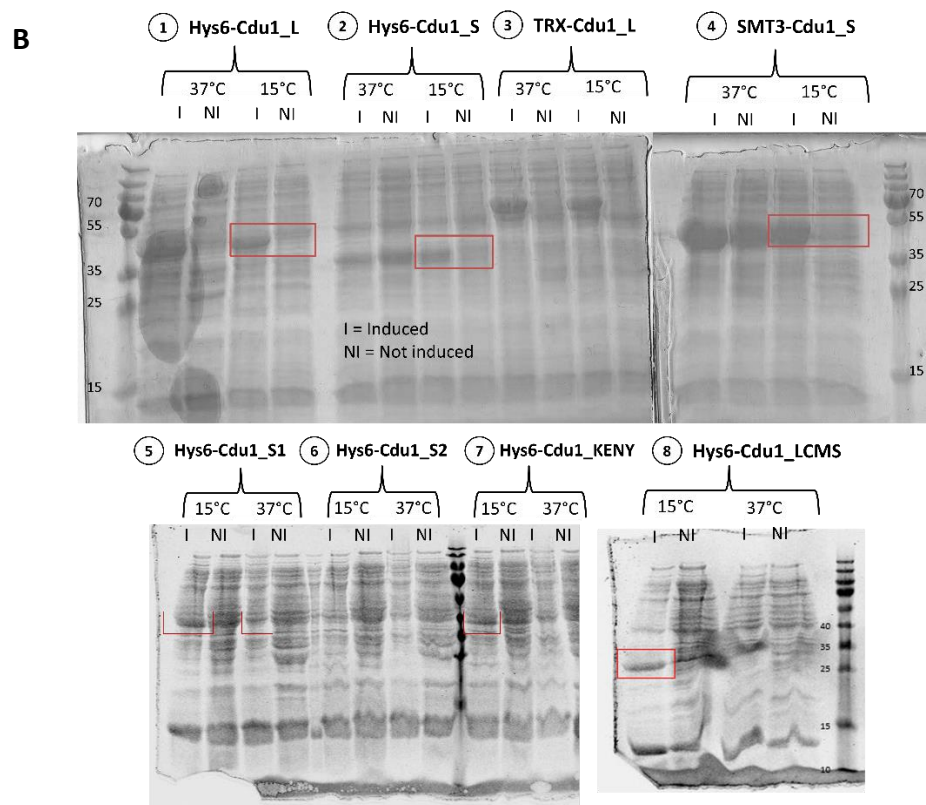
Figure 3.4. (A) Overview of the cloned Cdu1 constructs. The affinity tags (His6, SMT3 or TRX) can be all cleaved off by the precision protease (3C). **(B)** Expression test for the constructs depicted in **A**. Clear overexpression bands can be observed on the induced (I) lanes compared to the non-induced (NI) lanes, especially for the bacteria that were grown at 15°C.

3.2.2 Purification by IMAC and SEC

All Cdu1 constructs were expressed using BL21 DE3 STAR cells (*E. coli*) and purified using the strategies described in section 2.2.3. After IMAC, fractions containing the Cdu1_S and Cdu1_L fragments (Figure 3.5, A and C) were dialyzed overnight while the tag cleaving reaction was taking place. Pure protein fragments (Figure 3.5, B and D) were concentrated until a volume of 5 ml was achieved in preparation for SEC. However, during this concentration step, a precipitate was observed for the Cdu1_L fragment possibly because of the presence of the unstructured region towards the N-terminus. The precipitate was not observed

for the Cdu1_S fragment. After centrifugation at 20 min at 25,000 x g and 4 ∘C, the samples were applied to a SD 75 16/60 SEC column.

The SEC chromatogram for Cdu1_S (Figure 3.5 E) revealed a main peak eluting at approximately 0.5 CVs with two shoulders, probably representing protein aggregates. Only the fractions belonging to the main peak were pooled (represented in brackets in Figure 3.5 E). A final yield of 4.0 mg per littler of cell culture for the Cdu1_S fragment and 0.1 mg per littler of culture for the Cdu1_L fragment, was achieved. During the final concentration step, intense precipitation was again observed for the Cdu1_L fragment. Therefore, the more stable Cdu1_S construct was preferentially used for subsequent crystallization experiments. A similar precipitation effect as for the Cdu1_L construct was observed when purifying the S1 and KENY constructs, in the same conditions above described. This led to a minimal yield as shown in Figure 3.6 A and B. However, similarly to the Cdu1_S construct, sufficient amounts of the Cdu1_LCMS construct could be obtained after IMAC (Figure 3.6 C). After SEC, a final yield of approximately 5.0 mg per litter of culture was achieved for the Cdu1_LCMS construct (Figure 3.6 D). The final fractions were pooled and concentrated to 10 mg/ml using and Amicon centrifugal concentrator with a MW cutoff of 10000 Da.
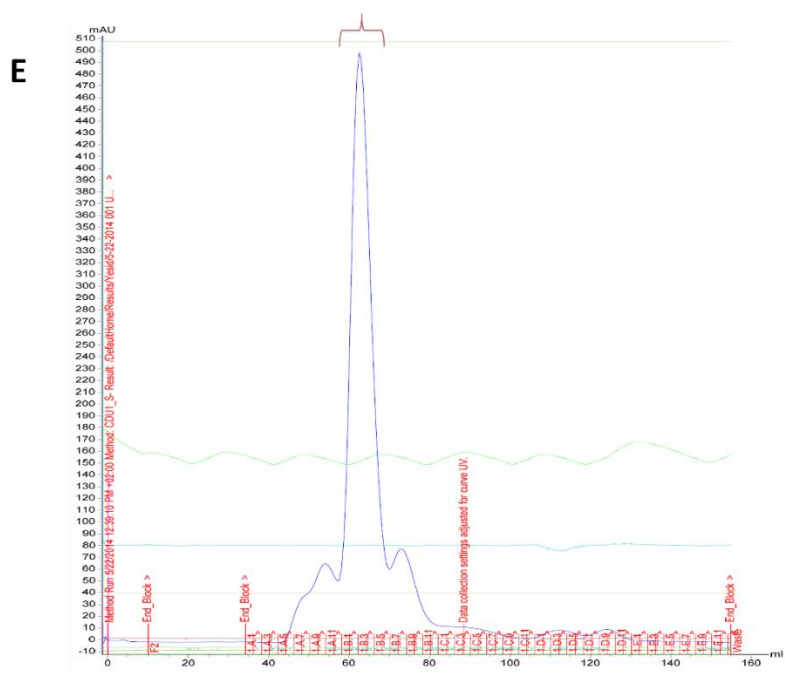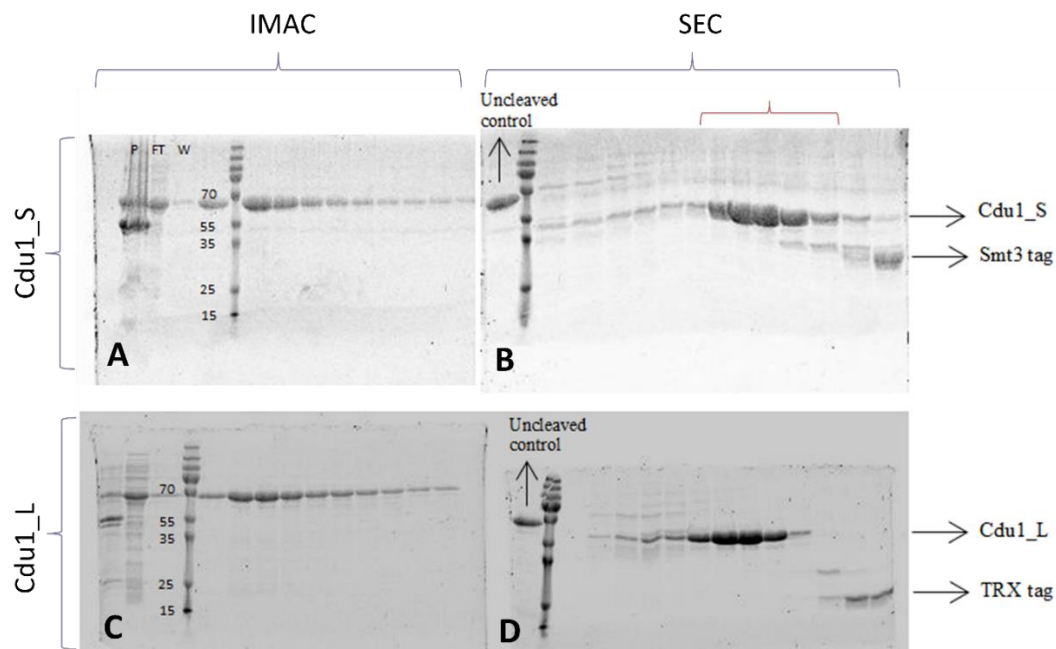
Figure 3.5. **(A)** SDS PAGE after IMAC of the Cdu1_S fragment. Pellet (P), Flow through (FT), Wash (W) and 10 elution fractions. **(B)** SDS PAGE after SEC of Cdu1_S after the overnight tag cleavage reaction. **(C)** SDS PAGE after IMAC of the Cdu1_L fragment. Pellet (P), Flow through (FT), Wash (W) and 10 elution fractions. **(D)** SDS PAGE after SEC of Cdu1_L after the overnight tag cleavage reaction. **(E)** Typical SEC chromatogram of a Cdu1 purification. The Cdu1_S fragment eluted at approx. 0.5 CV on a Superdex 75 16/60 column.
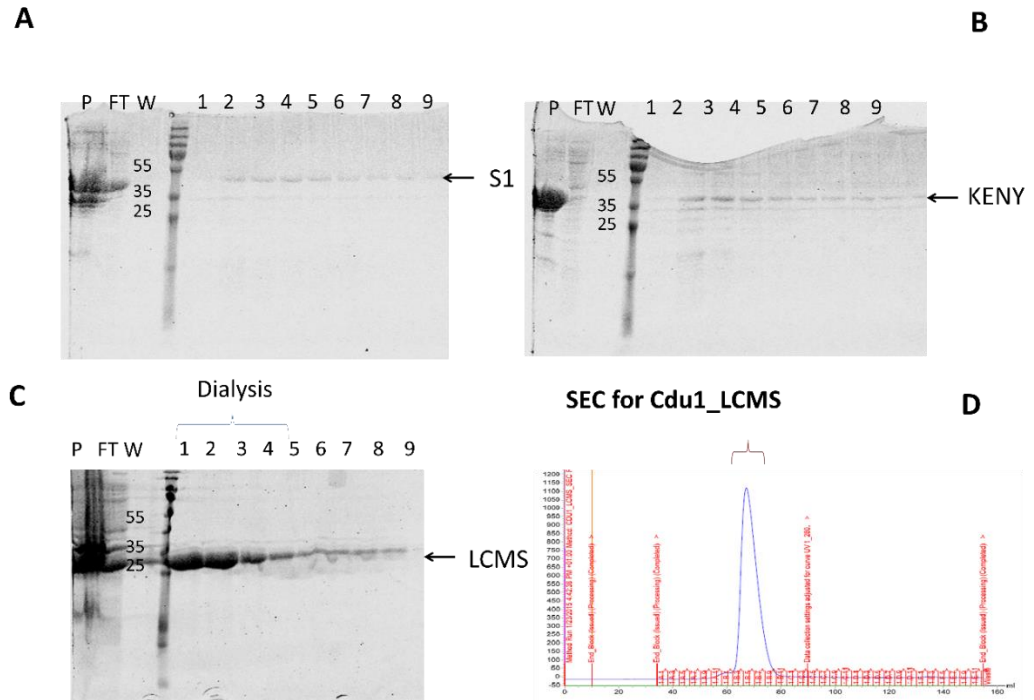
Figure 3.6. **(A)** (a) SDS PAGE after IMAC of the Cdu1_S1 fragment. Pellet (P), Flow through (FT), Wash (W) and 10 elution fractions. **(B)** SDS PAGE after IMAC of the Cdu1_KENY fragment. **(C)** SDS PAGE after IMAC of the Cdu1_LCMS. **(D)** Size exclusion chromatogram of the Cdu1_LCMS construct. Elution occurs around 0.5 CV on a Superdex 75 16/60 column, which is similar to previous Cdu1 constructs.

3.3 Obtaining a Cdu1 oxidation resistant variant

We speculated that the surface exposed cysteines may oxidize leading to the formation of inter-molecular disulfide bridges and thus to the formation of heterogeneous aggregates, thereby impairing crystallization. Wildtype Cdu1 was therefore assessed via Multi Angle Light Scattering (MALS).

3.3.1 MALS experiments on Cdu1

A batch of Cdu1_LCMS was purified with a 5 mM β-mercaptoethanol (BME) supplemented lysis buffer and a 2 mM dithiothreitol (DTT) supplemented

gel filtration buffer as reducing agents. This batch was analyzed via SEC- MALS, along with a Cdu1 batch that contained no reducing agent as control.
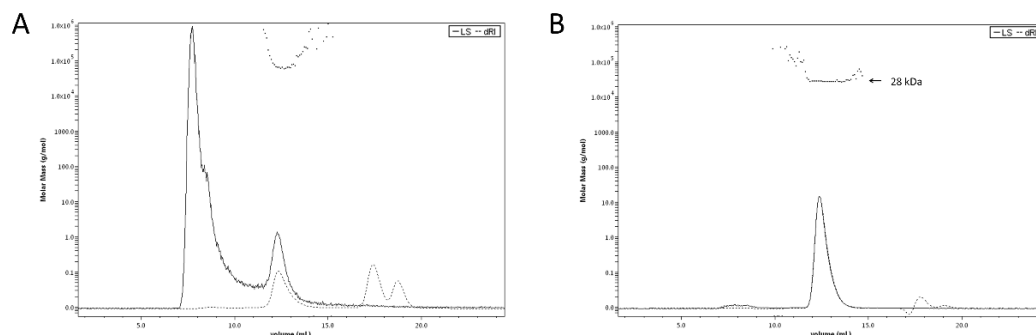
Figure 3.7. (A) SEC-MALS analysis of Cdu1-LCMS (155-401) in the absence of reducing agent. The prominent light scattering signal (solid line) at low elution volumes indicates sample aggregation. The dynamic refractive index signal (dashed line) indicates that Cdu1_LCMS (155-401) elutes at a retention volume of around 12 ml **(B)** SEC-MALS analysis of Cdu1_LCMS (155-401) in the presence of 2 mM DTT. The absence of a prominent light scattering signal at low elution volumes indicates increased sample homogeneity

SEC-MALS experiments showed a prominent light scattering signal at low elution volumes in the absence of a reducing agent, where the presence of aggregates leads to an overestimation of the molecular weight (Figure 3.7 A). In the presence of a reducing agent (2 mM DTT), the absence of such a signal at low elution volumes indicated better sample homogeneity (Figure 3.7 B). The determined mass of 28 kDa is consistent with the calculated size of a monomeric protein. Although Pruneda *et al.* demonstrated that native Cdu1 (130-401) crystals could be obtained in the presence of reducing agent (5 mM DTT), I was not able to obtain crystals in the presence of 4 mM DTT for the Cdu1_S construct (125-401). To prevent aggregation of Cdu1 due to oxidation of the cysteine residues and to obtain protein crystals, the Cdu11_Mut (155-401) variant was prepared.

3.3.2 Site directed mutagenesis on selected Cdu1 cysteines

In order to identify which cysteines could most likely be mutated without affecting the stability of the protein, an alignment of the CT Cdu1 sequence along with the sequences of seven related proteases was performed. The alignment included Cdu1 from other *Chlamydia* species. Only the non-conserved cysteines, apart from the catalytic cysteine, were mutated (Figure 3.8).
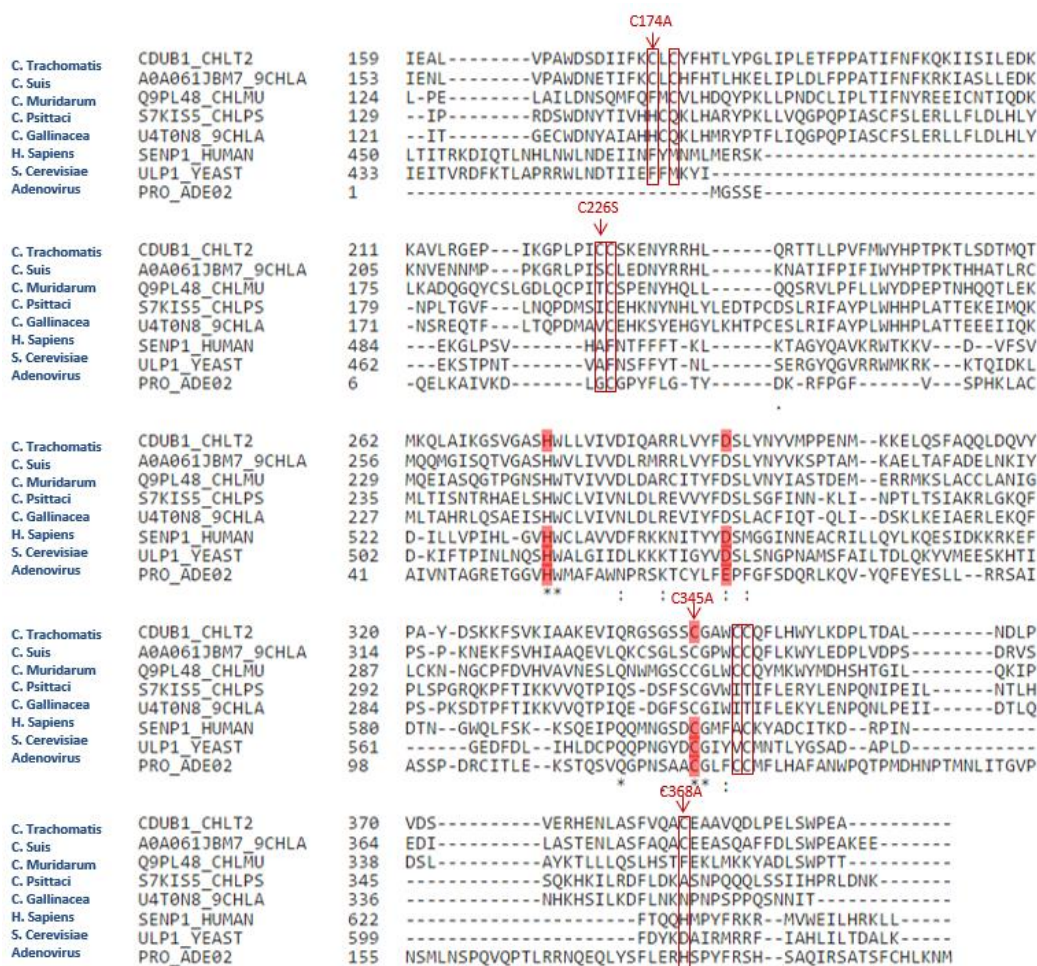
Figure 3.8. Sequence alignment of the protease domain of CT Cdu1 with related proteases. Point mutations within the CT construct are indicated in red above the sequence.

The non-conserved cysteines at positions 174 and 368 were mutated to alanine since no preference was observed for a specific amino acid in this position across

the CT Cdu1 related proteases. Cysteine 226 was replaced by a serine, because of the presence of this residue in the *C. Suis* (CS) Cdu1 protein and the active site cysteine C345 was mutated to alanine, a mutation reported to catalytically inactivate the protein without affecting its solubility (67, 90). Cysteines 176, 349 and 350 were left intact because they are conserved across the CT, CS and CM (*Chlamydia Muridarum)* genomes.

The resulting Cdu1_Mut construct was generated as described in section 2.2.1 and, after mutagenesis, positive clones were validated through sequencing (appendix, Figure 6.4). Expression and purification were performed utilizing the same methods as described in section 3.2.2. A final yield of 10 mg per liter of culture was achieved (Figure 3.9).
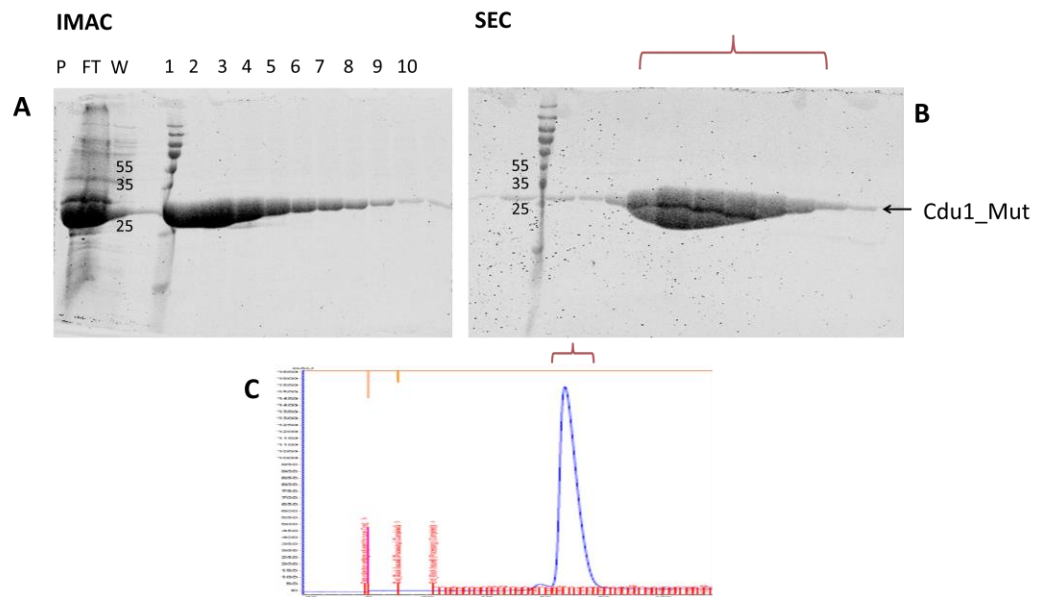
Figure 3.9. (A) SDS PAGE after IMAC of the Cdu1_Mut construct (155-401). Pellet (P), Flow through (FT), Wash (W) and 10 elution fractions. **(B)** SDS PAGE after SEC of the Cdu1_Mut construct. **(C)** Size exclusion chromatogram of the Cdu1_Mut construct on a Superdex 75 16/60 column. Fractions representing the peak in the SEC chromatogram are marked by brackets.

3.4   Crystallization and structural analysis of Apo-Cdu1

At the beginning of this work, no structural information about Cdu1 was available. The structure of Cdu1 promised to provide value insights into the molecular basis of Cdu1 function and would enable structure-based optimization of inhibitors for this promising drug target. Therefore, an extensive search for crystallization conditions of Cdu1 was pursued.

3.4.1   Crystallization

As an initial step towards suitable crystallization conditions, Sparse Matrix Sampling was performed for all Cdu1 constructs which could be purified in sufficient soluble form (Section 3.2). Vapor diffusion crystallization experiments were carried out with a crystallization robot. Each crystallization screen was performed at 4°C and 20°C for the Cdu1_L, Cdu1_S, and Cdu1_LCMS construct. However, no useful lead conditions were observed for any of the constructs, with either protein precipitation, phase separation or clear drops being prevalent even in the presence of reducing agent (4 mM DTT). To overcome this issue, the Cdu1_Mut variant was prepared as described in section 3.3.

An initial stock of Cdu1_Mut at a concentration of 10 mg/mL was used as input for Sparse Matrix Sampling using all the standard crystallization screens available at our lab (page 28, Section 2.1.1). Although this preliminary approach did not yield crystals, it was observed that the drop turned "oily" in the presence of 100 mM MES pH 6.5, 12% PEG 20000 (Crystal screen I+II, F11) and with 100 mM Bicine pH 9.0, 10% PEG 20000, 2% 1, 4 dioxane (Crystal screen I+II, H12). This oily state could be a hint that the protein concentration should be increased in order to achieve crystallization. Therefore, the Cdu1_Mut stock was further concentrated to 20 mg/mL and the complete screen was repeated. At this point, the first Cdu1 crystals were obtained in the conditions mentioned above (Figure 3.10 A and B). Fine screens were subsequently performed with the 24-well plate

format. Crystals obtained in the Crystal screen I+II, condition H12, could be reproduced (Figure 3.10 C) whereas the one obtained initially in condition F11 did not yield crystals.
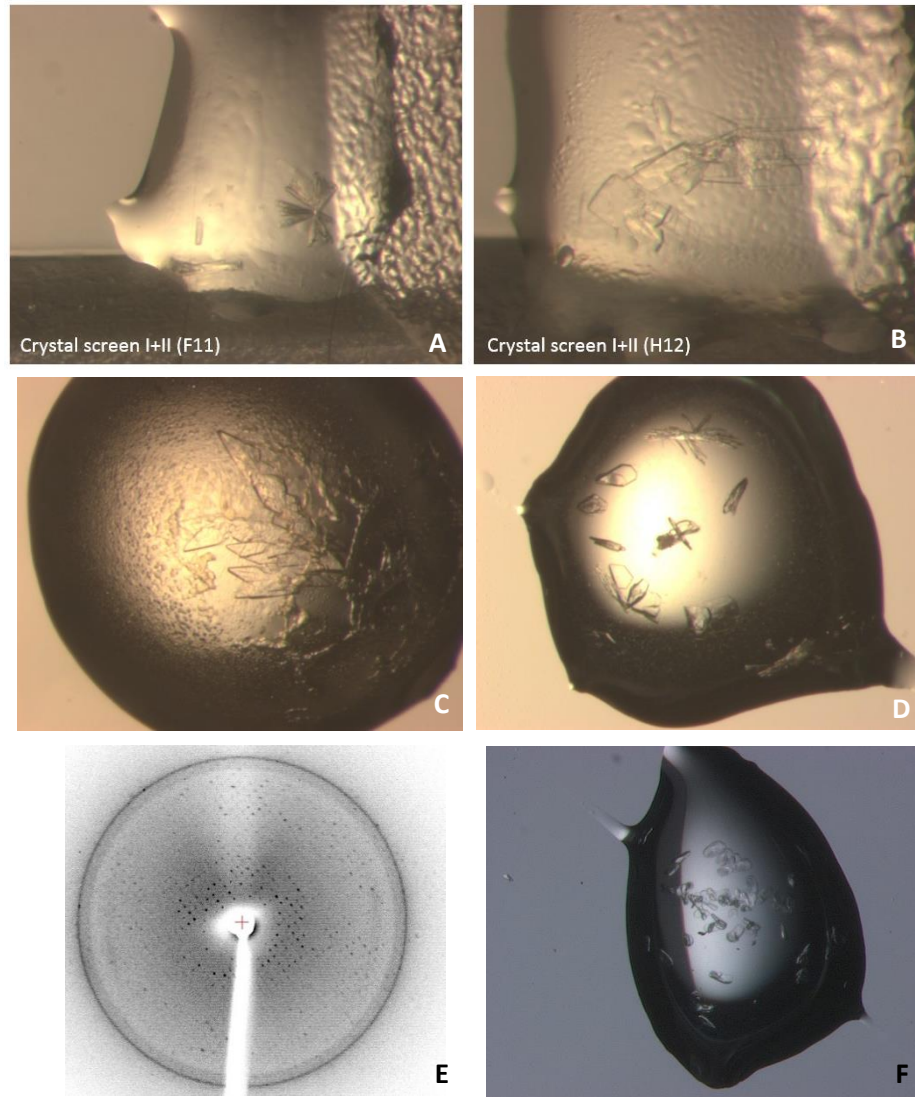
Figure 3.10. (A) Crystals of the Cdu1_Mut construct grown overnight at 20°C (100 mM MES pH 6.5, 12% PEG 20000). **(B)** Plate-like crystals of Cdu1_Mut grown overnight at 20°C (100 mM Bicine pH 9.0 + 10% PEG 20000 and 2% 1, 4 dioxane). **(C)** Fine screen on Cdu1_Total_Mut at 20°C (100 mM Bicine pH 9.0 + 10% PEG 20000 and 2% 1, 4 dioxane). **(D)** Single Cdu1_LCMS crystals obtained by serial seeding at 4°C. **(E)** Diffraction pattern of a Cdu1_LCMS crystal. Preliminary data revealed a maximum resolution of 2.3 Å recorded at the RVZ X-ray source. **(F)** Cdu1_LCMS_SeMet crystals obtained by using Cdu1_LCMS crystals as seeds 4°C.

To obtain crystals from the Native Cdu1_LCMS construct, a new fine screen was performed and the recently obtained Cdu1_Mut crystals were used as seeds (see Materials and Methods, section 2.2.6). This approach produced needle like Cdu1_LCMS crystals that were then used for a new round of serial seeding. After 3 rounds, plate like crystals were obtained (Figure 3.10 D). All crystals were flash frozen in liquid nitrogen using the mother liquor containing additionally 40% PEG 400, as cryo protectant and their ability to produce a diffraction pattern was evaluated at the RVZ X-ray home source. Diffraction up to 2.3 Å was observed at the edge of the detector (Figure 3.10 E).

Spontaneous nucleation is an essential process that needs to take place prior to crystallization, which also requires higher concentrations than crystallization itself (160). However, at this concentration range (10 mg/mL) Cdu1 aggregates were also forming most likely due to cysteine oxidation, thereby imparting crystallization. The cysteine variant permitted a doubling of the protein concentration, generating the appropriate conditions for nucleation to occur without the formation of aggregates. Using the crystals of the variant as nucleation points through seeding, permitted the use of lower concentrations of wild-type Cdu1 for crystallization where aggregates do not tend to form, but also where spontaneous nucleation would otherwise not occur.

3.4.2 Structure elucidation

Cdu1_LCMS was overexpressed in a M9-Minimal-Medium supplemented with an amino acid mix containing Selenium-methionine (Se-Met). During expression of a selenium-methionine derivative, a Se-Met residue is incorporated in the protein instead of the usual Met residue. The protein was purified by the methods described in section 3.2.2 and crystals were readily obtained by using Cdu1_LCMS crystals as seeds (Figure 3.10 F).

Data Collection

	Cdu1 (155-401)	Cdu1 (155-401) _SeMet[a]
Crystal		
Space Group	$P2_1$	$P2_1$
Unit Cell Parameters		
a, b, c (Å)	43.67, 77.24, 67.65	39.76, 78.33, 69.37
ß (°)	107.28	96.18
Resolution Limits (Å)	77.24 – 1.65 (1.68 – 1.65)	78.33 – 1.90 (1.94 – 1.90)
R_{pim} (%)	5.3 (47.4)	5.9 (56.3)
$CC_{1/2}$ (%)	99.3 (56.4)	99.8 (72.6)
Multiplicity	5.5 (5.4)	3.5 (3.6)
Unique Reflections	50612 (2543)	32833 (2093)
Completeness	98.1 (98.3)	97.2 (96.4)
$<I/\sigma I>$	12.4 (1.8)	12.4 (2.3)
Phase Determination		
No. of sites		10
FOM after density modification (PARROT)		0.423
Refinement		
R_{work} (%)	16.5	
R_{free} (%)	21.3	
RMSD Bond Lengths (Å)	0.004	
RMSD Bond Angles (°)	0.732	
Ramachandran statistics (%)	97.7 / 2.1 / 0.2	

Table 3.1: Data Collection and Refinement Statistics for Cdu1 (155-401) and Cdu1_SeMet. Data for the highest resolution shell are given in parentheses.
[a] Bijvoet pairs were kept separate for processing
Taken from Fischer and Ramirez et al. 2017 with permission.

Crystals from the native protein (Cdu1_LCMS), and the selenomethionine derivative (Cdu1_SeMet) were flash frozen and taken to the synchrotron radiation facility DESY in Hamburg. Diffraction data for the native data set were recorded up to 1.7 Å, whilst the Se-Met derivative crystals diffracted up to 1.8 Å. In preparation for crystal soaking and co-crystallization experiments, new Cdu1_LCMS crystals were generated in the conditions described in section 3.4.1 containing additional 5% dimethyl sulfoxide (DMSO) in the mother liquor. These crystals were subsequently taken to BESSY in Berlin, where a 1.7 Å dataset was collected and the structure of Cdu1 (155-401) was solved using molecular

replacement based on the Se-Met derivative. Data collection and refinement statistics are summarized in Table 3.1.

3.4.3 The unique structural features of Cdu1 among the deubiquitinylase family

Sections in between " " are taken from Fischer and Ramirez et al. 2017 with permission

"The Cdu1 (155-401) fragment crystallized in space group $P2_1$ with two molecules in the asymmetric unit and at a resolution of 1.7 Å. Each monomer is composed of a central four-stranded β-sheet, which is surrounded by eight α-helixes resembling the architecture of the C48 proteases (Figure 3.11 A). The catalytic cysteine, Cys345, is located at the N-terminus of α-helix F, which is 'sandwiched' in between a set of three anti-parallel β-strands (1,2,3) and two extra α-helices, A and G, which are perpendicular to each other. β-strands 2 and 3 also contain the two other residues of the catalytic triad, His275 and Asp292 (Figure 3.11 B). The structure of Cdu1 was also solved by Pruneda et al. at a resolution of 2.1 Å (67). Their fragment consists of residues 130 to 401 and can be superimposed with our Cdu1 structure with an rms deviation of 0.563 Å including residues 161 to 401 using the program LSQ superpose in COOT (164). Due to the difference in length of the construct the structure of Pruneda et al. contains an additional α-helix at the N-terminus which is not present in our Cdu1 structure".

The N-terminus of Cdu1 is unstructured up to residue 167, providing a possible explanation why the Cdu1_L construct (65-401) and the Cdu1_S (125-401) failed to initially crystallize, as they contained a longer portion of this disordered area. Since the Cdu1_S1, Cdu1_S2 and Cdu1_KENY constructs were starting at residues 173, 202 and 229 respectively, the secondary structure was being disturbed, which could explain why these constructs were difficult to purify in a soluble form. As demonstrated by Pruneda et al., a truncated version of Cdu1 only five amino acids

shorter than Cdu1_S can be crystallized without the seeding approach described in section 3.4.1. However, the absence of the α-helix at the N-terminus of the Cdu1_LCMS construct may have had a positive impact in terms of crystal packing, leading to the superior resolution of the Cdu1 model described herein this thesis (1.7 Å) as compared to Pruneda's model (2.1 Å).
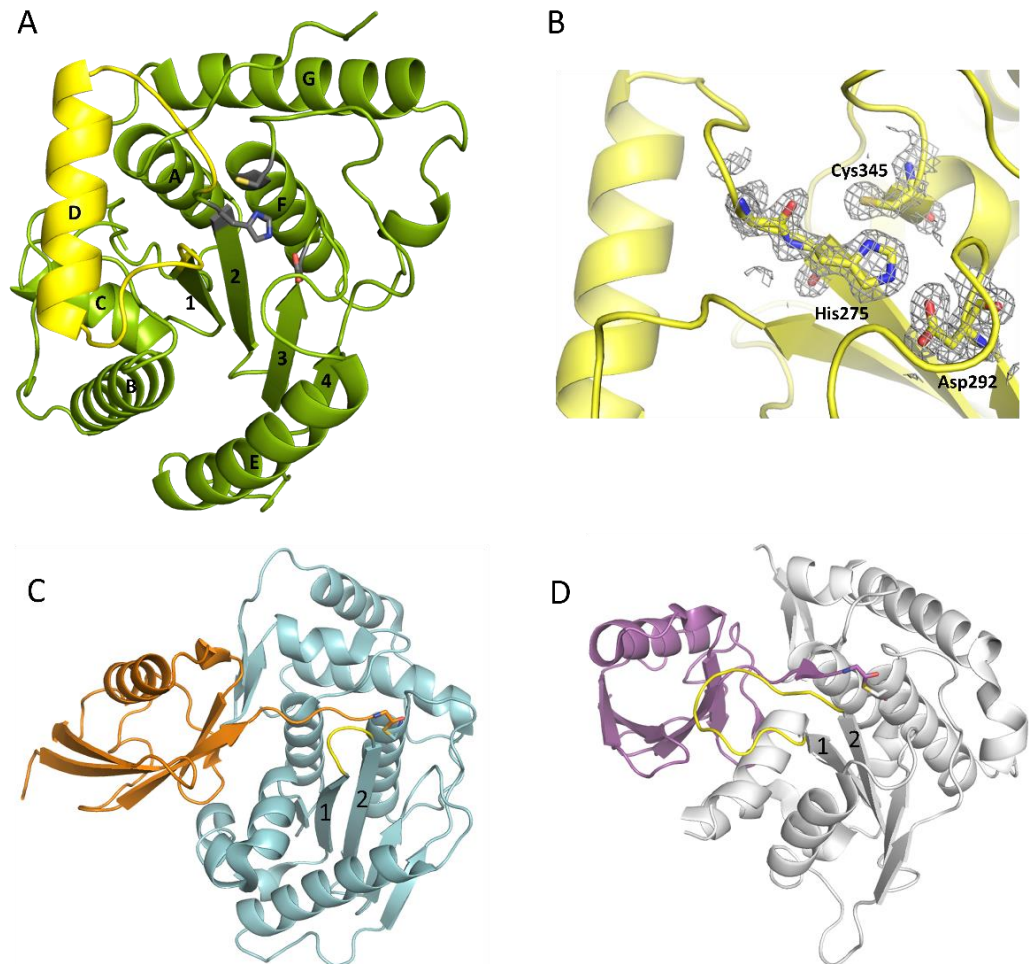
Figure 3.11. (A) Overall structure of Cdu1 showing the catalytic triad (grey C atoms shown in stick representation) and the Lid helix motif between β-strands 1 and 2 (yellow). **(B)** $2F_0 - F_C$ omit maps for the catalytic triad of Cdu1 contoured at a sigma level of 1.0. **(C)** Structure of the Ulp1 (cyan)-SMT3 (orange) complex. The loop between β-strands 1 and 2 is shown in yellow (PDB1EUV). **(D)** Structure of the SENP8 (gray) – Nedd8 (purple) complex. The loop between β-strands 1 and 2 is shown in yellow. (PDB:1XT9). Taken from Fischer and Ramirez et al. 2017 with permission.

"A search for other related structures utilizing the program DALI (116), indicated that Cdu1 (155-401) shares a high degree of structural similarity with Ulp1 (103), despite the fact that they only share 17% sequence identity within a stretch of 221 amino acids. The crystal structure of Ulp1 in complex with the ubiquitin like protein SMT3 (Figure 3.11 C) reveals the direction in which the C-terminal glycine of SMT3 approaches the active site cysteine. This path seems to be conserved between the two enzymes. Remarkably, however, a small loop between β-strands 1 and 2 (shown in yellow) in Ulp1 contrasts with the much bigger α-helix D in Cdu1 (Figure 3.11 A) in between the equivalent β-strands of both structures.

This region corresponds to VR-3 within the CE fold (67). The same phenomenon can also be observed when comparing Cdu1 with the human SENP8 protein in complex with a Nedd8 aldehyde (165), with whom Cdu1 shares 11% identity within 208 amino acids. Although larger than in Ulp1, VR-3 of SEMP8 fails to adopt the helical architecture observed in Cdu1 (Figure 1.9 and 3.11 D). The additional α-helix D in Cdu1, a feature not observed in the other structures may play a regulatory role in substrate binding or recognition. This hypothesis is further strengthened by the analysis of Pruneda et al. where deletion of this region (residues 250 to 272) prevented Cdu1 from binding ubiquitin/Nedd8 suicide probes (67).

Importantly, an insertion of identical length as observed in Cdu1 seems to be present in the Cdu1 proteins encoded by the genomes of *C. suis, C. muridarum, C. psittaci, C. gallinaceai* and Cdu2 from *C. trachomatis* suggesting that it is a conserved feature across the chlamydial deubiquitinases (Figure 3.12)"

```
CDU1 C. Trachomatis    CDUB1_CHLT2       159  IEALVPAWDSDIIFKCLCYFHTLYPGLIPLETFPPATIFNFKQKIISILED---KKAVLR
CDU1 C. Suis           A0A061JBM7_9CHLA  153  IENLVPAWDNETIFKCLCHFHTLHKELIPLDLFPPATIFNFKRKIASLLED---KKNVEN
CDU1 C. Muridarum      Q9PL48_CHLMU      123  RLPELAILDNSQMFQFMCVLHDQYPKLLPNDCLIPLTIFNYREEICNTIQD---KLKADQ
CDU1 C. Psittaci       S7KIS5_CHLPS      127  DLIPRDSWDNYTIVHHCQKLHARYPKLLVQGPQPIASCFSLERLLFLDLHLYNPLTGVFL
CDU1 C. Gallinaceai    U4T0N8_9CHLA      119  SPITGECWDNYAIAHHCQKLHMRYPTFLIQGPQPIASCFSLERLLFLDLHLYNSREQTFL
CDU2 C. Trachomatis    CDUB2_CHLT2        85  TGPPLPIWDNEHLFQFSCLMQNKHRRVLPIDICNPLTKFNFLECICNCLMT---KQSVNV
                                              *.   :  :      ::  :  .:       : *  .  :: :  : ::
                                                                                    Insertion between β strands 1 and 2
CDU1 C. Trachomatis    CDUB1_CHLT2       216  GEP-I--KGPLPICCSKENYRRHL------QRTTLLPVFMWYHPTPKTLSDTMQTMKQLA
CDU1 C. Suis           A0A061JBM7_9CHLA  210  NMP-P--KGRLPISCLEDNYRRHL------KNATIFPIFIWYHPTPKTHHATLRCMQQMG
CDU1 C. Muridarum      Q9PL48_CHLMU      180  GQYCSLGDLQCPITCSPENYHQLL------QQSRVLPFLLWYDPEPTNHQQTLEKMQEIA
CDU1 C. Psittaci       S7KIS5_CHLPS      187  NQP-------DMSICEHKNYNHLYLEDTPCDSLRIFAYPLWHHPLATTEKEIMQKMLTIS
CDU1 C. Gallinaceai    U4T0N8_9CHLA      179  TQP-------DMAVCEHKSYEHGYLKHTPCESLRIFAYPLWHHPLATTEEEIIQKMLTAH
CDU2 C. Trachomatis    CDUB2_CHLT2       142  N-ETDMCELFCPPTCTPENYRRLL------CTSSVFPFVMWHDPSADTQEAMLTKMDQTM
                                              *   ..*.:                ::    :*:.*   .  :  *

CDU1 C. Trachomatis    CDUB1_CHLT2       267  IKGSVGASHWLLVIVDIQARRLVYFDSLYNYVMPPENMKKELQSFAQQLDQVYPAYD---
CDU1 C. Suis           A0A061JBM7_9CHLA  261  ISQTVGASHWVLIVVDLRMRRLVYFDSLYNYVKSPTAMKAELTAFADELNKIYPSPK---
CDU1 C. Muridarum      Q9PL48_CHLMU      234  SQGTPGNSHWTVIVVDLDARCITYFDSLVNYIASTDEMERRMKSLACCLANIGLCKN---
CDU1 C. Psittaci       S7KIS5_CHLPS      240  NTRHAELSHWCLVIVNLDLREVVYFDSLSGFINN-KLINPTLTSIAKRLGKQFPLSPG--
CDU1 C. Gallinaceai    U4T0N8_9CHLA      232  RLQSAEISHWCLVIVNLDLREVIYFDSLACFIQT-QLIDSKLKEIAERLEKQFPS-PK--
CDU2 C. Trachomatis    CDUB2_CHLT2       195  SSGRVGNSHWVLVIVDIEYRCVTFFDSLCDYVASPQQMREQLEGLAVSLGAIYPKEGGAD
                                              *  ::*:  * : :* ** :       :   :  :* *

CDU1 C. Trachomatis    CDUB1_CHLT2       324  ------SKKFSVKIAAKEVIQRGSGSSCGAWCCQFLHWYLKDPLTDALNDLPVDSVERHE
CDU1 C. Suis           A0A061JBM7_9CHLA  318  ------NEKFSVHIAAQEVLQKCSGLSCGPWCCQFLKWYLEDPLVDPSDRVSEDILASTE
CDU1 C. Muridarum      Q9PL48_CHLMU      291  -----NGCPFDVHVAVNESLQNWMGSCCGLWCCQYMKWYMDHSHTGILQKIPDSLAYKTL
CDU1 C. Psittaci       S7KIS5_CHLPS      297  -----RQKPFTIKKVVQTPIQS-DSFSCGVWITIFLERYLENPQNIPE-ILNTLHSQKHK
CDU1 C. Gallinaceai    U4T0N8_9CHLA      288  -----SDTPFTIKKVVQTPIQE-DGFSCGIWITIFLEKYLENPQNLPE-IIDTLQNHKHS
CDU2 C. Trachomatis    CDUB2_CHLT2       255  SDQEELLSPFQVRIGSTVKVQSPGEFTCGAWCCQFLAWYLENPDFDLEEKVPTNPSERRA
                                                 * ::     :*      ** *  *   ::..         :

CDU1 C. Trachomatis    CDUB1_CHLT2       378  NLASFVQACEAAVQDLPELSWPEA---
CDU1 C. Suis           A0A061JBM7_9CHLA  372  NLASFAQACEEASQAFFDLSWPEAKEE
CDU1 C. Muridarum      Q9PL48_CHLMU      346  LLQSLHSTFEKLMKKYADLSWPTT---
CDU1 C. Psittaci       S7KIS5_CHLPS      350  ILRDFLDKASNPQQQLSSIIHPRLDNK
CDU1 C. Gallinaceai    U4T0N8_9CHLA      341  ILKDFLNKNPNPSPPQSNNI-------
CDU2 C. Trachomatis    CDUB2_CHLT2       315  LLADFISTTEQAMSRYSSLSWPTTD--
                                              * .: .           .
```

Figure 3.12. Sequence alignment of the protease domain of *C. trachomatis* Cdu1 with related proteases. VR-3 connecting β-strands 1 and 2 (highlighted in yellow) in the catalytic domain of Cdu1 is present in several *Chlamydia* species. The catalytic triad is marked in blue. Taken from Fischer et al, 2017 with permission.

3.5 Crystallization and structural analysis of Cdu1 in complex with ubiquitin.

Dissecting the molecular basis of host-protein recognition by bacterial effector proteins allows us to broaden our understanding on the structural mechanisms of infection, a notion that may readily be used for the rational design of novel anti-infective agents. Therefore, we aimed to solve the crystal structure of Cdu1 in complex with ubiquitin, as described in the following section.

3.5.1 Generation and crystallization of the Cdu1~ubiquitin complex

Ubiquitin-Propargylamide (Ubi-PA) is an active site directed probe where the C-terminal carboxylate of ubiquitin is replaced with an alkyne group. This long

considered unreactive warhead can nevertheless establish a covalent bond with the active site cysteine of deubiquitinating enzymes due to correct positioning of the warhead directly into the binding pocket, in a process driven by the natural complementarity between ubiquitin and DUBs (166). The generation and purification of Ubi-PA was performed by Dr. Florian Sauer who kindly provided it for the present experiment.

To avoid cysteine oxidation derived aggregation of the covalent adduct, the Cdu1_CT (155-401) was generated by gene synthesis (GenScript). Cloning was pursued as described for the wild type variant in the framework of the bachelor thesis of Christian Tiesmeyer (Tiesmeyer, 2016). This variant was identical to the Cdu1_Mut construct (Section 3.3), except for the active cysteine being left intact in order for it to be able to react with the C-terminus of Ubi-PA.
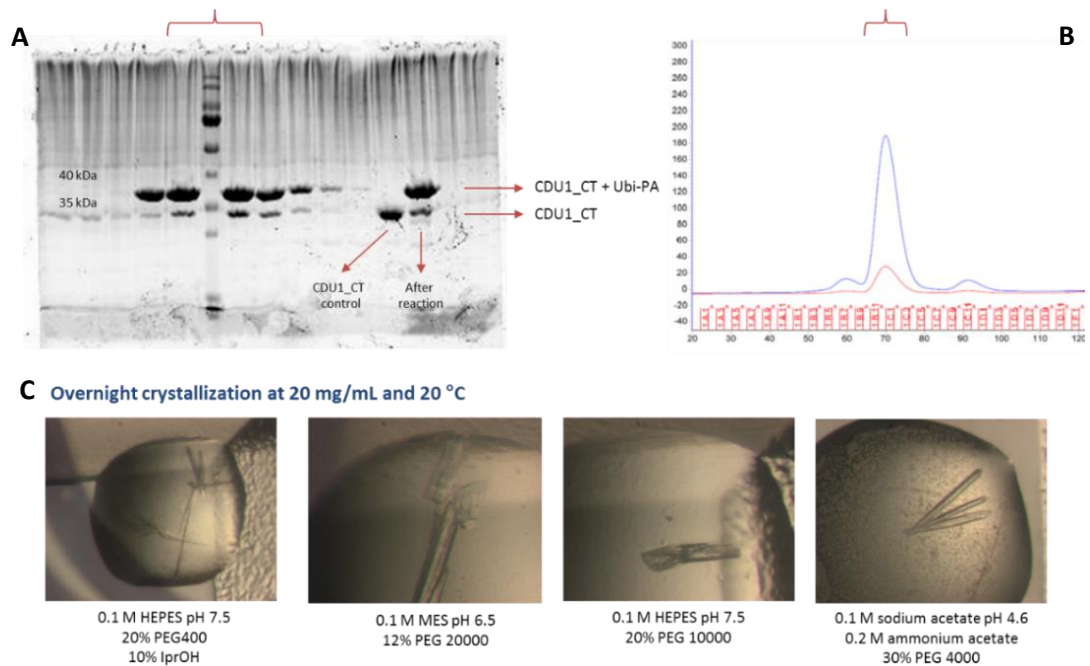
Figure 3.13. (A) SDS-PAGE after SEC purification of the Cdu1-ubiquitin-PA complex displaying a mass shift coherent with the binding of ubiquitin. Brackets above the gel and the chromatogram indicate the SDS-PAGE analysis of main chromatographic peak **B)** Elution profile of the Cdu1-Ubiquitin complex on a Superdex 75 16/60 column. **(C)** Overnight crystals of the purified fractions (red brackets) obtained at 20 mg/mL and 20 °C.

Cdu1_CT (155-401) was incubated at a 1:2 molar ratio with Ubi-PA overnight at 4°C. The product was purified by size exclusion chromatography (see Materials and Methods, section 2.2.1).

Although total separation of the Cdu1-Ubiquitin complex from the free Cdu1_CT variant could not be achieved (Figure 3.13 A and B), crystals could be obtained after overnight crystallization in 0.1 M MES (pH 6.5) and 12% PEG 20000 at 20 mg/ml and 20 °C (Figure 3.13 C).

| Data Collection	
PDB code	6FDK
Crystal	Cdu1~Ubiquitin
Space Group	$P2_1$
Unit Cell Parameters	
a, b, c (Å)	47.81 56.94 60.69 104.86
ß (°)	
Proteins per ASU	1
Solvent content (%)	44.58
Resolution Limits (Å)	46.22 - 1.60 (1.63 - 1.60)
R_{pim} (%)	2.7 (68.3)
$CC_{1/2}$ (%)	99.9 (49.9)
Multiplicity	6.7 (7.1)
Unique Reflections	41342 (2039)
Completeness	99.3 (100)
$<I/\sigma I>$	14.8 (1.1)
Average B-factor	24.28
macromolecules	22.51
ligands	19.14
solvent	35.10
Refinement	
R_{work} (%)	16.9
R_{free} (%)	20.3
RMSD Bond Lengths (Å)	0.03
RMSD Bond Angles (°)	2.51
Ramachandran statistics (%)	97.76 / 1.92 / 0.32

Table 3.2 Data Collection and Refinement Statistics for the Cdu1~ubiquitin complex. Data for the highest resolution shell are given in parentheses. Taken from Ramirez et al, 2018 with permission.

All crystals were soaked in a solution identical to the mother liquor with additional 40% PEG 400 as cryo protectant and flash frozen in liquid nitrogen. Diffraction data were collected at the European Synchrotron Radiation Facility (ESRF), beam line ID23-1, Grenoble. The best diffraction data were recorded for the crystals grown in 0.1 M MES (pH 6.5) and 12% PEG 20000 (up to 1.6 Å). Structure determination and refinement for the Cdu1~ubiquitin complex was carried out using the apo Cdu1 structure (PDB 5B5Q) and ubiquitin (PDB 1UBQ) as search model for molecular replacement. Manual model building was performed in COOT (115) and model refinement was carried out with REFMAC (123). Data collection and refinement statistics are summarized in Table 3.2.

3.5.2 The Cdu1~ubiquitin covalent complex
The complete section in taken from Ramirez et al, 2018.

The structure of Cdu1 in complex with ubiquitin reveals that the flexible C-terminal tail of the Ub-PA reaches Cdu1's catalytic cysteine (Figure 3.14 A). The Cdu1~Ub interface is characterized by a buried surface area of approximately 1089 $Å^2$ (VMD, "calculate SASA" tool (167)). Cdu1 residues Ile225 and Phe191 address the Ub Ile44 hydrophobic patch through van der Waals interactions with Ile44 and Val70. Remarkably, the side chain of Phe191 flips to make the binding pocket accessible for the C-terminus of ubiquitin. As hypothesized in section 3.4.3, Cdu1's α-helix D, which is unique for chlamydial DUBs, might play a role in substrate recognition by assuming a different orientation when the substrate is bound (90). However, our ubiquitin-bound structure shows that Cdu1 does not undergo significant conformational changes upon binding to ubiquitin apart from α-helix D tilting forward to address the Ub Ile36 patch through a hydrophobic contact with Ile267 (Figure 3.14 B). This movement positions the backbone carbonyl of Cdu1 Gly269 in a favorable orientation to address the side chain of Ub

Arg74 which is consistent with the reported inability of Cdu1 to bind ubiquitin upon deletion of α-helix D (i.e. VR-3 within the CE fold) (67).

Moreover, the side chain of Ub Leu73 forms a van der Waals contact with Cdu1 Trp247, thereby disturbing the internal sulphur$\cdots\pi$ interaction (168) between this residue and Met262, as observed in the apo-form (Figure 3.11 B, inserted panel). Ubiquitin Ile73 can be described as a wedge that separates the Trp247-Met262 pair, which in turn facilitates the motion of the chlamydial α-helix D to permit proper binding to the substrate. Cdu1 Trp247, Ser168 and Asp167 play a role in recognizing the C-terminus of Ub by establishing hydrogen bonds with the backbone carbonyl of Arg74, Leu73 and Arg72, respectively. Additionally, the side chains of Cdu1 Asp169 and Glu395 assume conformations towards ubiquitin's Arg42 and Gln49 to further stabilize the complex.

Contacts with the hydrophobic Ub Ile44 and Ile36 patches are the most commonly observed DUB-Ub interaction motives. Both features are conserved in NEDD8, but not in other Ubls such as SMT3 (99). Other conserved residues in NEDD8 include Arg42, Gln49, His68, Val70 and Arg74 (Figure 3.14 C) – all of which are also present in ubiquitin and are important for the Cdu1-Ub interaction. Recognition of these conserved residues provides a molecular explanation of the reported deneddylating activity of Cdu1 (92).
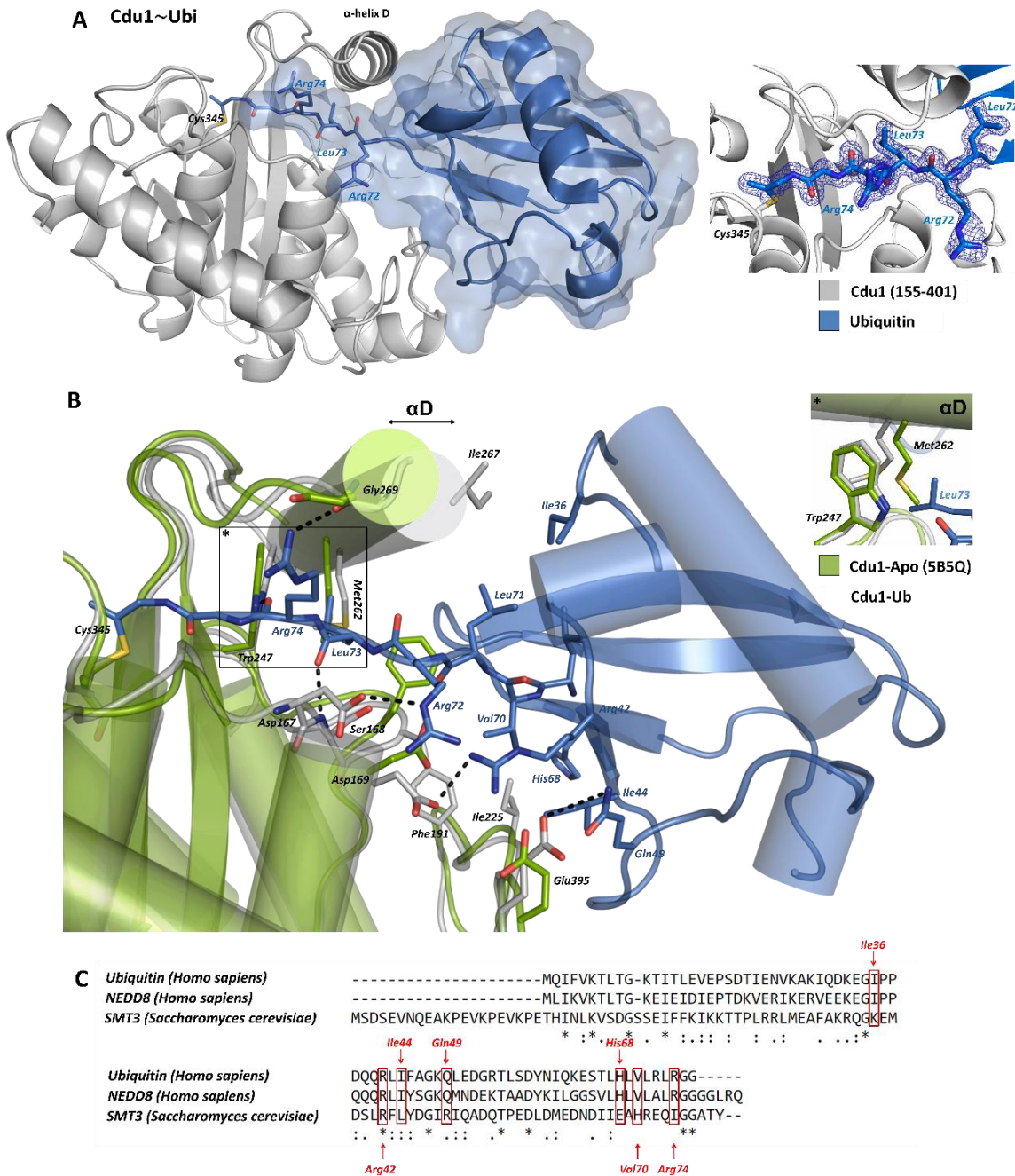
Figure 3.14. (A) Crystal structure of the Cdu1~Ub complex (cartoon representation) and 2Fo-Fc map of the C-terminus of ubiquitin (Leu71-Arg74) within the active site of Cdu1 (1.0 σ, blue). **(B)** Superposition of the Cdu1 apo (PDB: 5B5Q, green) and the Cdu1-Ub complex structure (PDB: 6FDK, grey) illustrating the interaction network and the movement of α-helix D towards ubiquitin upon binding of the substrate. **(C)** Sequence alignment of ubiquitin, SMT3 and NEDD8. Key interaction spots between Cdu1 and ubiquitin (Ile36, Arg42, Ile44, Gln49, Val70 and Arg74) are also possible in NEDD8 at equivalent positions. Taken from Ramirez et al. 2018 with permission.

Our observations thus suggest that ubiquitin recognition is triggered by a concerted mechanism involving the interaction of both the C-terminus (Leu71 to Gly76) and the globular core of ubiquitin with key Cdu1 residues. This facilitates a conformational change on α-helix D to allow proper binding to the substrate. Since the key residues for the described interactions are conserved between ubiquitin and NEDD8 we assume that NEDD8 is bound to Cdu1 in a very similar way as observed for ubiquitin and Cdu1 may undergo similar movements to accommodate NEDD8.

3.6   Screening of potential Cdu1 covalent inhibitors

3.6.1   Setup and optimization of a DUB enzymatic assay

As stated by Ramirez et al, 2018, the first step towards the evaluation of Cdu1 covalent ligands was the identification of suitable and standardized enzymatic assay conditions. The second-generation DUB substrate Ub-Rh110Gly was used based on its higher dynamic range and unlikeliness to interfere with the fluorescence of test molecules, as compared to other ubiquitinylated substrates (137). The dependency of Cdu1's enzymatic activity on the Ub-Rh110Gly concentration was explored by titrating 0.3 nM enzyme with increasing concentrations of substrate. This enzyme concentration displayed a linear phase of at least 2 min reaction time in control experiments. In contrast to the Cdu1_LCMS (155-401) construct used for crystallization, the Cdu1_S (125-401) variant proved to be superior with respect to replicate tests. However, the longer construct did not display a different catalytic activity compared to the shorter construct.

The following conditions were found to improve the reliability of the DUB assay in terms of consistency among replicates:

- Buffer conditions: 20 mM NaCl, 20 mM HEPES pH 7.5, 1mM TECEP, 50 µg/mL BSA.

- Substrate concentration (Ub-Rh110Gly): 250 nM

- Construct used: 0.3 nM ChlaDUB1_S (125-401)

- Total reaction volume: 40 µL per well

- Initiation of the reaction: 0.6 nM protein mixed with 500 nM substrate (20 + 20 µL)

- Fluorescence measurement at: Every 20 s for 2 h

Michaelis–Menten kinetic analyses of the initial reaction velocities yielded a K_M of 1.7 µM and a V_{max} of 179.4 RFU/s (Figure 3.15). This K_M value is comparable to the CE clan deubiquitylase SseL (*Salmonella typhimurium*), with a K_M of 1.5 µM for ubiquitin-1-amido-methyl-coumarin (Ub-AMC) (101).
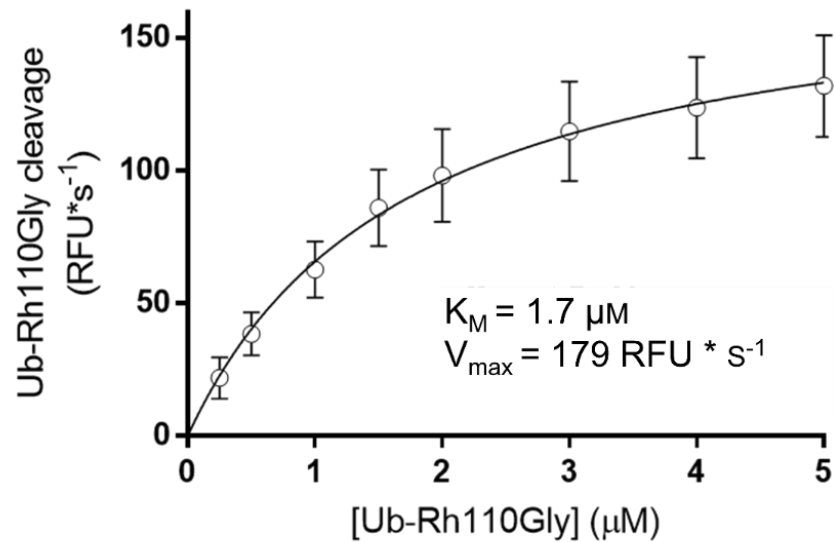
Figure 3.15. Michaelis-Menten analyses of Ub-Rh110Gly substrate cleavage in the presence of 0.3 nM Cdu1_S (residues 125-401). Taken from Ramirez et al, 2018 with permission.

3.6.2 Structure similarity-based target hopping of adenain inhibitors on Cdu1

The complete section in taken from Ramirez et al, 2018.

As described in section 1.4.5, adenain is the only CE protease to date which has been the subject of a structure-informed drug design campaign, making its covalent inhibitors an attractive starting point for the development of ligands for other CE proteases. A structural alignment (116) of Cdu1 and adenain revealed striking similarities between both proteins (Figure 3.16 A). These similarities are not limited to the overall fold and spatial distribution of the secondary structure elements, but also include the topology of the active sites which are remarkably conserved (Figure 3.16 B). Therefore, we hypothesized that the homology between adenain and Cdu1 offers the possibility to identify primary features for the development of specific Cdu1 covalent inhibitors based on the already known adenain inhibitors.

A set of 7 adenain inhibitors previously reported in the literature (109, 139) (Figure 1.11) with different covalent warheads (one vinyl- sulfone, two cyano-pyrimidines and four glycine nitriles) were screened against Cdu1 (Figure 3.17 A and B). Remarkably, only cyano-pyrimidines **3** and **5** yielded more than 50% of inhibition at a concentration of 100 µM after 1.5 h incubation with Cdu1 (Figure 3.13 7) suggesting that the cyano-pyrimidine warhead might be a suitable electrophile to covalently label the active site Cys345.

Reversible binding of cyano-pyrimidines **3** and **5** could be demonstrated by a rapid dilution assay, in which Cdu1 was pre-incubated with a 10-fold excess of the equivalent IC_{50} for each compound to ensure loss of enzymatic activity. After rapid-dilution into assay buffer containing the substrate, DUB activity was restored within 2 hours. The lag phase during the first 30 min of the progression curve followed by a linear phase (Figure 3.17 D) indicates a slow reversible binding, as Cdu1 recovers its activity while the inhibitors gradually dissociate (145). The

compounds were found to have a modest potency against Cdu1, with IC_{50} values of 204 μM and 113 μM for **3** and **5,** respectively (Figure 3.17 E), in contrast to 0.06 μM and 24 μM for adenain.
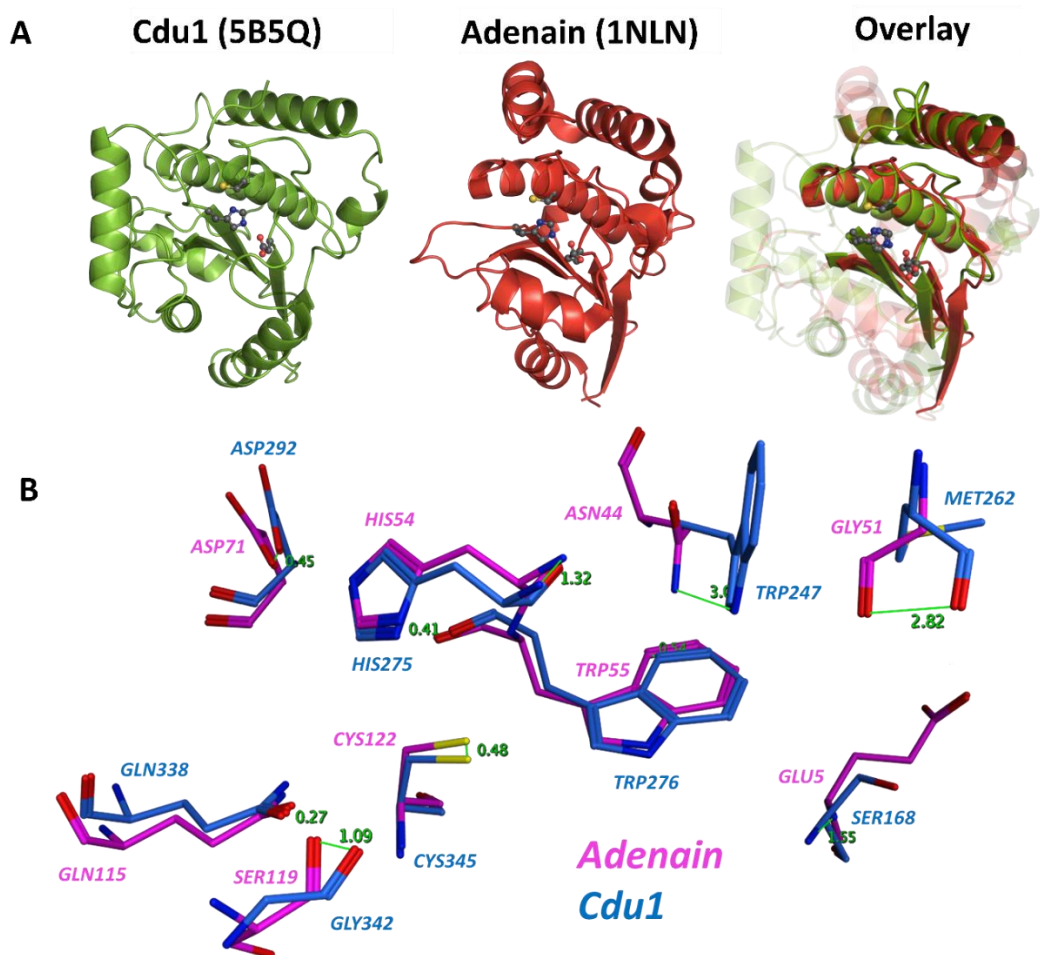
Figure 3.16. (A) Cartoon representations of Cdu1 (PDB: 5B5Q, green) and adenain (PDB: 1NLN red). In the overlay, the structures were aligned based on the cysteine-containing α-helix (catalytic triad depicted in gray, ball and stick representation). **(B)** 3D conservation mapping of the residues surrounding the active site cysteines of Cdu1 (blue) and adenain (magenta). Distances between equivalent atoms (Å) are depicted in green. Taken from Ramirez et al, 2018 with permission.

The IC_{50} value represents the concentration of inhibitor needed to impair the enzymatic activity by half (142). It should be highlighted that, although it has been suggested that IC_{50} values might not be the best way to describe covalent

inhibitors as this parameter is known depend on the incubation time (140), they are certainly useful to get an idea of the potency of a given covalent drug.
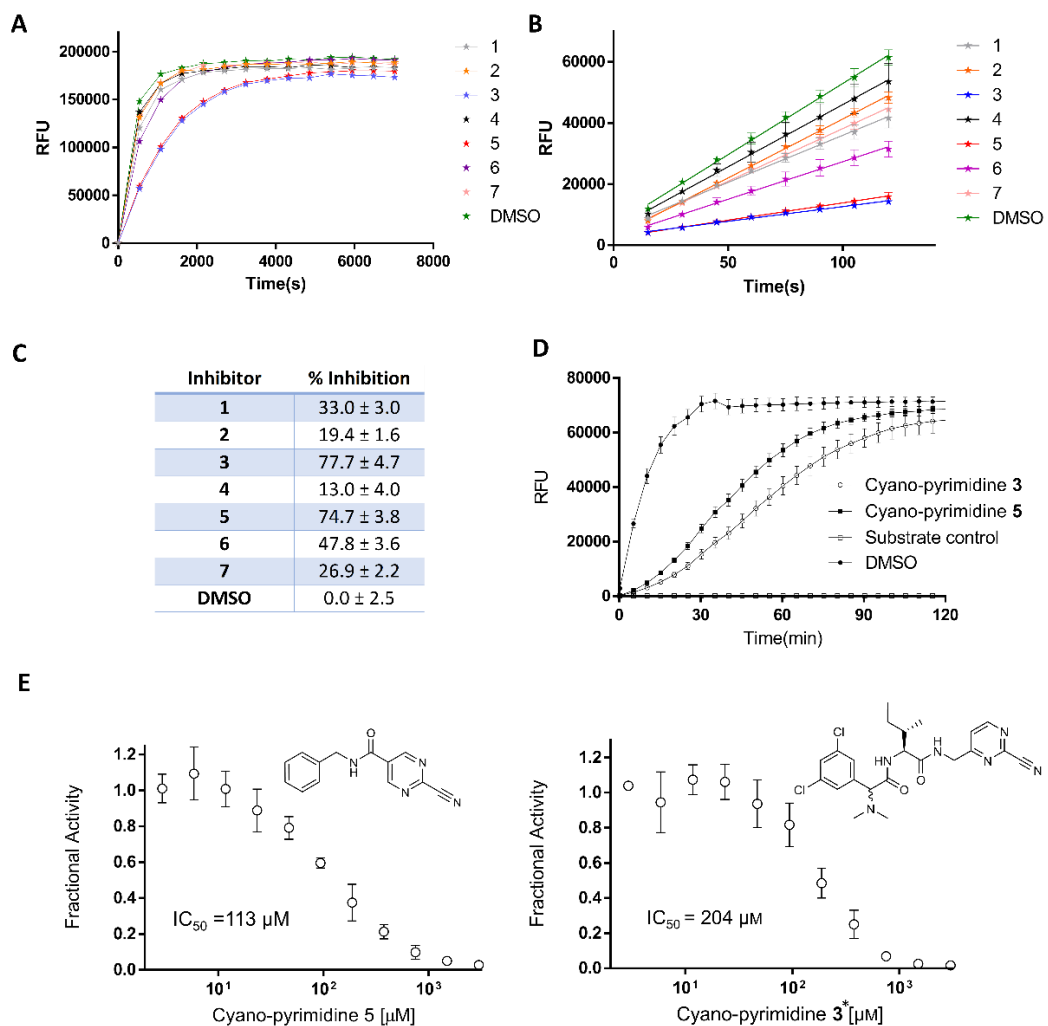
Figure 3.17. (A) Progression curve of Cdu1 substrate cleavage (Ub-Rh110Gly) after 1.5 h incubation with 100 µM compounds **1** to **7** (n_{rep}=3). **(B)** Initial phase of the substrate cleavage progression curve (Ub-Rh110Gly, t=120 s) after incubation with 100 µM compounds **1** to **7** for 1.5 h (n_{rep}=3). **(C)** Percentage of inhibition of compounds **1** to **7** at 100 µM. **(D)** Rapid dilution assay. Time course of substrate cleavage: cyano-pyrimidine 3 (n_{rep}=6); cyano-pyrimidine 5 (n_{rep}=6); substrate control (n_{rep}=3). The DMSO control shows the uninhibited reaction (n_{rep} = 6). Error bars indicate the standard error of the mean. **(E)** Concentration-response curve for Cdu1 (125-401) activity after 1 hour of incubation with cyano-pyrimidine 3 (right) and cyano-pyrimidine 5 (left). The data represent the mean of at least 3 independent experiments performed in duplicate. (*)mixture of two diastereoisomers. Taken from Ramirez et al, 2018 with permission.

3.6.3 Screening of vinyl-methyl-ester electrophilic fragments as potential Cdu1 binders

A set of 180 thiol-reactive vinyl-methyl-ester (VME) compound library (kindly provided by Prof. Alexander Statsyuk) was tested in the DUB assay as described above (see appendix, Table 6.3). This library has already proven to be effective for identifying weak covalent binders for other proteases (141) and might serve as starting point for inhibitor design. Initial phase kinetics were determined for each ligand and the percentage of inhibition was calculated (Figure 3.18). Unfortunately, no hit could be identified out of this library. This result suggests that the VME warhead might lack sufficient reactivity to address the active site cysteine in Cdu1 through a covalent bond.
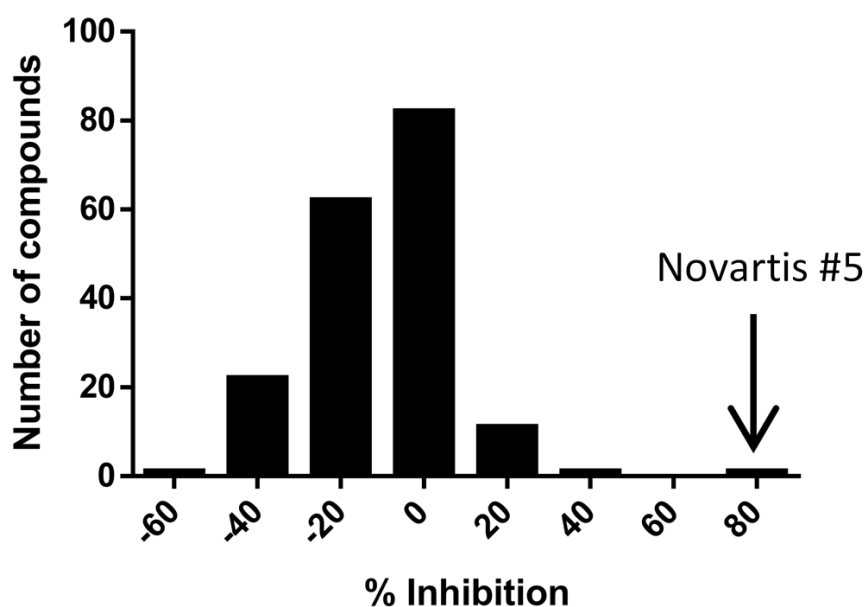
Figure 3.18. Percentage of inhibition distribution of 180 VME ligands (n=2) on Cdu1 activity, with cyano-pyrimidine **5** (Novartis) as positive control. Taken from Ramirez et al, 2018 with permission.

3.7 Crystallization and structural analysis of Cdu1 in complex with cyano-pyrimidines 3 and 5

The complete section in taken from Ramirez et al, 2018.

3.7.1 Crystallization and structure elucidation

Crystals of Cdu1_LSMS (155-401) were soaked in a solution containing 5 mM of cyano-pyrimidine **5**, and 50 mM of cyano-pyrimidine **3** (racemic) for 30 mins and 65 hours respectively. The crystals were taken to the European Synchrotron Radiation Facility (ESRF) for X-ray data collection at beam line ID23-1 and solved to a resolution of 2.3 Å (Table 3.3).

Data Collection		
PDB code	6FDQ	6FDU
Crystal	Cdu1~Cyano-pyrimidine **5**	Cdu1~Cyano-pyrimidine **3**
Space Group	$P2_1$	$P2_1$
Unit Cell Parameters		
a, b, c (Å)	43.59 57.04 114.59	44.08 57.09 116.38
ß (°)	94.73	96.00
Proteins per ASU	2	2
Solvent content (%)	51.51	52.74
Resolution Limits (Å)	43.44 - 2.30 (2.38 - 2.30)	43.84 - 2.30 (2.38 - 2.30)
R_{pim} (%)	10.2 (51.8)	5.3 (38.5)
$CC_{1/2}$ (%)	98.3 (73.8)	99.8 (70.1)
Multiplicity	3.3 (3.2)	5.0 (5.2)
Unique Reflections	24607 (2344)	25784 (2502)
Completeness	98.8 (96.3)	99.9 (99.9)
$<I/\sigma I>$	3.5 (1.1)	8.8 (2.1)
Average B-factor	47.58	44.32
macromolecules	47.54	44.39
ligands	64.29	69.88
solvent	44.61	42.25
Refinement		
R_{work} (%)	22.9	20.0
R_{free} (%)	28.3	24.0
RMSD Bond Lengths (Å)	0.02	0.02
RMSD Bond Angles (°)	2.01	2.08
Ramachandran statistics (%)	95.02 / 4.36 / 0.62	96.63 / 2.74 / 0.63

Table 3.3 Data Collection and Refinement Statistics for the Cdu1~ cyano-pyrimidine complexes. Data for the highest resolution shell are given in parentheses. Taken from Ramirez et al, 2018 with permission.

3.7.2 The Cdu1~cyano-pyrimidine complexes

Despite their modest potency, crystal structures of **3** (6FDU) and **5** (6FDQ) in complex with Cdu1 could be obtained, demonstrating that the reactivity of a cyano-pyrimidine warhead is sufficient to form a covalent bond to Cys345 in Cdu1 (Figure 3.19 A). The electron density for **5** is better defined in chain A of our structure, while **3** is only visible in chain B. The stabilizing effect caused by a hydrogen bond of the thioimidate group of the reaction product to Gln338 might play a crucial role. This hydrogen bond is clearly established in the crystal structure of **3** and **5** (Figure 3.19 B). The spatial features of Cdu1's substrate binding pocket permit an almost planar arrangement of the thioimidate group and the pyrimidine ring. Such a conformation represents an extended delocalized system upon reaction and contributes to the stabilization of the reaction product.

An intermolecular interaction especially important for stabilizing the binding to Cdu1 is a hydrogen bond donated by Trp247. **3** and **5** are able to establish this interaction through a carbonyl group of the amide close to the pyrimidine ring. Another important interaction observed for **3** is a hydrogen bond donated by the main chain nitrogen of Ser168.

The electron density of the benzene ring of **5** and the dichloro-benzene ring of **3** was not completely defined in the X-ray structures (Figure 3.15 A). A B-factor analysis suggest that this might reflect flexibility of these parts of the ligands within the active site of Cdu1 (Figure 3.19 C). The aryl group in **5** might be temporarily stabilized by a weak van der Waals contact with Ala266. The same may be the case for the dichloro-benzene ring of **3,** which is mainly solvent exposed.
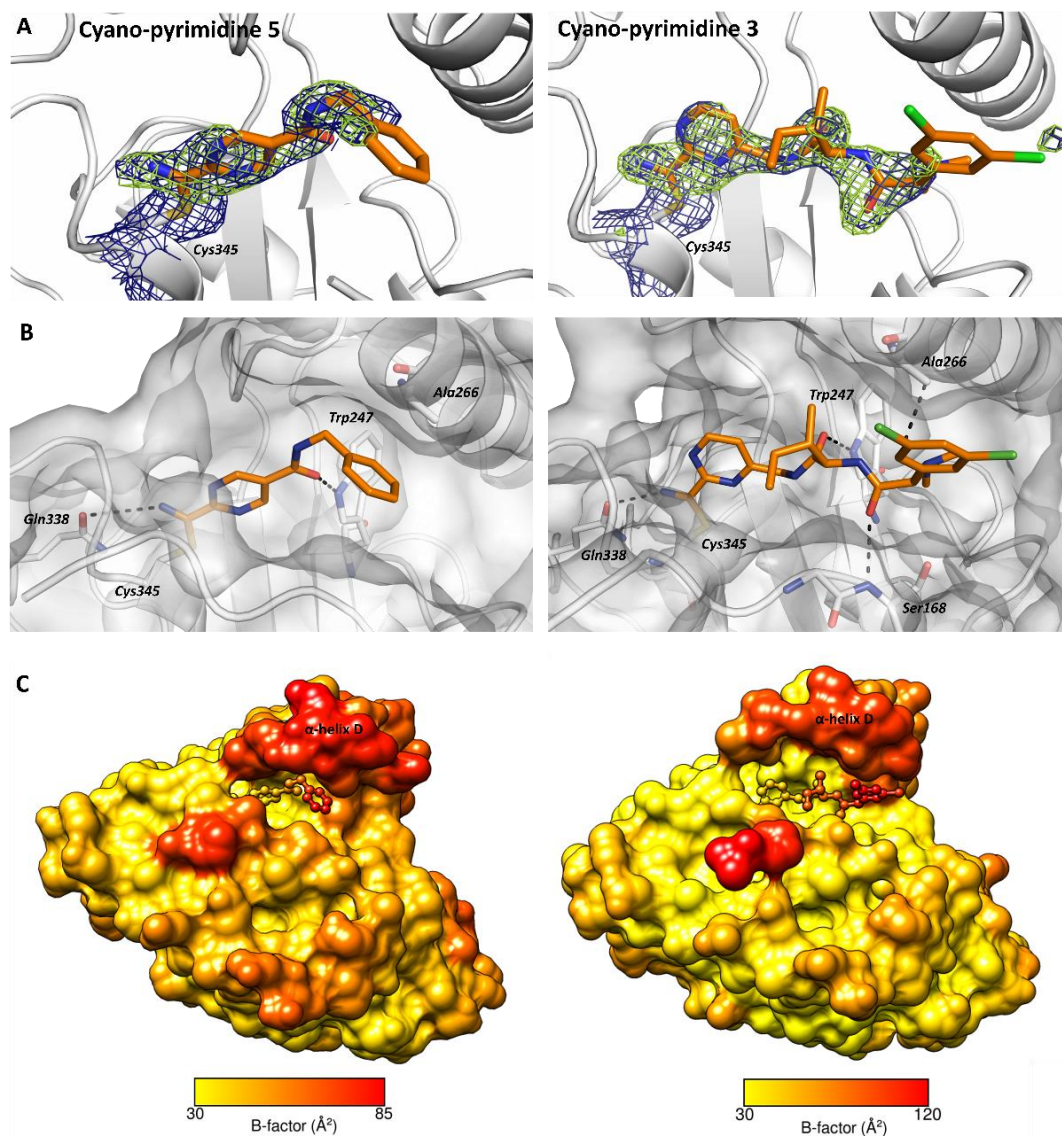
Figure 3.19. (A) 2Fo – Fc (1.0 σ, blue) and Fo-Fc (3.0 σ, green) omit maps for cyano-pyrimidine 5 (Cdu1 chain A, left) and for cyano-pyrimidine 3 (Cdu1 chain B, right). **(B)** Experimentally observed binding modes of cyano-pyrimidines **3** (PDB: 6FDU, right) and **5** (PDB: 6FDQ, left). **(C)** Surface representation of Cdu1 bound to cyano-pyrimidines **5** and **3** (represented as balls and sticks) colored based on crystallographic B-factor values (yellow: low, red: high). High B-factors (red) correlate with increased flexibility. Taken from Ramirez et al, 2018 with permission.

3.7.3 Cdu1's pocket flexibility

The Cdu1~cyano-pyrimidine bound structures also reveal important features with respect to the flexibility of the substrate binding site. In contrast to **5**, which is *para*-substituted at the cyano-pyrimidine ring, **3** is *meta*-substituted. To establish the observed hydrogen bond with Trp247, the warhead on **3** is forced to move towards Val241. To avoid a steric clash, the flexible loop between Cdu1's α-helix D and β-strand 2 moves up towards an open position (Figure 3.20, depicted in gray), thus breaking the internal hydrogen bond between Val271 and Ser274, which otherwise stabilizes a closed conformation (Figure 3.20, depicted in orange). The presence of the bulky iso-pentyl substituent in **3** might also contribute to the loop opening, breaking an additional hydrogen bond between Ala266 and Gly269. The flexibility of α-helix D is also reflected in high temperature factors for this region as compared to the rest of the protein (Figure 3.19 C).
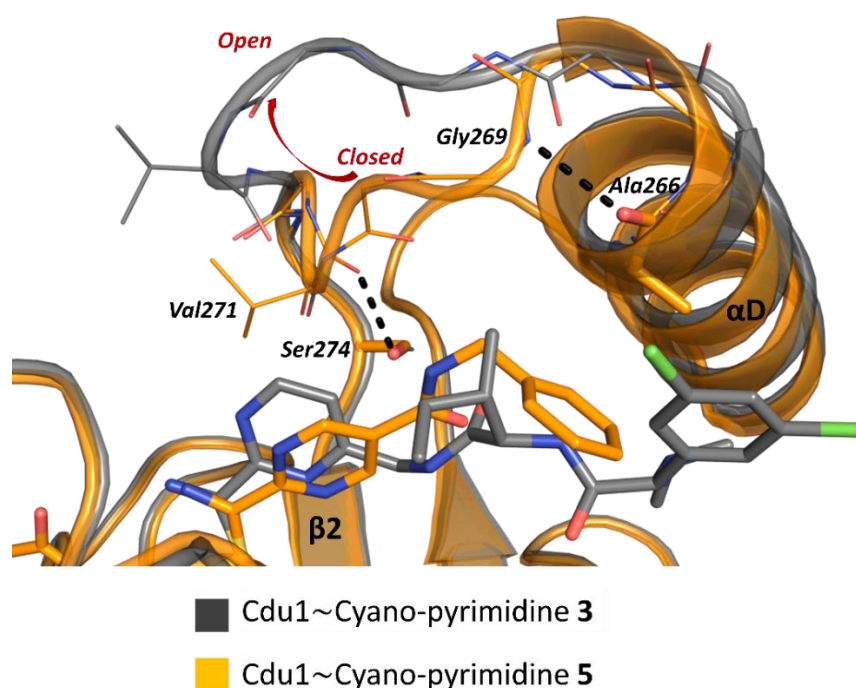
Figure 3.20. Superposition of Cdu1 in complex with **3** (gray) and **5** (orange). The loop between α-helix D and β-strand 2 moves towards an open conformation to accommodate cyano-pyrimidine **3**. Taken from Ramirez et al, 2018 with permission.

3.7.4 Cross-target specificity of cyano-pyrimidines 3 and 5

As hypothesized, the cyano-pyrimidine warhead of both ligands binds to Cdu1 in a similar way as to adenain with an RMSD (169) over the heavy atoms of the warhead of 0.56 Å for **3** and 1.57 Å for **5** (Figure 3.21, right). There are, however, a few key differences in the binding modes beyond the covalent warhead. The terminal aromatic ring of **5** establishes a van der Waals contact with adenain's Val53, moving the ligand outside of the substrate binding site (Figure 3.21, top-center). In contrast, this ring establishes the same type of interaction with Cdu1's Ala266, keeping the ligand below α-helix D and positioning the ligand's carbonyl group at an optimal distance to form a hydrogen bond with Trp247 (Figure 3.21, top-left). The backbone of cyano-pyrimidine **3** addresses Cdu1's Glu338, Trp247 and Ser168 through hydrogen bonds (Figure 3.21, bottom-left). This is analogous to how the ligand interacts with adenain's Glu115, Asn44 and Glu5 (Figure 3.21, bottom-center).
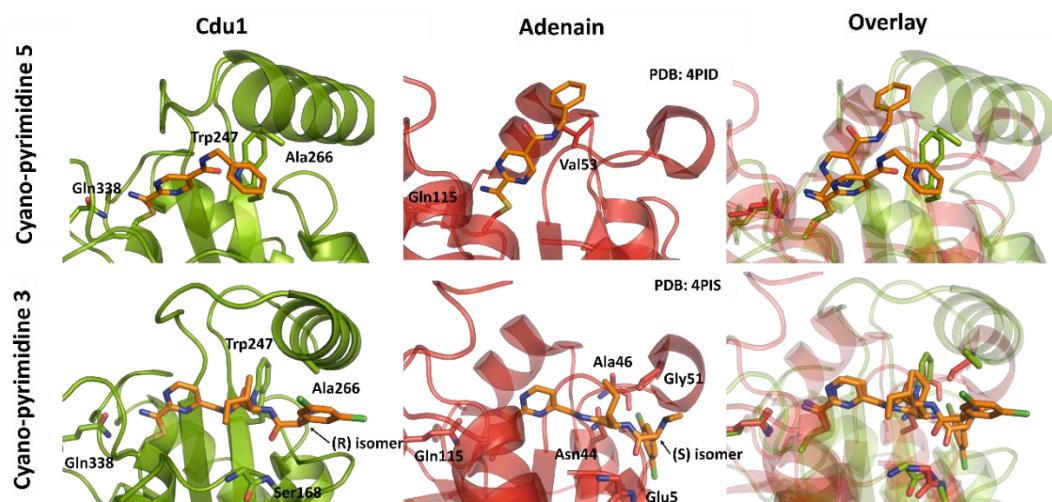
Figure 3.21. Comparison of the experimental binding modes of cyano-pyrimidines **3** and **5** to Cdu1 (green, left column) and adenain (red, central column), and superposition of the complexes (right column). The R- and the S-configurations of **3** are indicated in the lower left and the lower middle panel. Taken from Ramirez et al, 2018 with permission.

It is worth to highlight that, unlike **5**, **3** has improved specificity for adenain (109).In adenain, the terminal dichloro-benzene ring is positioned within a hydrophobic pocket while establishing a weak halogen bond with Ala46. These features are absent in Cdu1 where the dichloro-benzene group is oriented towards the solvent and lacks shape complementarity with the pocket. This might be one of the factors explaining the considerable difference in potency of **3** between the two proteases. Interestingly, the 2-(S)-dimethylamino isomer of **3** seems to preferentially bind to adenain while the 2-(R)-dimethylamino isomer binds to Cdu1, where the bigger α-helical insert above the active site impairs the dichloro-benzene ring to fit inside the pocket, as observed for the 2-(S) isomer in adenain (Figure 3.21). This different preference might be one way to increase inhibitor specificity.

3.8   Virtual screening and covalent docking of Cdu1 ligands

3.8.1   Pose reproduction of cyano-pyrimidines by covalent docking
This section is adapted from Ramirez et al, 2018 with permission

A detailed analysis of the interaction patterns and conformations in the crystal structures of **3** and **5** provides a basis for our design efforts towards the identification of more potent and specific cyano-pyrimidine based inhibitors. The structures permit the calibration of a Cdu1-optimized covalent docking procedure to quickly predict the binding orientation of cyano-pyrimidine ligands in the binding pocket of Cdu1. In this type of docking, a covalent bond between the electrophilic atom within the ligand's warhead and the nucleophilic sulphur atom of the active cysteine is defined as a constraint prior to the docking routine.

Via crystal pose reproduction studies on **3** and **5** using GOLD's (126) covalent docking routines and by probing all scoring functions available in GOLD and different sampling settings, the empirical ChemScore scoring function (147) was

found to be the most reliable model to predict the key interactions and ligand conformations observed in the experiment. For **5**, the docking result predicted a weaker hydrogen bond to Gln338 with an unfavorable geometry, whereas the crystal structure exhibits optimal spatial positioning of the thioimidate for this interaction (Figure 3.22, left). For **3,** the hydrogen bond donated by the main chain nitrogen of Ser168 to the carbonyl moiety, was correctly predicted **(**Figure 3.22, right). The ChemScore docking routine of both **3** and **5** also correctly predicted the hydrogen bond to Trp247, a key interaction also featured in the crystal structures, whereas the other GOLD scoring functions failed consistently (Figure 3.22, highlighted in red). To decide with more confidence on the final representative docking pose to be compared with the experiment, we relied not only on ChemScore, but also included the predictive power of the DSX scoring functions (125) by rescoring the 50 best GOLD poses and calculating a consensus rank (see Methods section). The thus selected, representative docking poses deviate from the experiment by an RMSD of 1.04 Å (rank 2) and 0.93 Å (rank 1) for **3** and **5**, respectively (further details about the obtained poses can be found in the appendix).

Figure 3.22. Comparison of the representative top-ranked docking poses (green) modeled with GOLD and a consenus-ranking approach with the experimentally observed binding modes (orange) of cyano-pyrimidines **3** (PDB: 6FDU, right) and **5** (PDB: 6FDQ, left). The docking poses of **3** (consensus rank 2) and **5** (consensus rank 1) differ by 1.04 and 0.93 Å from the corresponding experimentally observed poses. The docking pose predicted by GoldScore for **5** (red, left), fails to establish the experimentally observed interaction with Trp247. Adapted from Ramirez et al, 2018 with permission.

3.8.2 Analysis of cyano-pyrimidine 3 derivatives

The experimentally observed superiority of the cyano-pyrimidines as compared to other covalent warheads (section 3.6.2) brought our attention to other adenain inhibitors based on this moiety. The ability of four additional cyano-pyrimidine **3** derivatives to fit within Cdu1's active site cleft, which were available from our collaborators at Novartis, was evaluated in our optimized covalent docking routine as described in section 3.8.1. Since the backbone of **3** induces an opened loop conformation in Cdu1 (section 3.7.3), this structure was used for the docking of all cyano-pyrimidine **3** derivatives (PDB ID: 6FDU). Like for the pose reproduction of cyano-pyrimidines **3** and **5**, the empirical scoring function ChemScore and the knowledge-based DSX scoring function were taken into consideration. A known liability of scoring functions is that they generally tend to score bigger molecules higher, and hence, large molecules may show up as false positives in structure-based virtual screens (170). To compensate this issue, all scores derived from our docking analysis were normalized by the cubic root of the heavy atom count of the ligand for the remaining of the discussion. Thereby, the cubic root score (CRS) (171) is defined according to the equation:

$$CRS = Score * \frac{1}{\sqrt[3]{Heavy\ atom\ number}}$$

Cyano-pyrimidines **138**, **139**, **140** and **141** were found to potentially be able to address the active site pocket of Cdu1 in an analogous way to **3** and may reveal further interaction hotspots (Figure 3.23 A). For instance, **140** in which the bulky isopentyl group is replaced by a 2-propanol moiety, seems to have the capability to simultaneously address **Trp276** and **Trp247** (Figure 3.23 B). Similarly, **141** could potentially also address **Asp167** through its flexible amino-dimethyl group (Figure 3.23 C). CRS analysis revealed that **140** is favored by ChemScore (rank 1), and by DSX_CSD (rank 2) in comparison to other derivatives of **3** (Table 3.4).
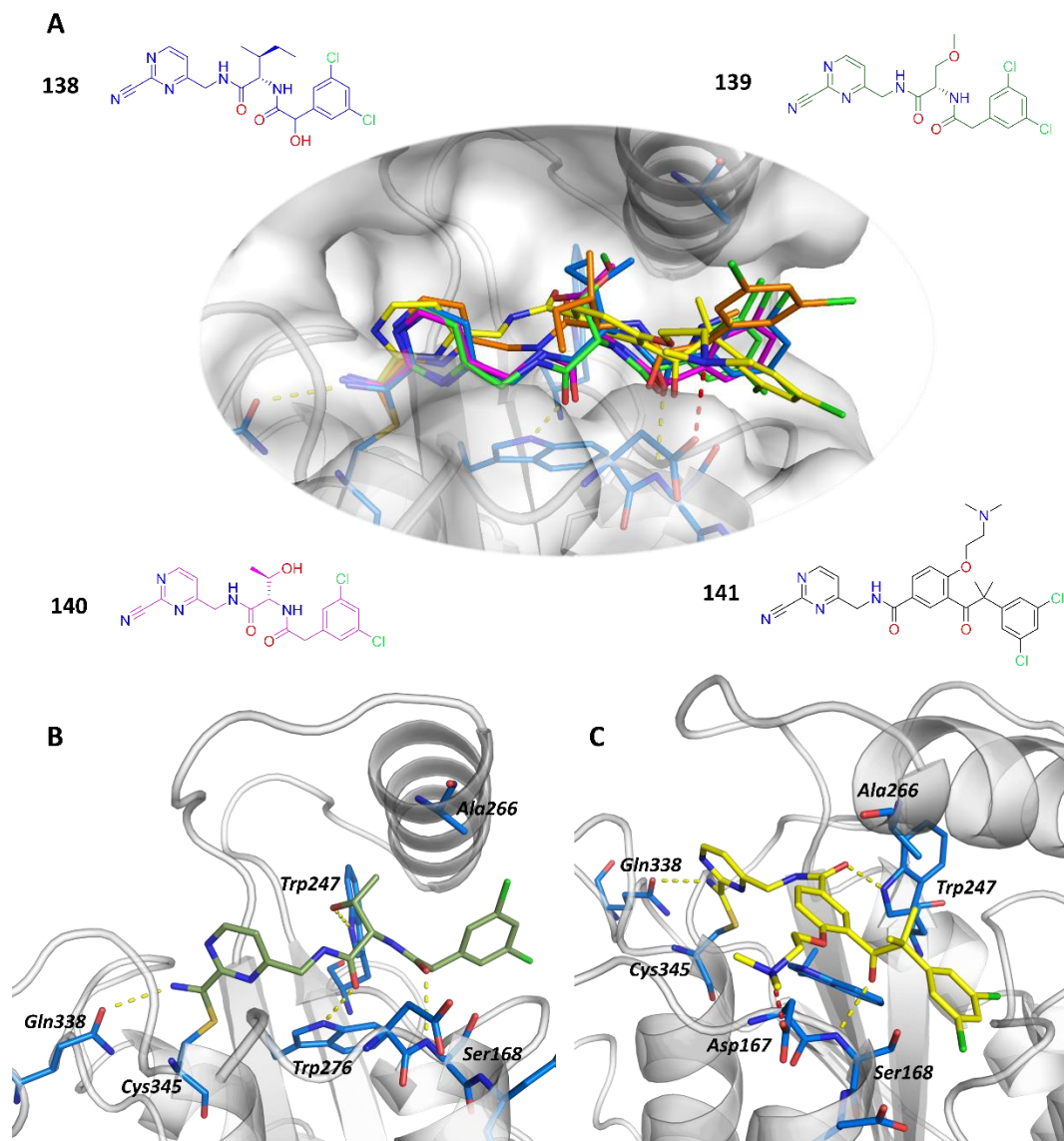
Figure 3.23. (A) ChemScore poses for cyano-pyrimidine **3** derivatives **138**, **139**, **140** and **141** in complex with Cdu1. **(B)** ChemScore pose for Cyano-pyrimidine **140**, simultaneously addressing Cdu1's **Trp276** and **Trp247**. **(C)** ChemScore pose for Cyano-pyrimidine **141**, showing how it could potentially address Cdu1's **Asp167**.

Experimental validation of cyano-pyrimidine **3** derivatives was performed enzymatically through the percentage inhibition experiment (section 2.2.4). This revealed that, within the experimental uncertainty of the method, all compounds are equally good Cdu1 inhibitors, comparable to **3** (Figure 3.24). This suggests that the backbone of **3** is indeed promising for Cdu1 inhibition. Although the

determination of the IC_{50} of **138**, **139**, **140**, and **141** was not performed within this work, the docking results readily provide some ideas for specific ligand development, as will be described in section 3.9.1.

Ranking according to ChemScore

	Lig #	CRS	ChemScore	HAN
1	140	13.22	40.15	28
2	139	12.79	38.83	28
3	141	11.95	39.83	37
4	138	11.86	36.84	30
5	3	10.38	32.97	32

Ranking according to DSX_CSD

	Lig #	CRS	DSX_CSD	HAN
1	3	32.14	102.03	32
2	140	30.23	105.13	28
3	138	27.31	84.86	30
4	139	25.88	78.59	28
5	141	25.05	83.46	37

Table 3.4 Docking data for **3, 138**, **139**, **140** and **141**

Lig #	% inhibition	IC_{50} (µM)
3	77.7± 4.7	204
138	95.1 ± 12.0	--
139	87.8 ± 8.9	--
140	85.0 ± 6.4	--
141	96.3 ± 31.5	--
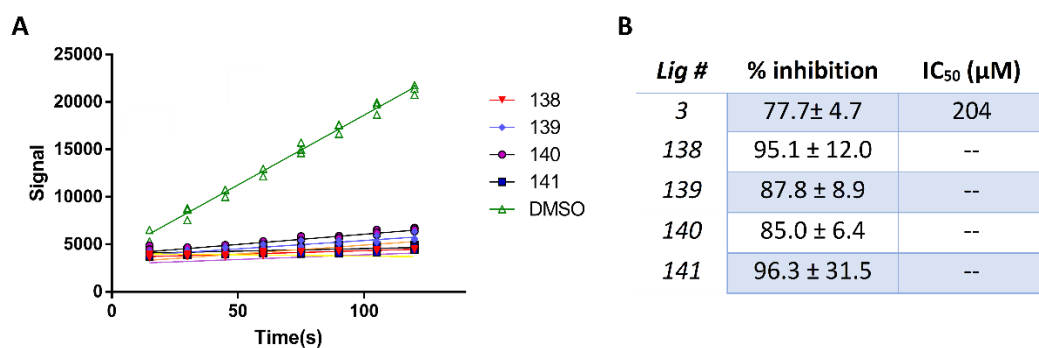
Figure 3.24 (A) Initial phase of Cdu1 substrate cleavage progression curve (Ub-Rh110Gly, t=120 seconds) after 1.5 h incubation with 100 µM compounds 138 to 141 (n=3). **(B)** % of inhibition compounds **138** to **141**. Inhibition data for cyano-pyrimidine **3** is included for comparison.

3.8.3 Analysis of cyano-pyrimidine 5 derivatives and Cathepsin K (CatK) inhibitors

To further exploit the potential of the cyano-pyrimidine warhead for Cdu1 inhibition, a search for commercially available molecules containing this warhead in the ZINC database was performed (130). This work was done within the framework of a research internship by Gerhard Keller (Keller, 2017). From a total of 5000 cyano-pyrimidines initially found, a substructure search for the scaffold of **3** and **5** was applied as a selection criterion (Figure 3.25).
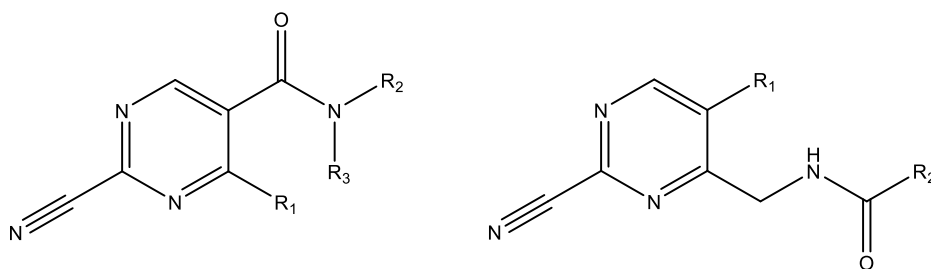
Figure 3.25. Substructure search for molecules sharing the backbone of cyano-pyrimidines **3** and **5** within the ZINC database. (R_n = Any atom)

On the basis of the substructure of **5,** 104 molecules were found. The search for the substructure of **3** produced no hits. This set of molecules was submitted to our covalent docking routine on the closed conformation of Cdu1 (PDB ID: 6FDQ), using ChemScore as primary scoring function and for DSX rescoring. For comparison, cyano-pyrimidine **5** was included in this analysis as molecule 105. The 5 ligands with the best CRS values (normalized against ChemScore) were visually inspected using a custom rational hit-picker tool, developed by Dr. Thomas Adler (Table 3.5).

Interestingly, compounds **79** and **90**, developed by Novartis for the treatment of disorders involving cathepsin K (CatK) (172), were identified to be among the favorites (Figure 3.26).
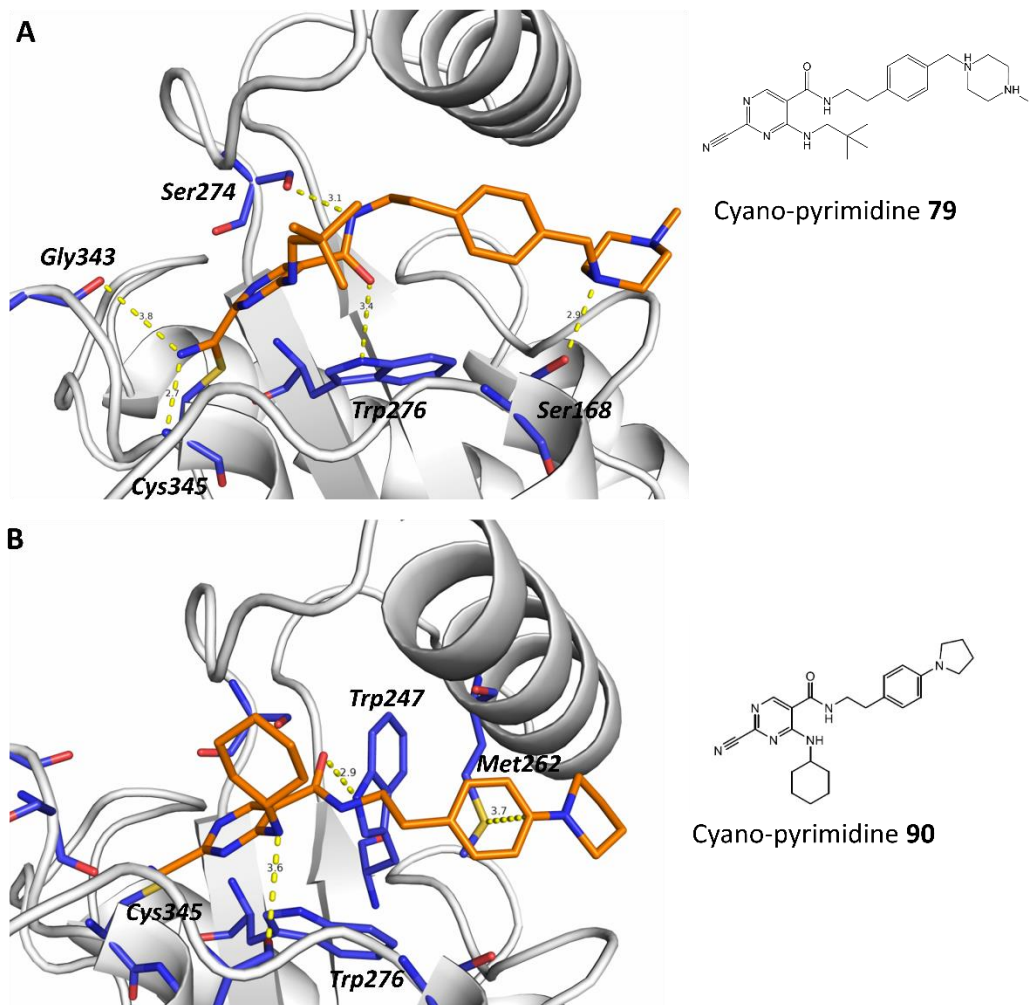
Figure 3.26. (A) Top ChemScore pose for cyano-pyrimidine **79** in complex with Cdu1, addressing Cdu1's **Trp276, Ser274** and **Ser168.** **(B)** Top ChemScore pose for cyano-pyrimidine **90**, addressing Cdu1's **Trp247** and **Met262.**

	Lig #	ChemScore	CRS	HAN
1	90	39.13	12.46	31
2	79	38.75	12.21	32
3	38	42.12	11.76	46
4	1	41.76	11.66	46
5	67	38.03	11.63	35

Table 3.5 Docking data for cyano-pyrimidines **1**, **38**, **67**, **79,** and **90**

Cyano-pyrimidine **79** may potentially address Cdu1's Trp276 and Ser 274 (Figure 3.26 A), while **90** seems to be able to establish Sulphur—π interaction with Met262 (Figure 3.26 B). CatK inhibitors are of special interest, as they can be readily obtained from Novartis within the framework of our collaboration. However, they have a few key differences compared to the previously analyzed molecules. All adenain inhibitors (except for **5**) display meta-substitutions with respect to the cyano-pyrimidine ring, as para-substituted ligands have proven to be not optimal to reach the active site cleft in adenain (Altmann 2017, personal communication). CatK inhibitors differ in the sense that they are both meta and para-di-substituted, making them unable to inhibit adenain (Figure 3.27).

Adenain inhibitor

CatK inhibitor Dutacatib ®
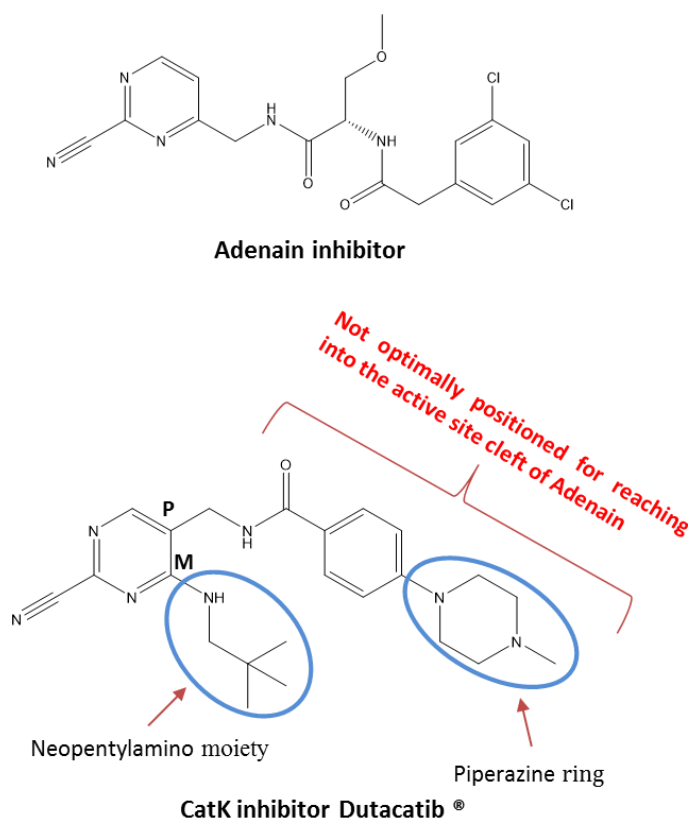
Figure 3.27. Meta-substituted Adenain inhibitor (cyano-pyrimidine **138**) and the di-substituted CatK inhibitor Dutacatib®, exemplifying the different substitution patterns on these ligands.

A lipophilic meta-substitution, such as the neopentylamino moiety of **79** (Figure 3.26 A) or the aniline group in **90** (Figure 3.26 B), is present in all Novartis CatK inhibitors. The para-substitution is variable, sometimes displaying a piperazine ring. The docking results for **79** and **90** suggested that the active cleft of Cdu1 might not only allocate these disubstituted ligands, but that the piperazine moiety could potentially address Cdu1's Ser168 (Figure 3.26 A).

It was noted that compounds **79** and **90** are part of Novartis patents WO 2003020278A1 (173) and WO 2004020441A1 (172) . These patents contain more than 200 cyano-pyrimidines, many of which are not part of the ZINC library. Therefore, the search for other promising cyano-pyrimidine Cdu1 inhibitors was continued by evaluating these molecules through our docking routine (a total of 370 molecules including protonated species were analyzed). After scoring and visual inspection of the top ranked ligands, cyano-pyrimidine **93** was requested to Novartis for enzymatic evaluation (Table 3.6).

	Lig #	CRS	ChemScore	HAN
		Ranking according to ChemScore		
1	5	13.00	34.070	18
2	230	11.37	33.240	25
3	232	11.37	32.790	24
4	93	10.86	34.480	32
5	231	10.83	32.490	27
		Ranking according to DSX_CSD		
	Lig #	CRS	DSX_CSD	HAN
1	93	42.05	133.500	32
2	5	39.82	104.350	18
3	230	36.21	105.880	25
4	232	35.36	102.000	24
5	231	33.30	99.890	27

Table 3.6 Docking data and % of inhibition of cyano-pyrimidines **93**, **230**, **231** and **232.** Data for **5** are shown for reference.

Cyano-pyrimidine **93** was similar to the originally identified **79**, except for having a meta-substitution pattern in the central benzene ring. It also lacks a carbon atom in between this ring and the terminal piperazine group, when compared to **79**. Covalent docking suggested that these modifications may permit the molecule to fit inside Cdu1's active site more appropriately than **79** (Figure 3.28 A).

Figure 3.28. (A) ChemScore pose for cyano-pyrimidine **93** in complex with Cdu1. **(B)** ChemScore poses for cyano-pyrimidines **230**, **231** and **232**. The experimental binding mode of **5** is displayed in orange.
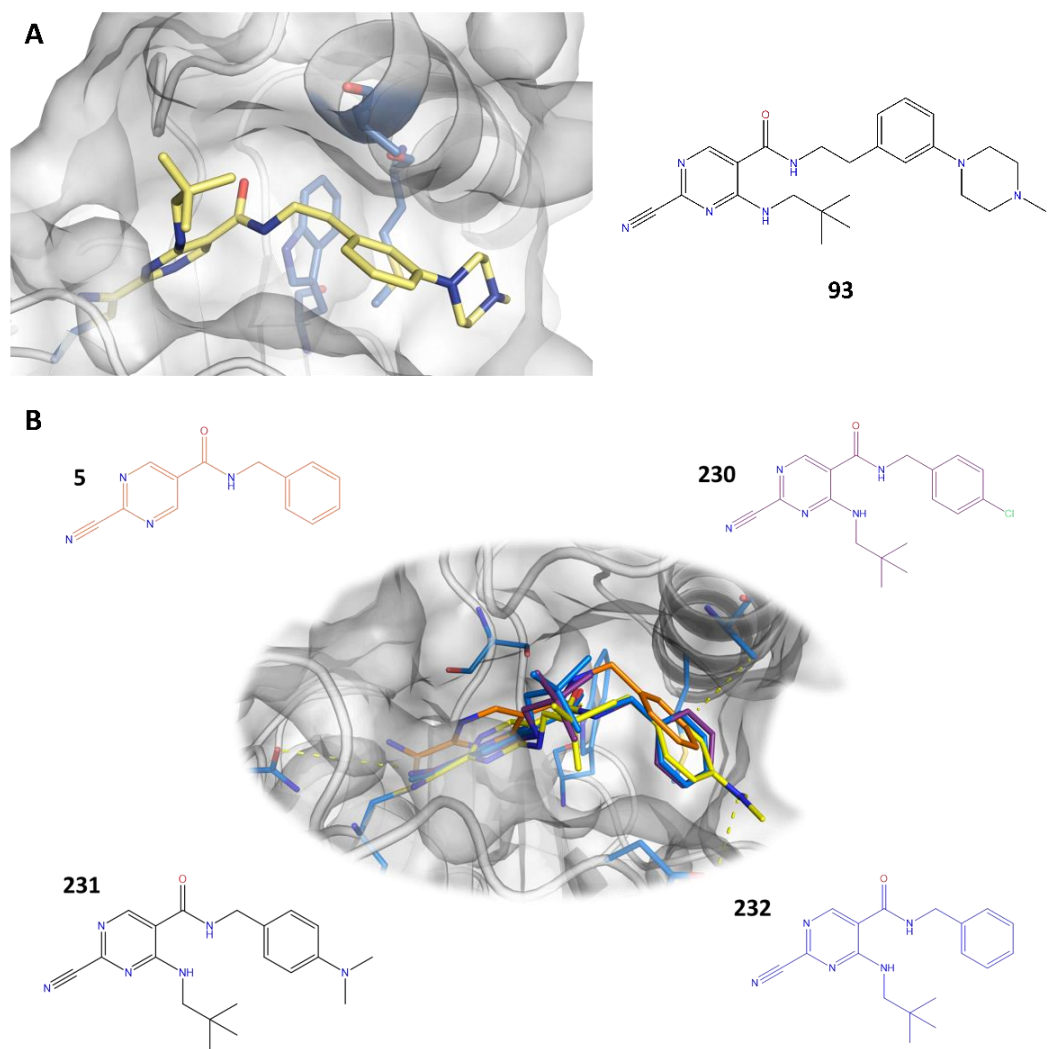
Additionally, cyano-pyrimidines **230**, **231** and **232** (Table 3.6) within these patents were requested on the basis of their structural similarity to the already known binder, cyano-pyrimidine **5**. Covalent docking suggested that these molecules may have the potential to bind Cdu1 in an analogous way as observed for **5** (Figure 3.28 B).

However, **93** showed no inhibition but rather an activating effect (Figure 3.29). The reasons behind this remain unclear, however, one may hypothesize that the ligand may enhance the activity of Cdu1 by stabilizing the protein upon binding to a site different to the active cleft. Cyano-pyrimidines **231** and **232** showed no effect (Table 3.6 and Figure 3.29), suggesting that the neopentylamino moiety might not be adequate for Cdu1 inhibition. Nevertheless, the fact that cyano-pyrimidine **230** showed an inhibitory effect suggests that the inclusion of a polarizable halogen group (in this case a chlorine) might compensate for the inclusion of the neopentylamino moiety to yield a positive effect on inhibition.

Lig #	% inhibition	IC$_{50}$ (µM)
5	74.7± 3.8	113
93	-117 ± 6.7	--
230	73 ± 4.3	--
231	17.3 ± 2.6	--
232	17.6 ± 2.7	--
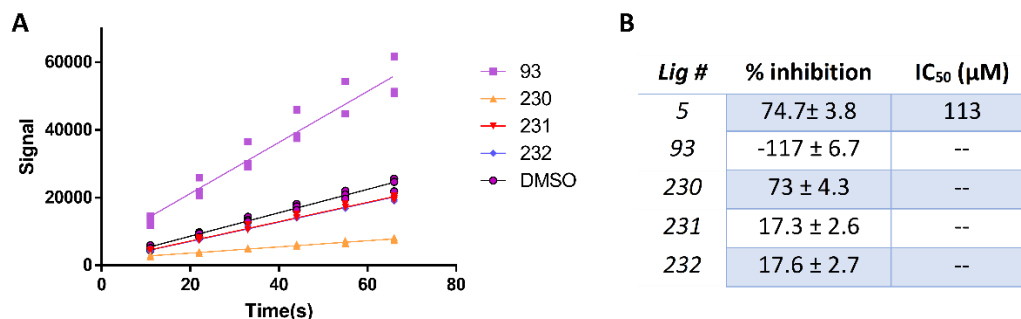
Figure 3.29 (A) Initial phase of Cdu1 substrate cleavage progression curve (Ub-Rh110Gly, t=120 seconds) after 1.5 h incubation with 100 µM compounds **93**, **230**, **231** and **232** (n=3). **(B)** % of inhibition cyano-pyrimidines **93**, **230**, **231** and **232**. Inhibition data for cyano-pyrimidine **5** is included for comparison.

3.9   Development of HJR108

3.9.1   Design of a suitable backbone for Cdu1 specific inhibition
Sections in between " " are taken from Ramirez et al, 2018 with permission

"A comparison of the Cdu1~ubiquitin complex with the inhibitor bound structures reveals that the peptidic backbone of cyano-pyrimidine **3** reaches into the active site of Cdu1 in a similar way as the C-terminus of ubiquitin (Figure 3.30 A). Both substrate and inhibitor address Cdu1 residues Gln338, Trp247 and Ser168, suggesting that these residues are hotspots for interactions. However, Gln338 is a conserved feature across the human SENP family while Trp247 is replaced either by an asparagine (SENP6, SENP7 and SENP8) or a histidine (SENP1, SENP2, SENP3, and SENP5) that can as well act as a hydrogen bond donor. The conserved nature of these residues might pose a challenge for the design of ligand specificity over other DUBs. This is especially the case for Ser168, where only the main chain nitrogen is required to establish an interaction. Therefore, addressing residues on the unique α-helix D of Cdu1 (i.e. VR-3) such as Met262, Ala266 and Gly269 may need to be preferentially taken into consideration for the design of new Cdu1 ligands. Interestingly, the side chain of Ub Arg74 and Leu73 are located at equivalent positions as the iso-pentyl substituent and the dimethylamine substituent in **3**, respectively.

Modifying the substitution pattern of 3 to mimic Arg74 and Leu73 may improve the ligand towards complementarity and yield an increased potency. However, the loop opening triggered by **3** (section 3.7.3) might represent a net energetic penalty which must first be compensated by the ligand upon binding. In contrast, the backbone of **5** (i.e., para-substituted with respect to the cyano-pyrimidine ring) keeps the active site closer to the Ub-bound conformation (Figure 3.30 B), while addressing Trp247 and Ala266. This closed conformation may favor the backbone of **5** to initiate Cdu1 inhibitor optimization efforts."
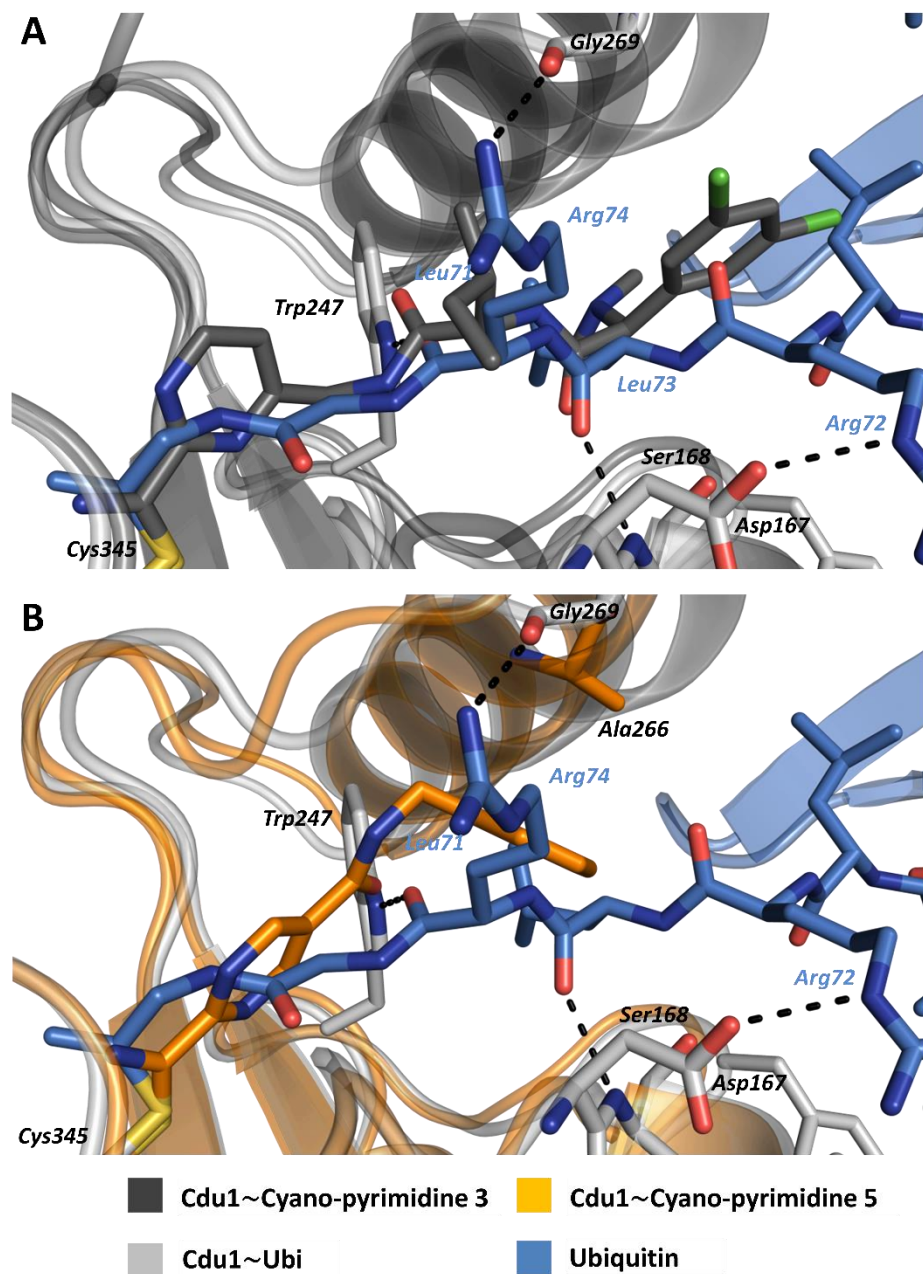
Figure 3.30. Superposition of the Cdu1~Ub complex with the Cdu1~cyano-pyrimidine **3** **(A)** and the Cdu1~cyano-pyrimidine **5 (B)** complex. Taken from Ramirez et al, 2018 with permission.

To design a Cdu1 specific ligand, the backbone of **5** was first improved in a way that it can address **Ser168**, analogous to **3**. When superimposing the experimental pose of **5** with the docking result of cyano-pyrimidine **141** (section 3.8.2), it was

observed that the central aromatic ring of **141** was positioned near the terminal ring of **5.** This observation prompted the merging of the ligands at this position to yield cyano-pyrimidine **106** (Figure 3.31).
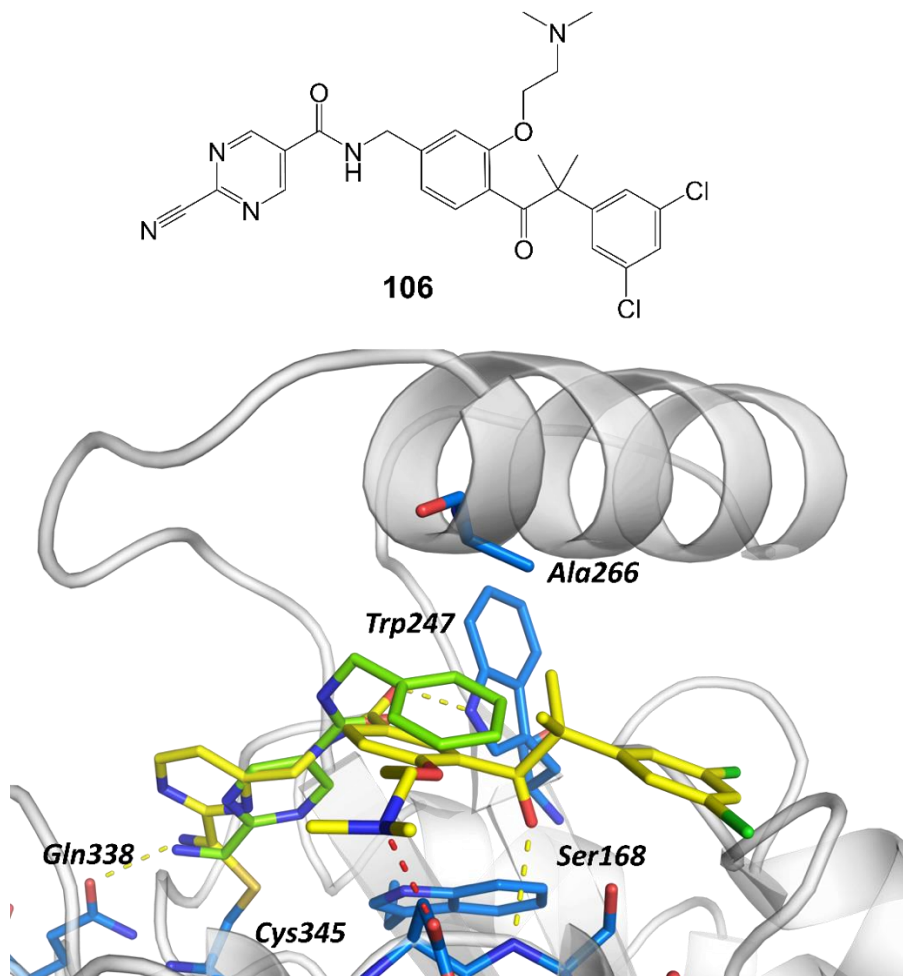
Figure 3.31. ChemScore pose for Cyano-pyrimidine **3** derivative **141** (yellow) superimposed with the experimental pose of **5** (green). This result led to the idea of merging the ligands yielding compound 106.

Since the dichloro-benzene ring of **3** was experimentally observed to be solvent exposed (Section 3.7.2), it was hypothesized that a single chlorine group in the para position could be more suitable for Cdu1's narrower binding site as compared to adenain. With this consideration in mind, a new set of ligands based on the backbone of **106** was designed (Figure 3.32). These molecules were evaluated by means of covalent docking along with the 104 cyano-pyrimidines retrieved from the

ZINC database, which also shared the backbone of **5.** It should be mentioned that cyano-pyrimidine **5** itself was included in this docking run as control and renamed as molecule 105.

CRS analysis revealed that seven of the designed ligands were among the best ranked when scoring with ChemScore. Rescoring using DSX_CSD resulted in four of the designed molecules being among the top ten best scored molecules (Table 3.7).
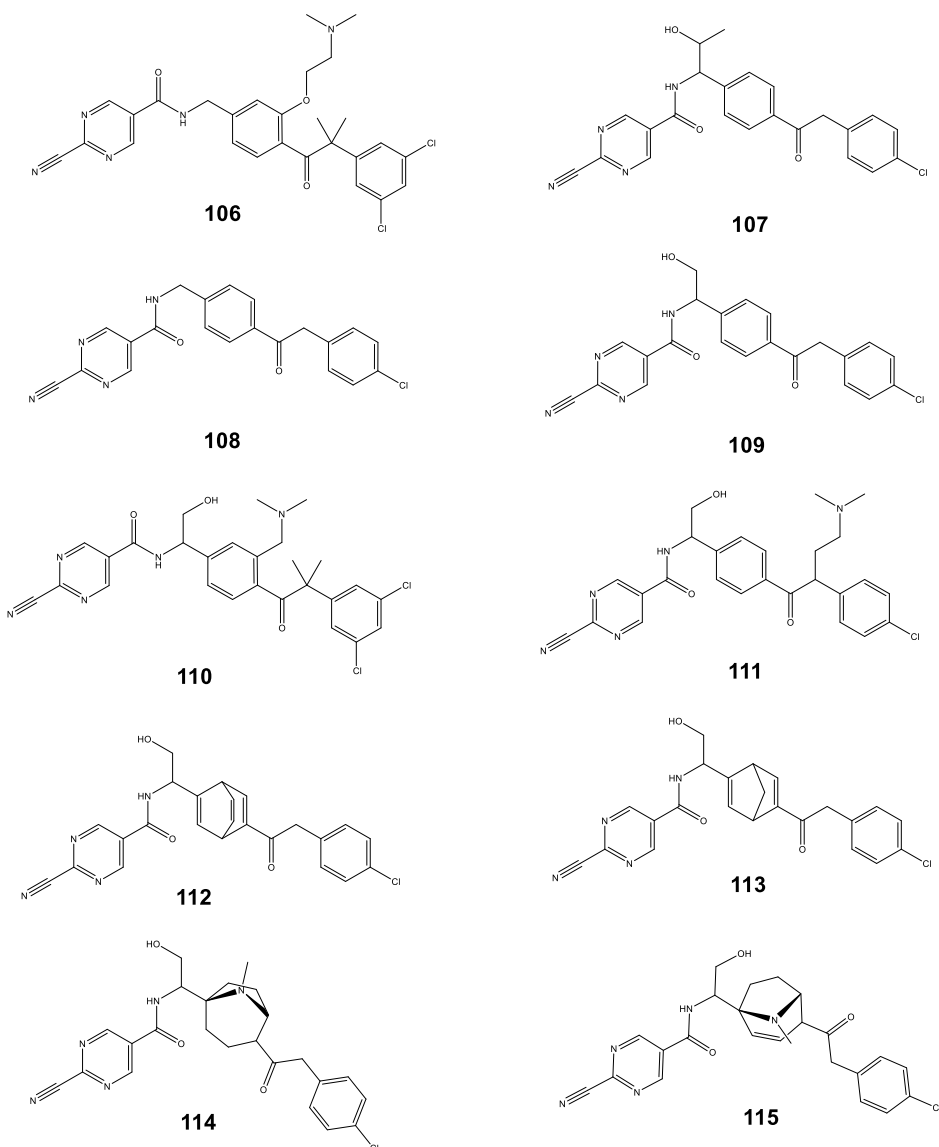
Figure 3.32. New series of ligands designed based on the backbone of cyano-pyrimidine **106.**

Remarkably, cyano-pyrimidine **108** had the best CRS in both scoring functions. Visual inspection revealed that this molecule combines the desirable features of both crystalized ligands while being able to address **Met262** through a sulphur---π interaction, a key residue located within Cdu1's α-helix D (Figure 3.33).

	Ranking according to ChemScore					Ranking according to DSX_CSD		
	Lig #	CRS	ChemScore	HAN		Lig #	CRS	DSX_CSD	HAN
1	108	14.69	44.610	28		108	51.87	157.50	28
2	109	14.13	43.920	30		38	49.63	177.81	46
3	111	13.78	45.060	35		83	49.62	159.16	33
4	112	13.77	43.720	32		39	48.77	174.74	46
5	113	13.56	42.590	31		109	48.51	150.73	30
6	106	13.56	45.190	37		79	47.37	150.39	32
7	110	13.29	43.880	36		87	47.25	150.01	32
8	105 (**5**)	13.18	34.550	18		106	46.80	155.95	37
9	90	13.06	41.030	31		110	46.22	152.6	36
10	95	12.94	39.77	29		82	46.07	147.77	33

Table 3.7 Docking data for the new series of designed ligands and ZINC cyano-pyrimidines. The top ten poses according to ChemScore and DSX_CSD (CRS score) are shown for cyano-pyrimidines **93**, **230**, **231** and **232.** Data for **5** are shown for reference.
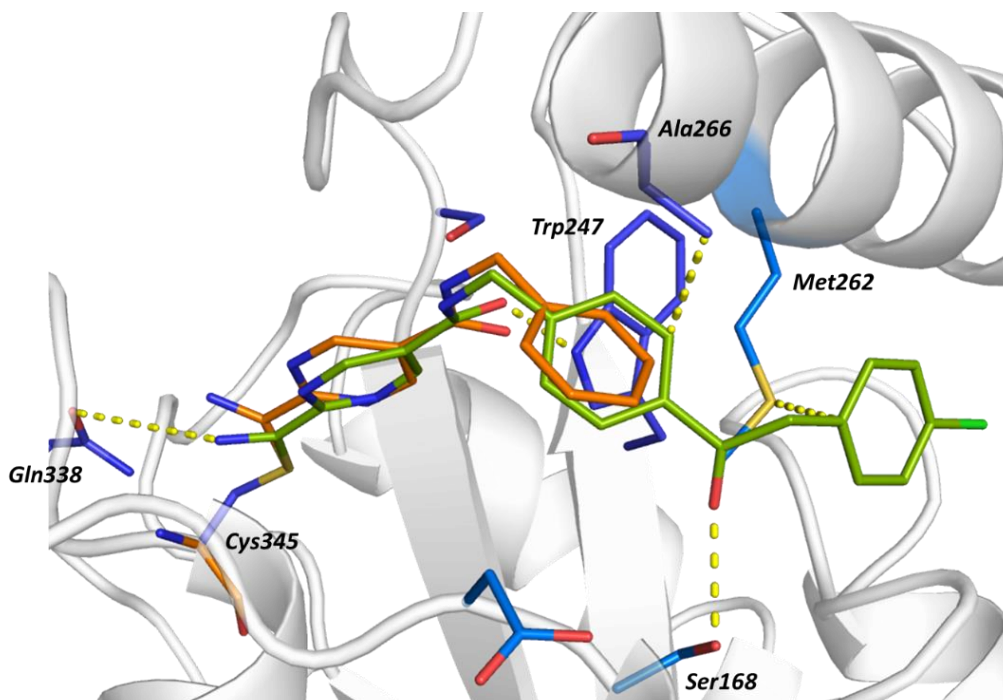
Figure 3.33. Top ChemScore pose for **108** (green) in complex with Cdu1. The experimental pose of **5** (orange) is shown for reference.

Other examples of high-scored designed ligands include cyano-pyrimidine **109,** which is the second and fifth best ranked compound according to ChemScore and DSX_CSD, respectively. This ligand appears to be able to address the backbone of **Ser270** due to a CH$_2$OH substituent (Figure 3.34 A). As this residue is located within the flexible loop between α-helix D and β-strand 2, this interaction could aid the ligand to favor the closed conformation of Cdu1. Similarly, cyano-pyrimidine **113** where the central aromatic ring has been replaced by a Norbornadiene moiety, could potentially be advantageous with respect to cavity filling (Figure 3.34 B).
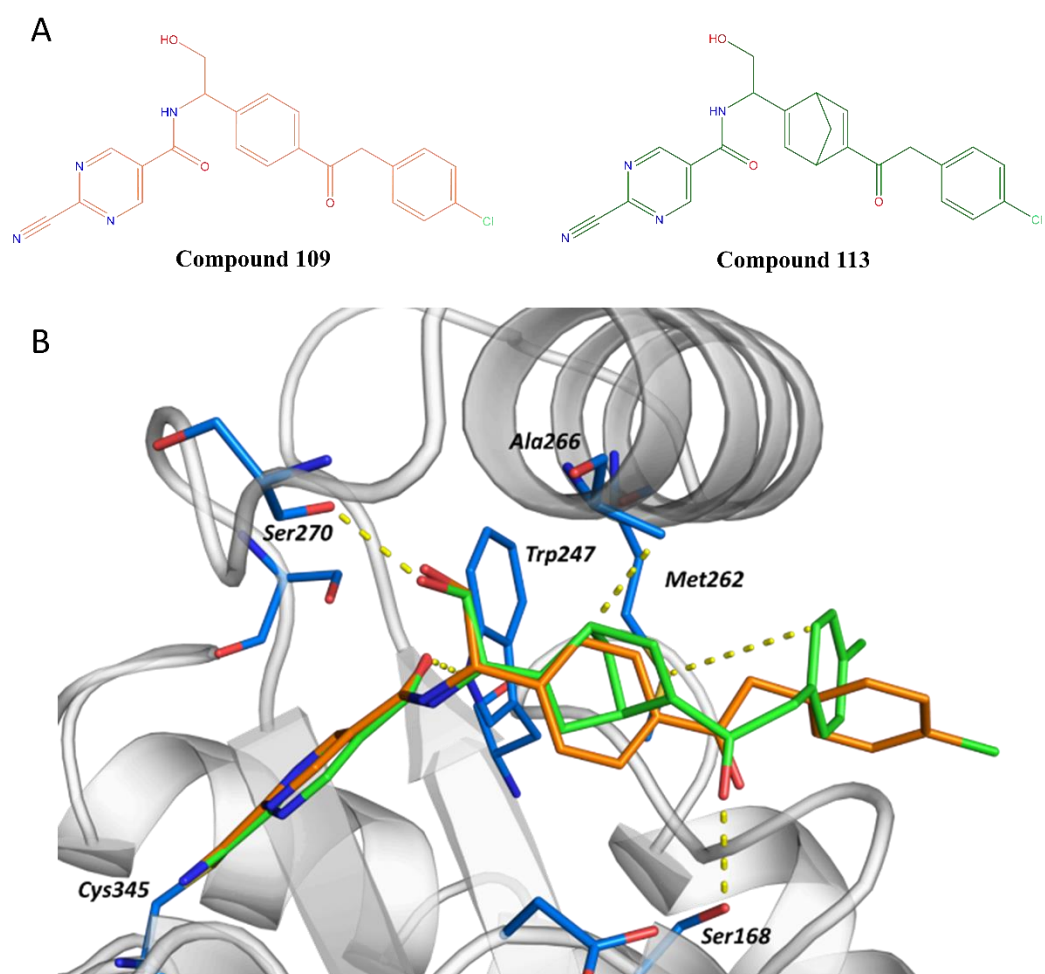
Figure 3.34. (A) Structure of the proposed Cdu1 covalent binders **109** and **113 (B)** Top ChemScore pose for compound **109** (orange) and **113** (green) in complex with Cdu1.

3.9.2 The HJR108 series

Despite being promising Cdu1 inhibitors, molecules like **109** and **113** might represent a challenge from the synthetic point of view. The presence of a leaving group at the terminal benzene ring of both molecules (i.e., the chorine atom) may facilitate undesired aromatic nucleophilic substitutions during synthesis and lead to side products. The synthesis of the norbornadiene moiety of **113,** a bicyclic hydrocarbon, may be too laborious for a backbone of which the effectiveness has not yet been tested experimentally. In contrast, **108** displays a satisfactory shape complementarity within the active cleft of Cdu1 while remaining a relatively simple molecule. Removal of the chlorine atom at the terminal benzene ring further simplifies the molecule. Covalent docking of this simplified backbone, designated as **HJR108** (Figure 3.35 A), revealed that this modification does not affect the docking pose (figure 3.35 B).
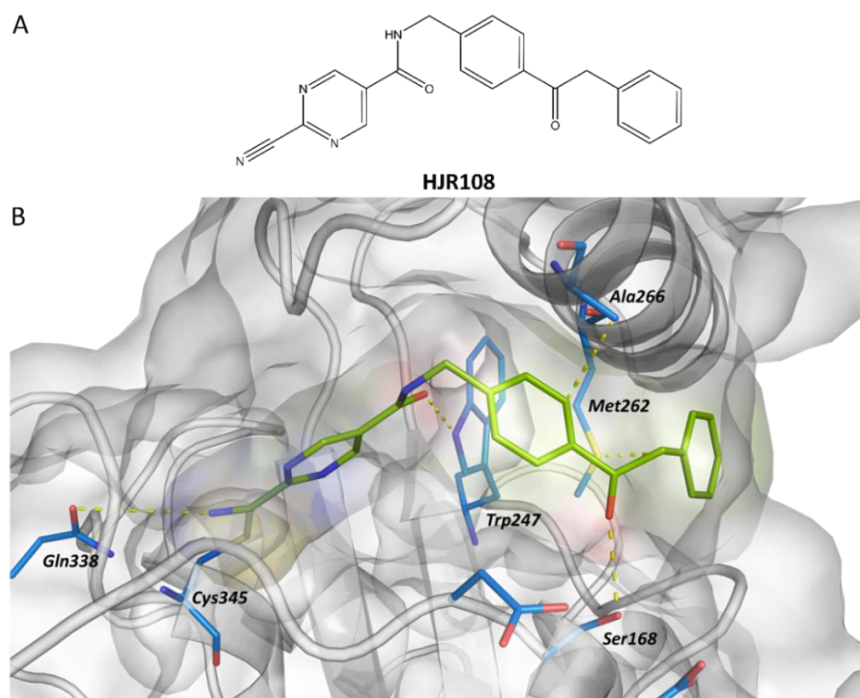
Figure 3.35. (A) Structure of proposed Cdu1 lead molecule **HJR108 (B)** Top ChemScore pose for **HJR108** in complex with Cdu1 showing adequate shape complementarity within the active site cleft.

To explore the effect of derivatization of the backbone of **HJR108**, a new series of molecules was generated by modifying the substitution pattern of this ligand. This set of molecules were designed in the framework of a research internship by Gerhard Keller (Keller, 2017) and designated as the **HJR108 series** (Figure 3.36).

HJR_108

HJR_108a

HJR_108g

HJR_108b

HJR_108h

HJR_108c

HJR_108i

HJR_108d

HJR_108j

HJR_108e

HJR_108k

HJR_108f

HJR_108I
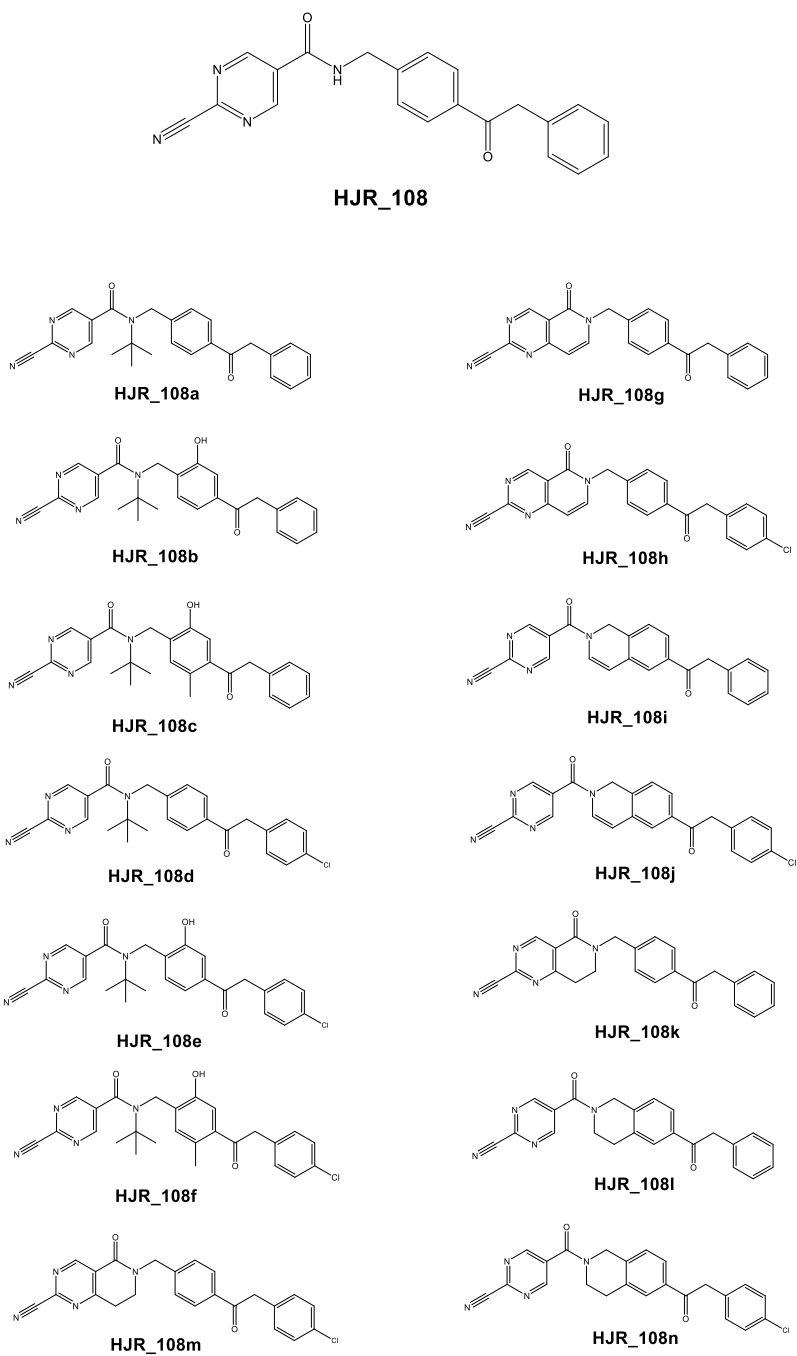
HJR_108m

HJR_108n

Figure 3.36. Possible cyano-pyrimidine derivatives of **HJR108** for Cdu1 inhibition (HJR108 series)

Subsequently, a new docking round including the analogs of **5** from the ZINC library, the newly designed ligands (Figure 3.32) and the HJR108 series cyano-pyrimidines, was performed. Nine of the ligands from the HJR108 series were among the ChemScore top ten, with HJR108 itself having the best CRS among all ligands and molecule 108 taking the position of the third best molecule. Six molecules of this series were among the ten best molecules in the DSX_CSD (Table 3.8).

	Ranking according to ChemScore				Ranking according to DSX_CDS			
	Lig #	CRS	ChemScore	HAN	Ligand	CRS	DSX_CSD	HAN
1	HJR_108	14.89	44.66	27	HJR_108g	54.10	166.21	29
2	HJR_108c	14.85	47.630	33	HJR_108h	53.20	167.13	31
3	108	14.69	44.610	28	HJR_108m	52.72	163.82	30
4	HJR_108g	14.65	45.010	29	HJR_108k	52.66	161.79	29
5	HJR_108a	14.56	45.740	31	108	51.26	157.50	29
6	HJR_108h	14.39	45.220	31	38	49.63	177.81	46
7	HJR_108i	14.37	44.150	29	83	49.62	159.16	33
8	HJR_108k	14.29	43.900	29	39	48.77	174.74	46
9	HJR_108e	14.31	45.890	33	HJR_108n	48.55	150.86	30
10	HJR_108b	14.17	44.980	32	HJR_108l	48.04	147.58	29

Table 3.8 Data for the HJR108 series docked along with the new series of designed ligands and ZINC cyano-pyrimidines. The top ten poses according to ChemScore and DSX_CSD (CRS score) are shown.

Other high-scored ligands include **HJR108c** (second best ChemScore) and **HJR108g** (best scored DSX_CSD). The idea towards the design of this molecules was to improve the cavity filling by derivatizing the nitrogen atom on the amide group of HJR108 with either an isobutyl moiety (Figure 3.37 A) or by creating a bicyclic system with the pyrimidine ring (Figure 3.37 B). The inclusion of these bulky substituents may also impede the motion of the carbonyl group towards the solvent and promote its interaction with Trp247. Additionally, making the central ring of **HJR108c** ortho-substituted with an OH group, may allow the molecule to address the main chain carbonyl of Met262 (Figure 3.37 A).

A
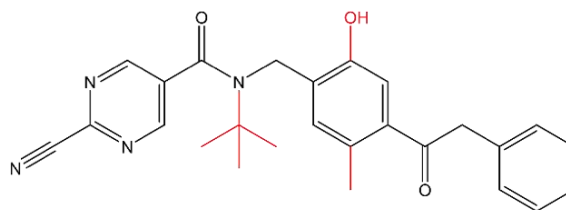
HJR_108c
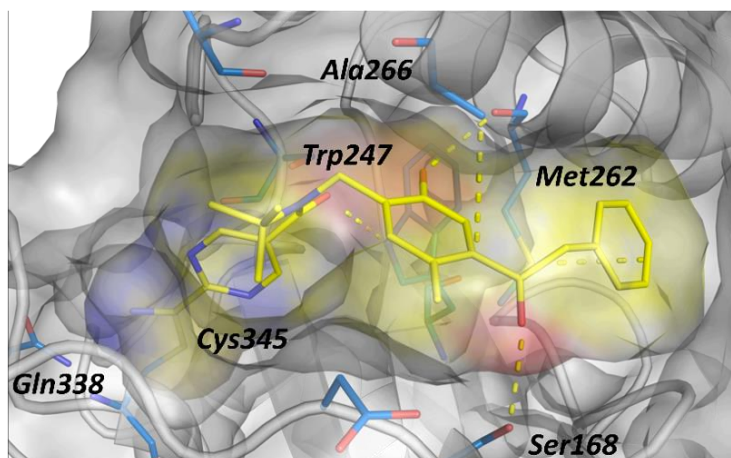
B
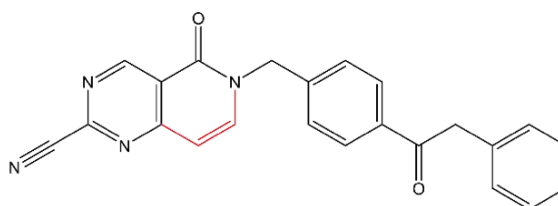
HJR_108g
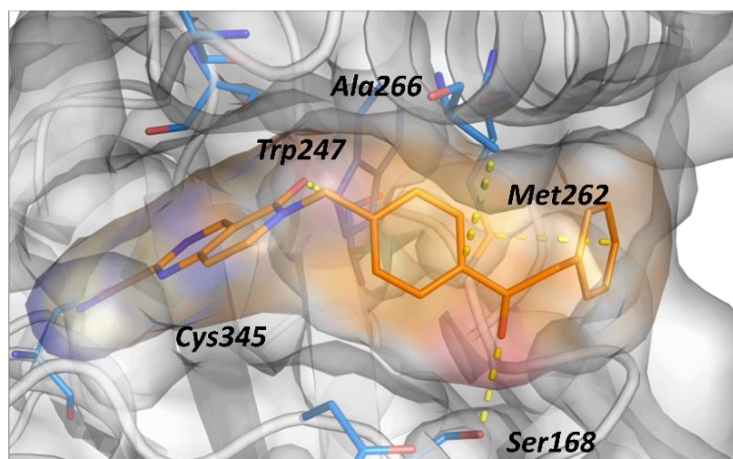
Figure 3.37. Top ChemScore pose for **HJR108c (A)** and **HJR108g (B)** in complex with Cdu1. Derivatizations with respect to HJR108 are highlighted in red (top panel).

3.9.3 Synthesis, IC$_{50}$ determination, and reversibility assay of HJR108

Among all designed ligands, **HJR108** was selected for synthesis based on its simplicity and satisfactory docking data. I proposed an initial synthesis description which was improved by Stefan Kletke, who started to work on the generation of HJR108 in the framework of a research collaboration with the laboratory of Prof. Dr Tanja Schirmeister (University of Mainz). The synthesis was later completed in the same lab by Kevin Schwickert.

The IC$_{50}$ for **HJR108** was determined utilizing the same conditions as for cyano-pyrimidines **3** and **5** (section 3.6.2). The obtained value of 19 μM represents a ten-fold improvement with respect to **3**, a molecule with a similar size to **HJR108** (Figure 3.38 A). The nearly six-fold improvement with respect to **5** indicates that the addition of the carbonyl moiety, followed by a benzene ring, was indeed beneficial as suggested by covalent docking. It was thus demonstrated that a substantial improvement of the potency of **5** could be achieved with the addition of only nine well placed heavy atoms.
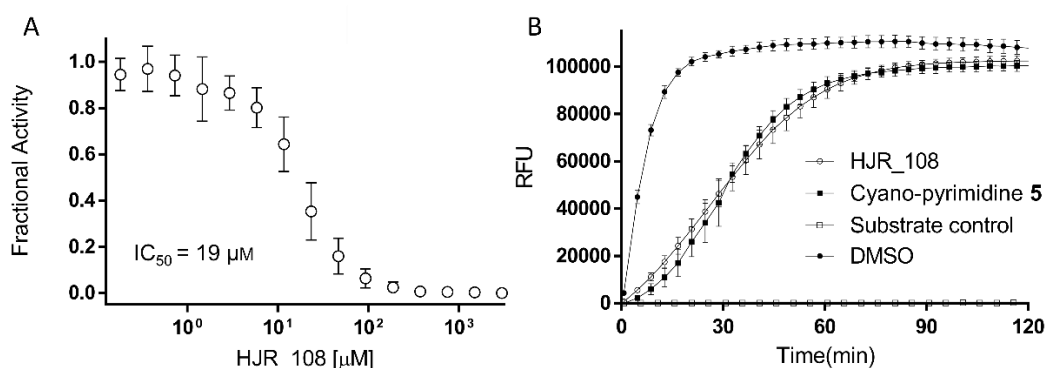
Figure 3.38. (A) Concentration-response curve for Cdu1 (125-401) activity after 1 hour of incubation with **HJR108**. The data represent the mean of at least 3 independent experiments performed in duplicate. **(B)** Rapid dilution assay. Time course of substrate cleavage: **HJR108** (n$_{rep}$=6); cyano-pyrimidine **5** (n$_{rep}$=6); substrate control (n$_{rep}$=3). The DMSO control shows the uninhibited reaction (n$_{rep}$ = 6). Error bars indicate the standard error of the mean.

Finally, the rapid dilution assay demonstrated a reversible mechanism for **HJR108**, comparable to **5** (Figure 3.38 B). Altogether, **HJR108** has been theoretically and experimentally demonstrated to be the first covalent inhibitor designed to address the active site of Cdu1, and that the improvement of its backbone may allow the structure-based optimization of even tighter binding and specific ligands.

3.9.4 Primary sequence alignment of Cdu1 and Cdu2 suggests a common inhibition mechanism

A primary sequence alignment of Cdu1 and Cdu2 revealed that the key residues which interact with **3** and **5** in Cdu1 are also present at equivalent positions in Cdu2, except for Ser168, which is replaced by an asparagine. Met262, predicted by covalent docking to stabilize the binding of HJR108 to Cdu1, is also present on Cdu2 (Figure 3.39). Although the crystal structure of Cdu2 is not yet available, this could be an indication that the three-dimensional arrangement of these residues is similar to the one from Cdu1. If this is the case, the Cdu1 inhibitors molecules could potentially inhibit Cdu2 as well.

```
                                                                      Ser168
CDU1_C. Trachomatis    118   STPPPLLPKAPKPVKTQEDLLPLVPEQVFVEMYEDMARRQTIEALVPAWDSDIIFKCLCY
CDU2_C. Trachomatis     78   ----------------------------QQ----LDDVLTGPPLPIWDNEHLFQFSCL
                                                                      :* **.: :*: *

CDU1_C. Trachomatis    178   FHTLYPGLIPLETFPPATIFNFKQKIISILEDKKAVLRGEP--IKGPLPICCSKENYRRH
CDU2_C. Trachomatis    104   MQNKHRRVLPIDICNPLTKFNFLECICNCLMTKQSVNVNETDMCELFCPPTCTPENYRRL
                             ::. : ::*:: * * *** : .* *:* .* : * *: *****
                                 Trp247         Met262      His275          ASP292
CDU1_C. Trachomatis    236   LQRTTLLPVFNWYHPTPKTLSDTMQTMKQLAIKGSVGASHWLLVIVDIQARRLVYFDSLY
CDU2_C. Trachomatis    164   LCTSSVFPFVNWHDPSADTQEAMLTKMDQTMSSGRVGNSHWVLVIVDIEYRCVTFFDSLC
                             * :::::*..**:.*: .* . : .* .* ** ***.******: * .:.****
                                                                      Trp276
                                                             Gln338       Cys345
CDU1_C. Trachomatis    296   NYVMPPENMKKELQSFAQQLDQVYPAYDSK---------KFSVKIAAKEVIQRGSGSSCG
CDU2_C. Trachomatis    224   DYVASPQQMREQLEGLAVSLGAIYPKEGGADSDQEELLSPFQVRIGSTVKVQSPGEFTCG
                             :** *::*:::*::* .* .** .. *.*:*.: :* . :**

CDU1_C. Trachomatis    347   AWCCQFLHWYLKDPLTDALNDLPVDSVERHENLASFVQACEAAVQDLPELSWPEA-
CDU2_C. Trachomatis    284   AWCCQFLAWYLENPDFDLEEKVPTNPSERRALLADFISTTEQAMSRYSSLSWPTTD
                             ******* ***::* * :::*.: **: **.*:.: * *:. .**** :
```

Figure 3.39. Sequence alignment of the protease domain of Cdu1 and Cdu2. Known protein-inhibitor interaction hotspots for Cdu1 (indicated in red) appear to be conserved in Cdu2.

4   C ONCLUSIONS AND OUTLOOK

Sections in between " " are taken from Ramirez et al, 2018 with permission

Cdu1 has been demonstrated to be a highly versatile protease able to address diverse substrates including ubiquitin and NEED8 (94), Mcl1 (90), GLU1 (95) and acetyl-CoA (93) thus acting as a deubiquitinase as well as an acetyltransferase. However, given the versatility of the CE fold (67), this could be only the tip of the iceberg regarding the moonlighting function of Cdu1. One may hypothesize that this enzyme addresses other cellular targets to improve the pathogenicity of *Chlamydia*. To further understand the cellular function of Cdu1, additional interaction partners should be identified utilizing techniques such as co-immunoprecipitation and mass spectrometry analysis. The protein domains relevant for the recognition of these interaction partners should be further explored using peptide array analysis, which would enable the detailed characterization of these interfaces via microscale thermophoresis, isothermal titration calorimetry and X-ray crystallography. The same may be done for Cdu2, which like Cdu1, is involved in the fragmentation of the host's cell Golgi apparatus (93). Moreover, the conserved primary sequence features between Cdu1 and Cdu2 suggest that the covalent Cdu1 inhibitors described in this thesis, might inhibit Cdu2 as well which should be experimentally validated.

''Exploiting the high active site homology between Cdu1 and the evolutionarily related protease adenain by utilizing cross-target screening of protease inhibitors permitted the identification of the first two small molecules that bind to the active site of Cdu1. This target-hopping strategy resulted in the successful repositioning of inhibitors and has delivered unprecedented starting points for Cdu1 inhibition. Our crystal structures of Cdu1 in complex with the inhibitors as well as the ubiquitin bound structure (PDB: 6FDK, 6FDU, 6FDQ) provide insights into the details of Cdu1 ligand recognition and binding pocket flexibility. The interaction network of the C-terminus of ubiquitin and of cyano-pyrimidines **3** and **5** indicate

that cyano-pyrimidine ligands do not only covalently bind to Cys345 but have the potential to be optimized to potent covalent inhibitors for this bacterial de-ubiquitylase by also addressing other hotspots for interactions in Cdu1's active site such as Asp167, Ser168, Trp247 and Gln338. Specificity for this bacterial DUB could be achieved by targeting residues located at the chlamydial α-helical insert of Cdu1 such as Met262, Ala266 and Gly269.''

Our covalent docking model readily enabled the *in-silico* optimization of Cdu1 covalent binders towards **HJR108**, the first reversible inhibitor designed to address the active site of Cdu1. The superiority of this ligand with respect to Cyano-pyrimidines **3** and **5** might be due to its ability to address Met262, as suggested by covalent docking. Experimental confirmation of this interaction through X-ray crystallography should be pursued.

Further inhibitor optimization might be pursued by synthetizing other promising members of the HJR108 series and evaluating their effect on Cdu1, both enzymatically and structurally. Special efforts should be put in the determination of the kinetic parameters of the covalent and non-covalent contributions of the inhibitors upon binding (time-dependent potency). Together, this data would help to further develop the HJR108 backbone into more potent and specific ligands. The specificity of the designed ligands should be tested, however, against a broader panel of DUBs. Special attention should be put into the members of the human SENP family (such as SENP8), given the conserved active site features that Cdu1 shares with these proteins.

Beyond covalent labeling of the active site cysteine, allosteric sites of Cdu1 may be addressed either by classic non-covalent molecules or by covalent targeting of non-active site cysteines. This could be an alternative to inhibit the enzyme by inducing conformational changes that would disturb its function. Alternatively, inhibition may also be achieved if some of these allosteric sites are located within protein interfaces relevant for substrate recognition.

Finally, other effector proteases should be studied as well to further unveil the mechanics of DUB-mediated pathogenicity in other pathogenic bacteria. A good starting point for this may be other proteins which belong to the CE clan and which are involved in disease progression (for instance, SseL from *Salmonella entercia* and RavZ from *Legionella pneumophila*). The conserved active-site topology within the CE clan may enable the possibility to identify polypharmacological covalent inhibitors which are active in diverse disease pathways.

5 BIBLIOGRAPHY

1.   Everett KDE, Bush RM, Andersen AA. Emended description of the order Chlamydiales, proposal of Parachlamydiaceae fam. nov. and Simkaniaceae fam. nov., each containing one monotypic genus, revised taxonomy of the family Chlamydiaceae, including a new genus and five new species, and standards for the identification of organisms. International Journal of Systematic and Evolutionary Microbiology. 1999;49(2):415-40.

2.   Gordon FB, Quan AL. OCCURENCE OF GLYCOGEN IN INCLUSIONS OF THE PSITTACOSIS-LYMPHOGRANULOMA VENEREUM-TRACHOMA AGENTS. J Infect Dis. 1965;115:186-96.

3.   Grayston JT, Kuo CC, Wang SP, Altman J. A new Chlamydia psittaci strain, TWAR, isolated in acute respiratory tract infections. The New England journal of medicine. 1986;315(3):161-8.

4.   Karyagina AS, Alexeevsky AV, Spirin SA, Zigangirova NA, Gintsburg AL. Effector proteins of chlamydiae. Molecular Biology. 2009;43(6):897-916.

5.   Gerard HC, Whittum-Hudson JA, Carter JD, Hudson AP. Molecular biology of infectious agents in chronic arthritis. Rheumatic diseases clinics of North America. 2009;35(1):1-19.

6.   Kern JM, Maass V, Maass M. Molecular pathogenesis of chronic Chlamydia pneumoniae infection: a brief overview. Clinical Microbiology and Infection. 2009;15(1):36-41.

7.   Littman AJ, Jackson LA, Vaughan TL. Chlamydia pneumoniae and lung cancer: epidemiologic evidence. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2005;14(4):773-8.

8.   Hahn DL, Dodge RW, Golubjatnikov R. Association of Chlamydia pneumoniae (strain TWAR) infection with wheezing, asthmatic bronchitis, and adult-onset asthma. Jama. 1991;266(2):225-30.

9.   Kleemola M, Saikku P, Visakorpi R, Wang SP, Grayston JT. Epidemics of pneumonia caused by TWAR, a new Chlamydia organism, in military trainees in Finland. J Infect Dis. 1988;157(2):230-6.

10.   Shima K, Kuhlenbaumer G, Rupp J. Chlamydia pneumoniae infection and Alzheimer's disease: a connection to remember? Medical microbiology and immunology. 2010;199(4):283-9.

11.   Assar O, Nejatizadeh A, Dehghan F, Kargar M, Zolghadri N. Association of Chlamydia pneumoniae Infection With Atherosclerotic Plaque Formation. Global journal of health science. 2015;8(4):260-7.

12.   Newman L, Rowley J, Vander Hoorn S, Wijesooriya NS, Unemo M, Low N, et al. Global Estimates of the Prevalence and Incidence of Four Curable Sexually Transmitted Infections in 2012 Based on Systematic Review and Global Reporting. PLOS ONE. 2015;10(12):e0143304.

13.   Peipert JF. Genital Chlamydial Infections. New England Journal of Medicine. 2003;349(25):2424-30.

14.   Plummer FA, Simonsen JN, Cameron DW, Ndinya-Achola JO, Kreiss JK, Gakinya MN, et al. Cofactors in Male-Female Sexual Transmission of Human Immunodeficiency Virus Type 1. The Journal of Infectious Diseases. 1991;163(2):233-9.

15.   Andrews WW, Goldenberg RL, Mercer B, Iams J, Meis P, Moawad A, et al. The Preterm Prediction Study: association of second-trimester genitourinary chlamydia infection with subsequent spontaneous preterm birth. American journal of obstetrics and gynecology. 2000;183(3):662-8.

16.   Owusu-Edusei K, Jr., Chesson HW, Gift TL, Tao G, Mahajan R, Ocfemia MC, et al. The estimated direct medical cost of selected sexually transmitted infections in the United States, 2008. Sexually transmitted diseases. 2013;40(3):197-201.

17.   WHO. WHO Alliance for the Global Elimination of Blinding Trachoma by the year 2020. Progress report on elimination of trachoma. Releve epidemiologique hebdomadaire. 2014;89(39):421-8.

18.   Kapoor G, Saigal S, Elongavan A. Action and resistance mechanisms of antibiotics: A guide for clinicians. Journal of anaesthesiology, clinical pharmacology. 2017;33(3):300-5.

19.   WHO. WHO Guidelines for the Treatment of Chlamydia trachomatis. Guidelines for the Treatment of Chlamydia trachomatis Geneva: World Health Organization Available from: https://wwwncbinlmnihgov/books/NBK379707/. 2016.

20.   Yoneyama H, Katsumata R. Antibiotic resistance in bacteria and its future for novel antibiotic development. Bioscience, biotechnology, and biochemistry. 2006;70(5):1060-75.

21.   Wise R. A review of the mechanisms of action and resistance of antimicrobial agents. Canadian respiratory journal. 1999;6 Suppl A:20a-2a.

22.   Fischbach MA, Walsh CT. Antibiotics for emerging pathogens. Science (New York, NY). 2009;325(5944):1089-93.

23.   Roberts MC. Update on acquired tetracycline resistance genes. FEMS microbiology letters. 2005;245(2):195-203.

24.   Sandoz KM, Rockey DD. Antibiotic resistance in Chlamydiae. Future microbiology. 2010;5(9):1427-42.

25.   Andersen AA, Rogers, D. G., editor Resistance to Tetracycline and Sulfadiazine in Swine C. Trachomatis Isolates. International symposium; 9th,

Human chlamydial infection; 1998; Napa, California, USA: International Chlamydia Symposium.

26.   Dugan J, Rockey DD, Jones L, Andersen AA. Tetracycline resistance in Chlamydia suis mediated by genomic islands inserted into the chlamydial inv-like gene. Antimicrobial agents and chemotherapy. 2004;48(10):3989-95.

27.   Lau SK, Wong GK, Li MW, Woo PC, Yuen KY. Distribution and molecular characterization of tetracycline resistance in Laribacter hongkongensis. The Journal of antimicrobial chemotherapy. 2008;61(3):488-97.

28.   Suchland RJ, Sandoz KM, Jeffrey BM, Stamm WE, Rockey DD. Horizontal Transfer of Tetracycline Resistance among Chlamydia spp. In Vitro. Antimicrobial Agents and Chemotherapy. 2009;53(11):4604-11.

29.   Unemo M, Shafer WM. Antimicrobial resistance in Neisseria gonorrhoeae in the 21st century: past, evolution, and future. Clinical microbiology reviews. 2014;27(3):587-613.

30.   Wi T, Lahra MM, Ndowa F, Bala M, Dillon J-AR, Ramon-Pardo P, et al. Antimicrobial resistance in Neisseria gonorrhoeae: Global surveillance and a call for international collaborative action. PLOS Medicine. 2017;14(7):e1002344.

31.   Stamm WE, Guinan ME, Johnson C, Starcher T, Holmes KK, McCormack WM. Effect of treatment regimens for Neisseria gonorrhoeae on simultaneous infection with Chlamydia trachomatis. The New England journal of medicine. 1984;310(9):545-9.

32.   Lin JS, Donegan SP, Heeren TC, Greenberg M, Flaherty EE, Haivanis R, et al. Transmission of Chlamydia trachomatis and Neisseria gonorrhoeae among men with urethritis and their female sex partners. J Infect Dis. 1998;178(6):1707-12.

33.   Komander D, Rape M. The ubiquitin code. Annual review of biochemistry. 2012;81:203-29.

34.   Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature. 1998;392(6676):605-8.

35.   Newton K, Matsumoto ML, Wertz IE, Kirkpatrick DS, Lill JR, Tan J, et al. Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. Cell. 2008;134(4):668-78.

36.   Wickliffe KE, Williamson A, Meyer HJ, Kelly A, Rape M. K11-linked ubiquitin chains as novel regulators of cell division. Trends in cell biology. 2011;21(11):656-63.

37.   Walczak H, Iwai K, Dikic I. Generation and physiological roles of linear Ubiquitin chains2012. 23 p.

38.   Spence J, Sadis S, Haas AL, Finley D. A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. Molecular and cellular biology. 1995;15(3):1265-73.

39.   Tenekeci U, Poppe M, Beuerlein K, Buro C, Muller H, Weiser H, et al. K63-Ubiquitylation and TRAF6 Pathways Regulate Mammalian P-Body Formation and mRNA Decapping. Mol Cell. 2016;62(6):943-57.

40.   Yau RG, Doerner K, Castellanos ER, Haakonsen DL, Werner A, Wang N, et al. Assembly and Function of Heterotypic Ubiquitin Chains in Cell-Cycle and Protein Quality Control. Cell. 2017;171(4):918-33.e20.

41.   Popovic D, Vucic D, Dikic I. Ubiquitination in disease pathogenesis and treatment. Nature Medicine. 2014;20:1242.

42.   Huang X, Dixit VM. Drugging the undruggables: exploring the ubiquitin system for drug development. Cell research. 2016;26:484.

43.   Hideshima T, Richardson P, Chauhan D, Palombella VJ, Elliott PJ, Adams J, et al. The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. Cancer research. 2001;61(7):3071-6.

44.   Chauhan D, Hideshima T, Mitsiades C, Richardson P, Anderson KC. Proteasome inhibitor therapy in multiple myeloma. Molecular cancer therapeutics. 2005;4(4):686-92.

45.   Kortuem KM, Stewart AK. Carfilzomib. Blood. 2013;121(6):893-7.

46.   Groll M, Kim KB, Kairies N, Huber R, Crews CM. Crystal Structure of Epoxomicin:20S Proteasome Reveals a Molecular Basis for Selectivity of α',β'-Epoxyketone Proteasome Inhibitors. Journal of the American Chemical Society. 2000;122(6):1237-8.

47.   Sakamoto KM, Kim KB, Kumagai A, Mercurio F, Crews CM, Deshaies RJ. Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. Proceedings of the National Academy of Sciences. 2001;98(15):8554-9.

48.   Winter GE, Buckley DL, Paulk J, Roberts JM, Souza A, Dhe-Paganon S, et al. Phthalimide conjugation as a strategy for in vivo target protein degradation. Science (New York, NY). 2015;348(6241):1376-81.

49.   Lai AC, Crews CM. Induced protein degradation: an emerging drug discovery paradigm. Nat Rev Drug Discov. 2017;16(2):101-14.

50.   Komander D, Clague MJ, Urbe S. Breaking the chains: structure and function of the deubiquitinases. Nat Rev Mol Cell Biol. 2009;10(8):550-63.

51.   Kwasna D, Abdul Rehman SA, Natarajan J, Matthews S, Madden R, De Cesare V, et al. Discovery and Characterization of ZUFSP/ZUP1, a Distinct Deubiquitinase Class Important for Genome Stability. Mol Cell. 2018;70(1):150-64.e6.

52.   Nijman SM, Luna-Vargas MP, Velds A, Brummelkamp TR, Dirac AM, Sixma TK, et al. A genomic and functional inventory of deubiquitinating enzymes. Cell. 2005;123(5):773-86.

53.   Murtaza M, Jolly LA, Gecz J, Wood SA. La FAM fatale: USP9X in development and disease. Cellular and molecular life sciences : CMLS. 2015;72(11):2075-89.

54.   Homan CC, Kumar R, Nguyen LS, Haan E, Raymond FL, Abidi F, et al. Mutations in USP9X are associated with X-linked intellectual disability and disrupt neuronal cell migration and growth. American journal of human genetics. 2014;94(3):470-8.

55.   Murali R, Wiesner T, Scolyer RA. Tumours associated with BAP1 mutations. Pathology. 2013;45(2):116-26.

56.   Chauhan D, Tian Z, Nicholson B, Kumar KG, Zhou B, Carrasco R, et al. A small molecule inhibitor of ubiquitin-specific protease-7 induces apoptosis in multiple myeloma cells and overcomes bortezomib resistance. Cancer cell. 2012;22(3):345-58.

57.   Hao YH, Fountain MD, Jr., Fon Tacer K, Xia F, Bi W, Kang SH, et al. USP7 Acts as a Molecular Rheostat to Promote WASH-Dependent Endosomal Protein Recycling and Is Mutated in a Human Neurodevelopmental Disorder. Mol Cell. 2015;59(6):956-69.

58.   Reincke M, Sbiera S, Hayakawa A, Theodoropoulou M, Osswald A, Beuschlein F, et al. Mutations in the deubiquitinase gene USP8 cause Cushing's disease. Nature genetics. 2015;47(1):31-8.

59.   Eichhorn PJ, Rodon L, Gonzalez-Junca A, Dirac A, Gili M, Martinez-Saez E, et al. USP15 stabilizes TGF-beta receptor I and promotes oncogenesis through the activation of TGF-beta signaling in glioblastoma. Nat Med. 2012;18(3):429-35.

60.   Huang X, Summers MK, Pham V, Lill JR, Liu J, Lee G, et al. Deubiquitinase USP37 is activated by CDK2 to antagonize APC(CDH1) and promote S phase entry. Mol Cell. 2011;42(4):511-23.

61.   Pan J, Deng Q, Jiang C, Wang X, Niu T, Li H, et al. USP37 directly deubiquitinates and stabilizes c-Myc in lung cancer. Oncogene. 2015;34(30):3957-67.

62.   Yao T, Cohen RE. A cryptic protease couples deubiquitination and degradation by the proteasome. Nature. 2002;419:403.

63.   Wu N, Liu C, Bai C, Han YP, Cho WC, Li Q. Over-expression of deubiquitinating enzyme USP14 in lung adenocarcinoma promotes proliferation through the accumulation of beta-catenin. International journal of molecular sciences. 2013;14(6):10749-60.

64.   Fang Y, Fu D, Tang W, Cai Y, Ma D, Wang H, et al. Ubiquitin C-terminal Hydrolase 37, a novel predictor for hepatocellular carcinoma recurrence, promotes cell migration and invasion via interacting and deubiquitinating PRP19. Biochimica et biophysica acta. 2013;1833(3):559-72.

65.   Harrigan JA, Jacq X, Martin NM, Jackson SP. Deubiquitylating enzymes and drug discovery: emerging opportunities. Nature Reviews Drug Discovery. 2017;17:57.

66.   Sun L, Xing Y, Chen X, Zheng Y, Yang Y, Nichols DB, et al. Coronavirus Papain-like Proteases Negatively Regulate Antiviral Innate Immune Response through Disruption of STING-Mediated Signaling. PLOS ONE. 2012;7(2):e30802.

67.   Pruneda Jonathan N, Durkin Charlotte H, Geurink Paul P, Ovaa H, Santhanam B, Holden David W, et al. The Molecular Basis for Ubiquitin and Ubiquitin-like Specificities in Bacterial Effector Proteases. Molecular Cell. 2016;63(2):261-76.

68.   Artavanis-Tsakonas K, Misaghi S, Comeaux CA, Catic A, Spooner E, Duraisingh MT, et al. Identification by functional proteomics of a deubiquitinating/deNeddylating enzyme in Plasmodium falciparum. Mol Microbiol. 2006;61(5):1187-95.

69.   Artavanis-Tsakonas K, Weihofen WA, Antos JM, Coleman BI, Comeaux CA, Duraisingh MT, et al. Characterization and structural studies of the Plasmodium falciparum ubiquitin and Nedd8 hydrolase UCHL3. The Journal of biological chemistry. 2010;285(9):6857-66.

70.   Frickel EM, Quesada V, Muething L, Gubbels MJ, Spooner E, Ploegh H, et al. Apicomplexan UCHL3 retains dual specificity for ubiquitin and Nedd8 throughout evolution. Cell Microbiol. 2007;9(6):1601-10.

71.   C. Noe M, Gilbert A. Targeted Covalent Enzyme Inhibitors. 472012. p. 413-39.

72.   Guterman L. COVALENT DRUGS FORM LONG-LIVED TIES. Chemical & Engineering News Archive. 2011;89(36):19-26.

73.   Roth GJ, Stanford N, Majerus PW. Acetylation of prostaglandin synthase by aspirin. Proceedings of the National Academy of Sciences of the United States of America. 1975;72(8):3073-6.

74.   Wright PM, Seiple IB, Myers AG. The evolving role of chemical synthesis in antibacterial drug discovery. Angewandte Chemie (International ed in English). 2014;53(34):8840-69.

75.   Bauer RA. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. Drug discovery today. 2015;20(9):1061-73.

76.   Bradshaw JM, McFarland JM, Paavilainen VO, Bisconte A, Tam D, Phan VT, et al. Prolonged and tunable residence time using reversible covalent kinase inhibitors. Nat Chem Biol. 2015;11(7):525-31.

77.   Copeland RA, Pompliano DL, Meek TD. Drug-target residence time and its implications for lead optimization. Nat Rev Drug Discov. 2006;5(9):730-9.

78.   Swinney DC. Biochemical mechanisms of drug action: what does it take for success? Nature Reviews Drug Discovery. 2004;3:801.

79.   Kuntz ID, Chen K, Sharp KA, Kollman PA. The maximal affinity of ligands. Proceedings of the National Academy of Sciences. 1999;96(18):9997-10002.

80.   Smith AJT, Zhang X, Leach AG, Houk KN. Beyond Picomolar Affinities: Quantitative Aspects of Noncovalent and Covalent Binding of Drugs to Proteins. Journal of Medicinal Chemistry. 2009;52(2):225-33.

81.   Polinsky RJ. Clinical pharmacology of rivastigmine: a new-generation acetylcholinesterase inhibitor for the treatment of alzheimer's disease. Clinical Therapeutics. 1998;20(4):634-47.

82.   Halgren TA. Identifying and Characterizing Binding Sites and Assessing Druggability. Journal of Chemical Information and Modeling. 2009;49(2):377-89.

83.   Mah R, Thomas JR, Shafer CM. Drug discovery considerations in the development of covalent inhibitors. Bioorganic & medicinal chemistry letters. 2014;24(1):33-9.

84.   Gunnell D, Murray V, Hawton K. Use of paracetamol (acetaminophen) for suicide and nonfatal poisoning: worldwide patterns of use and misuse. Suicide & life-threatening behavior. 2000;30(4):313-26.

85.   Wang F, Wang L, Wu J, Sokirniy I, Nguyen P, Bregnard T, et al. Active site-targeted covalent irreversible inhibitors of USP7 impair the functions of Foxp3+ T-regulatory cells by promoting ubiquitination of Tip60. PLOS ONE. 2017;12(12):e0189744.

86.   Peterson LF, Sun H, Liu Y, Potu H, Kandarpa M, Ermann M, et al. Targeting deubiquitinase activity with a novel small-molecule inhibitor as therapy for B-cell malignancies. Blood. 2015;125(23):3588-97.

87.   Erez E, Fass D, Bibi E. How intramembrane proteases bury hydrolytic reactions in the membrane. Nature. 2009;459(7245):371-8.

88.   Salmond GP, Reeves PJ. Membrane traffic wardens and protein secretion in gram-negative bacteria. Trends in biochemical sciences. 1993;18(1):7-12.

89.   Mueller KE, Plano GV, Fields KA. New frontiers in type III secretion biology: the Chlamydia perspective. Infection and immunity. 2014;82(1):2-9.

90.   Fischer A, Harrison KS, Ramirez Y, Auer D, Chowdhury SR, Prusty BK, et al. Chlamydia trachomatis-containing vacuole serves as deubiquitination platform to stabilize Mcl-1 and to interfere with host defense. eLife. 2017;6:e21465.

91.   Wang X, Hybiske K, Stephens RS. Direct visualization of the expression and localization of chlamydial effector proteins within infected host cells. Pathogens and Disease. 2018;76(2):fty011-fty.

92.   Misaghi S, Balsara ZR, Catic A, Spooner E, Ploegh HL, Starnbach MN. Chlamydia trachomatis-derived deubiquitinating enzymes in mammalian cells during infection. Molecular Microbiology. 2006;61(1):142-50.

93.   Pruneda JN, Bastidas RJ, Bertsoulaki E, Swatek KN, Santhanam B, Clague MJ, et al. A Chlamydia effector combining deubiquitination and acetylation activities induces Golgi fragmentation. Nature Microbiology. 2018;3(12):1377-84.

94.   Le Negrate G, Krieg A, Faustin B, Loeffler M, Godzik A, Krajewski S, et al. ChlaDub1 of Chlamydia trachomatis suppresses NF-κB activation and inhibits IκBα ubiquitination and degradation. Cellular Microbiology. 2008;10(9):1879-92.

95.   Wang X, Hybiske K, Stephens RS. Orchestration of the mammalian host cell glucose transporter proteins-1 and 3 by Chlamydia contributes to intracellular growth and infectivity. Pathogens and Disease. 2017;75(8):ftx108-ftx.

96.   Rawlings ND, Waller M, Barrett AJ, Bateman A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Research. 2014;42(D1):D503-D9.

97.   Barrett AJ, Rawlings ND. Evolutionary Lines of Cysteine Peptidases. Biological Chemistry2001. p. 727.

98.   Clague MJ, Barsukov I, Coulson JM, Liu H, Rigden DJ, Urbé S. Deubiquitylases From Genes to Organism. Physiological Reviews. 2013;93(3):1289-315.

99.   Ronau JA, Beckmann JF, Hochstrasser M. Substrate specificity of the ubiquitin and Ubl proteases. Cell research. 2016;26(4):441-56.

100.   Kim J-G, Taylor KW, Hotson A, Keegan M, Schmelz EA, Mudgett MB. XopD SUMO Protease Affects Host Transcription, Promotes Pathogen Growth, and Delays Symptom Development in Xanthomonas-Infected Tomato Leaves. The Plant Cell. 2008;20(7):1915-29.

101.   Rytkönen A, Poh J, Garmendia J, Boyle C, Thompson A, Liu M, et al. SseL, a Salmonella deubiquitinase required for macrophage killing and virulence. Proceedings of the National Academy of Sciences. 2007;104(9):3502-7.

102.   Mukherjee S, Keitany G, Li Y, Wang Y, Ball HL, Goldsmith EJ, et al. Yersinia YopJ acetylates and inhibits kinase activation by blocking phosphorylation. Science (New York, NY). 2006;312(5777):1211-4.

103.   Mossessova E, Lima CD. Ulp1-SUMO Crystal Structure and Genetic Analysis Reveal Conserved Interactions and a Regulatory Element Essential for Cell Growth in Yeast. Molecular Cell. 2000;5(5):865-76.

104.   Li S-J, Hochstrasser M. A new protease required for cell-cycle progression in yeast. Nature. 1999;398(6724):246-51.

105.   Yeh ET, Gong L, Kamitani T. Ubiquitin-like proteins: new wines in new bottles. Gene. 2000;248(1-2):1-14.

106.   Cotten M, Weber JM. The adenovirus protease is required for virus entry into host cells. Virology. 1995;213(2):494-502.

107.   Balakirev MY, Jaquinod M, Haas AL, Chroboczek J. Deubiquitinating Function of Adenovirus Proteinase. Journal of Virology. 2002;76(12):6323-31.

108. Grosche P, Sirockin F, Mac Sweeney A, Ramage P, Erbel P, Melkko S, et al. Structure-based design and optimization of potent inhibitors of the adenoviral protease. Bioorganic & medicinal chemistry letters. 2015;25(3):438-43.

109. Mac Sweeney A, Grosche P, Ellis D, Combrink K, Erbel P, Hughes N, et al. Discovery and Structure-Based Optimization of Adenain Inhibitors. ACS Medicinal Chemistry Letters. 2014;5(8):937-41.

110. Davies J. Where have All the Antibiotics Gone? The Canadian journal of infectious diseases & medical microbiology = Journal canadien des maladies infectieuses et de la microbiologie medicale. 2006;17(5):287-90.

111. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta crystallographica Section D, Biological crystallography. 2013;69(Pt 7):1204-14.

112. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403-10.

113. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta crystallographica Section D, Biological crystallography. 2011;67(Pt 4):235-42.

114. Potterton E, Briggs P, Turkenburg M, Dodson E. A graphical user interface to the CCP4 program suite. Acta Crystallographica Section D. 2003;59(7):1131-7.

115. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallographica Section D. 2010;66(4):486-501.

116. Holm L, Rosenström P. Dali server: conservation mapping in 3D. Nucleic Acids Research. 2010;38(suppl 2):W545-W9.

117. Touw WG, Baakman C, Black J, te Beek TAH, Krieger E, Joosten RP, et al. A series of PDB-related databanks for everyday needs. Nucleic acids research. 2015;43(Database issue):D364-D8.

118. Fortier A, Grosdidier A, Hernandez C, Baratin D, Kuznetsov D, de Castro E, et al. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Research. 2012;40(W1):W597-W603.

119. McCoy A. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallographica Section D. 2007;63(1):32-41.

120. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D. 2010;66(2):213-21.

121. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protocols. 2015;10(6):845-58.

122. Evans P. Scaling and assessment of data quality. Acta Crystallographica Section D. 2006;62(1):72-82.

123.   Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica Section D, Biological crystallography. 1997;53(Pt 3):240-55.

124.   Kabsch W. XDS. Acta Crystallographica Section D. 2010;66(2):125-32.

125.   Neudert G, Klebe G. DSX: A Knowledge-Based Scoring Function for the Assessment of Protein–Ligand Complexes. Journal of Chemical Information and Modeling. 2011;51(10):2731-45.

126.   Cole J, Nissink JWM, Taylor R. Protein-Ligand Docking Virtual Screening with GOLD: CRC Press; 2005. 379-415 p.

127.   Group CC. Molecular Operating Environment (MOE). 2018.

128.   Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28(1):235-42.

129.   UniProt C. The universal protein resource (UniProt). Nucleic acids research. 2008;36(Database issue):D190-D5.

130.   Sterling T, Irwin JJ. ZINC 15 – Ligand Discovery for Everyone. Journal of Chemical Information and Modeling. 2015;55(11):2324-37.

131.   Li MZ, Elledge SJ. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Meth. 2007;4(3):251-6.

132.   Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970;227(5259):680-5.

133.   Swinehart DF. The Beer-Lambert Law. Journal of Chemical Education. 1962;39(7):333.

134.   Block H, Maertens B, Spriestersbach A, Brinker N, Kubicek J, Fabis R, et al. Immobilized-metal affinity chromatography (IMAC): a review. Methods in enzymology. 2009;463:439-73.

135.   Structural Genomics C, Gräslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, et al. Protein production and purification. Nature Methods. 2008;5:135.

136.   Copeland RA. Kinetics of Single-Substrate Enzyme Reactions. Enzymes2002.

137.   Hassiepen U, Eidhoff U, Meder G, Bulber JF, Hein A, Bodendorf U, et al. A sensitive fluorescence intensity assay for deubiquitinating proteases using ubiquitin-rhodamine110-glycine as substrate. Analytical biochemistry. 2007;371(2):201-7.

138.   Copeland RA. Assay Considerations for Compound Library Screening. Evaluation of Enzyme Inhibitors in Drug Discovery2013.

139.   Grosche P, Sirockin F, Mac Sweeney A, Ramage P, Erbel P, Melkko S, et al. Structure-based design and optimization of potent inhibitors of the adenoviral protease. Bioorganic & medicinal chemistry letters. 2015;25(3):438-43.

140.   Kathman SG, Statsyuk AV. Covalent tethering of fragments for covalent probe discovery. MedChemComm. 2016;7(4):576-85.

141. Kathman SG, Xu Z, Statsyuk AV. A Fragment-Based Method to Discover Irreversible Covalent Inhibitors of Cysteine Proteases. Journal of Medicinal Chemistry. 2014;57(11):4969-74.

142. Copeland RA. Lead Optimization and Structure–Activity Relationships for Reversible Inhibitors. Evaluation of Enzyme Inhibitors in Drug Discovery: John Wiley & Sons, Inc.; 2013. p. 169-201.

143. Krippendorff BF, Neuhaus R, Lienau P, Reichel A, Huisinga W. Mechanism-based inhibition: deriving K(I) and k(inact) directly from time-dependent IC(50) values. Journal of biomolecular screening. 2009;14(8):913-23.

144. Copeland RA. Evaluation of enzyme inhibitors in drug discovery. A guide for medicinal chemists and pharmacologists. Methods of biochemical analysis. 2005;46:1-265.

145. Copeland RA. Testing for reversibility. Evaluation of Enzyme Inhibitors in Drug Discovery: John Wiley & Sons, Inc.; 2005. p. 125.

146. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. Proteins. 2003;52(4):609-23.

147. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. Empirical scoring functions - I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J Comput Aided Mol Des. 1997;11(5):425-45.

148. Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. Journal of Molecular Biology. 1995;245(1):43-53.

149. Groom CR, Bruno IJ, Lightfoot MP, Ward SC. The Cambridge Structural Database. Acta crystallographica Section B, Structural science, crystal engineering and materials. 2016;72(Pt 2):171-9.

150. Bashford D, Gerwert K. Electrostatic calculations of the pKa values of ionizable groups in bacteriorhodopsin. Journal of Molecular Biology. 1992;224(2):473-86.

151. Labute P. The generalized Born/volume integral implicit solvent model: Estimation of the free energy of hydration using London dispersion instead of atomic surface area. J Comput Chem. 2008;29(10):1693-8.

152. Looger LL, Hellinga HW. Generalized dead-end elimination algorithms make large-scale protein side-chain structure prediction tractable: implications for protein design and structural genomics11Edited by F. E. Cohen. Journal of Molecular Biology. 2001;307(1):429-45.

153. Desmet J, Spriet J, Lasters I. Fast and accurate side-chain topology and energy refinement (FASTER) as a new method for protein structure optimization. Proteins: Structure, Function, and Genetics. 2002;48(1):31-43.

154.   Gilson MK. Multiple-site titration and molecular modeling: Two rapid methods for computing energies and forces for ionizable groups in proteins. Proteins: Structure, Function, and Bioinformatics. 1993;15(3):266-82.

155.   Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation 1 1Edited by J. Thornton. Journal of Molecular Biology. 1999;285(4):1735-47.

156.   Labute P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. Proteins: Structure, Function, and Bioinformatics. 2008;75(1):187-205.

157.   Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem. 1996;17(5-6):490-519.

158.   Halgren TA. MMFF VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. J Comput Chem. 1999;20(7):730-48.

159.   Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking 1 1Edited by F. E. Cohen. Journal of Molecular Biology. 1997;267(3):727-48.

160.   Bergfors T. Seeds to crystals. Journal of Structural Biology. 2003;142(1):66-76.

161.   Collaborative. The CCP4 suite: programs for protein crystallography. Acta Crystallographica Section D. 1994;50(5):760-3.

162.   Skubák P, Pannu NS. Automatic protein structure solution from weak X-ray data. Nat Commun. 2013;4.

163.   . PyMOL, The PyMOL Molecular Graphics System, Version 1806 Schrödinger, LLC.

164.   Kabsch W. A solution for the best rotation to relate two sets of vectors. Acta Crystallographica Section A. 1976;32(5):922-3.

165.   Reverter D, Wu K, Erdene TG, Pan Z-Q, Wilkinson KD, Lima CD. Structure of a Complex between Nedd8 and the Ulp/Senp Protease Family Member Den1. Journal of Molecular Biology. 2005;345(1):141-51.

166.   Ekkebus R, van Kasteren SI, Kulathu Y, Scholten A, Berlin I, Geurink PP, et al. On Terminal Alkynes That Can React with Active-Site Cysteine Nucleophiles in Proteases. Journal of the American Chemical Society. 2013;135(8):2867-70.

167.   Durham E, Dorr B, Woetzel N, Staritzbichler R, Meiler J. Solvent accessible surface area approximations for rapid and accurate protein structure prediction. Journal of Molecular Modeling. 2009;15(9):1093-108.

168.	Persch E, Dumele O, Diederich F. Molecular Recognition in Chemical and Biological Systems. Angewandte Chemie International Edition. 2015;54(11):3290-327.

169.	Neudert G, Klebe G. fconv: format conversion, manipulation and feature computation of molecular data. Bioinformatics. 2011;27(7):1021-2.

170.	Pan Y, Huang N, Cho S, MacKerell AD, Jr. Consideration of molecular weight during compound selection in virtual target-based database screening. Journal of chemical information and computer sciences. 2003;43(1):267-72.

171.	Jacobsson M, Karlén A. Ligand Bias of Scoring Functions in Structure-Based Virtual Screening. Journal of Chemical Information and Modeling. 2006;46(3):1334-43.

172.	Altmann E, Betschart C, Hayakawa K, Irie O, Sakaki J, Iwasaki G, et al., inventors; Novartis A.-G., Switz.; Novartis Pharma G.m.b.H. . assignee. Preparation of heteroaryl nitriles for treating disorders involving cathepsin K patent WO2004020441A1. 2004.

173.	Altmann E, Hayakawa K, Iwasaki G, inventors; Novartis A.-G., Switz.; Novartis Pharma G.m.b.H. . assignee. Preparation of 2-cyano-4-aminopyrimidines as cathepsin K inhibitors for the treatment of inflammations and other diseases patent WO2003020278A1. 2003.

6   APPENDIX

6.1   List of abbreviations

2'-Deoxyadenosine 5'-triphosphate	**dATP**
2'-Deoxycytidine 5'-triphosphate	dCTP
2'-Deoxyguanosine 5'-triphosphate	dGTP
2'-Deoxythymidine 5'-triphosphate	dTTP
Azithromycin	AZM
Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung	BESSY
Cambridge Structural Database	CSD
Chlamydia Muridarum	CM
Chlamydia pneumoniae	CP
Chlamydia suis	CS
Chlamydia trachomatis	CT
column volume	CV
deoxyribonucleic acid	DNA
Deubiquitylase	DUB
Deutsches Elektronen-Synchrotron	DESY
Dimethyl sulfoxide	DMSO
Dithiothreitol	DTT
elementary body	EB
Ethylenediaminetetraacetic acid	EDTA
European Molecular Biology Laboratory	EMBL
European Synchrotron Radiation Facility	ESRF
Genetic Optimization for Ligand Docking	GOLD
hours post infection	hpi
immobilized affinity chromatography	IMAC
Iodoacetamide	IAA
isopropyl β-D-1-thiogalactopyranoside	IPTG
Liquid chromatography–mass spectrometry	LCMS

Lysogeny broth	LB
Machado-Josephin domain-containing proteases	MJD
Michaelis-Menten constant	K_M
molecular operating environment	MOE
Molecular weight	MW
nonsteroidal anti-inflammatory drugs	NSAIDs
optical density of 1.0 at λ = 600 nm	OD_{600}

6.2 Screen composition

Cdu1_Ultimate

Well	Composition		
A1		0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
A2		0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
A3		0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
A4		0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
A5		0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
A6		0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
A7		0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
A8		0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
A9		0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
A10		0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
A11		0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
A12		0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
B1	2.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
B2	2.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
B3	2.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
B4	2.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
B5	2.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
B6	2.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
B7	2.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
B8	2.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
B9	2.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
B10	2.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
B11	2.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
B12	2.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
C1	5.0 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
C2	5.0 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
C3	5.0 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
C4	5.0 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
C5	5.0 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
C6	5.0 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
C7	5.0 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
C8	5.0 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
C9	5.0 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane

C10	5.0 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
C11	5.0 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
C12	5.0 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
D1	5.0 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
D2	5.0 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
D3	5.0 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
D4	5.0 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
D5	5.0 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
D6	5.0 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
D7	5.0 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
D8	5.0 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
D9	5.0 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
D10	5.0 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
D11	5.0 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
D12	5.0 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
E1	7.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
E2	7.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
E3	7.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
E4	7.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
E5	7.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
E6	7.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
E7	7.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
E8	7.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
E9	7.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
E10	7.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
E11	7.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
E12	7.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
F1	10 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
F2	10 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
F3	10 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
F4	10 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
F5	10 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
F6	10 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
F7	10 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
F8	10 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
F9	10 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
F10	10 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
F11	10 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
F12	10 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
G1	12.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
G2	12.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
G3	12.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
G4	12.5 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
G5	12.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
G6	12.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
G7	12.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
G8	12.5 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
G9	12.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
G10	12.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane

G11	12.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
G12	12.5 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
H1	15 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
H2	15 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
H3	15 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
H4	15 % v/v peg 20000	0.1 M bicine ph 8.8	2% v/v 1,4-dioxane
H5	15 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
H6	15 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
H7	15 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
H8	15 % v/v peg 20000	0.1 M bicine ph 9.0	2% v/v 1,4-dioxane
H9	15 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
H10	15 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
H11	15 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane
H12	15 % v/v peg 20000	0.1 M bicine ph 9.2	2% v/v 1,4-dioxane

6.3   Construct list

Designation	Organism	Amino acid range	Mutation	Vector	Tag	Cleavage
CDU1_L	*C. trachomatis*	Cdu1 (65-401)	none	pETM14	His6	3C
CDU1_S	*C. trachomatis*	Cdu1 (125-401)	none	pETM14	His6	3C
CDU1_L	*C. trachomatis*	Cdu1 (65-401)	none	pCDF	TRX-His6	3C
CDU1_S	*C. trachomatis*	Cdu1 (125-401)	none	pETM14	HIis6-SMT3	3C
CDU1_S1	*C. trachomatis*	Cdu1 (173-401)	none	pETM14	His6	3C
CDU1_S2	*C. trachomatis*	Cdu1 (202-401)	none	pETM14	His6	3C
CDU1_KENY	*C. trachomatis*	Cdu1 (229-401)	none	pETM14	His6	3C
CDU1_LCMS	*C. trachomatis*	Cdu1 (155-401)	none	pETM14	His6	3C
Cdu1_Mut	*C. trachomatis*	Cdu1 (155-401)	C174A, C226S, C345A, C368A	pETM14	His6	3C
Cdu1_CT	*C. trachomatis*	Cdu1 (155-401)	C174A, C226S, C368A	pETM14	His6	3C

6.4 Appendix figures
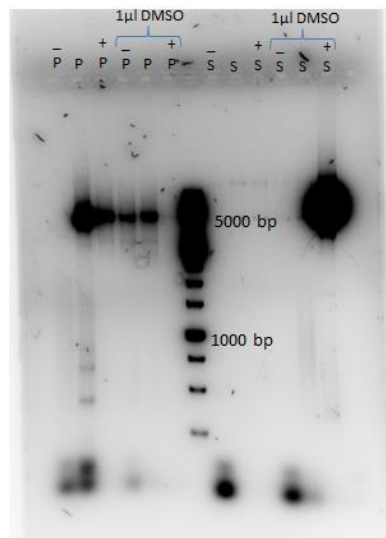
Figure 6.1. Linearization of the His6-pETM14 (P) and Smt3-pETM14 (S) plasmids. Amplification conditions were optimized by varying the annealing temperature by -5°C (-) and +5°C (+) in the presence and absence of DMSO.
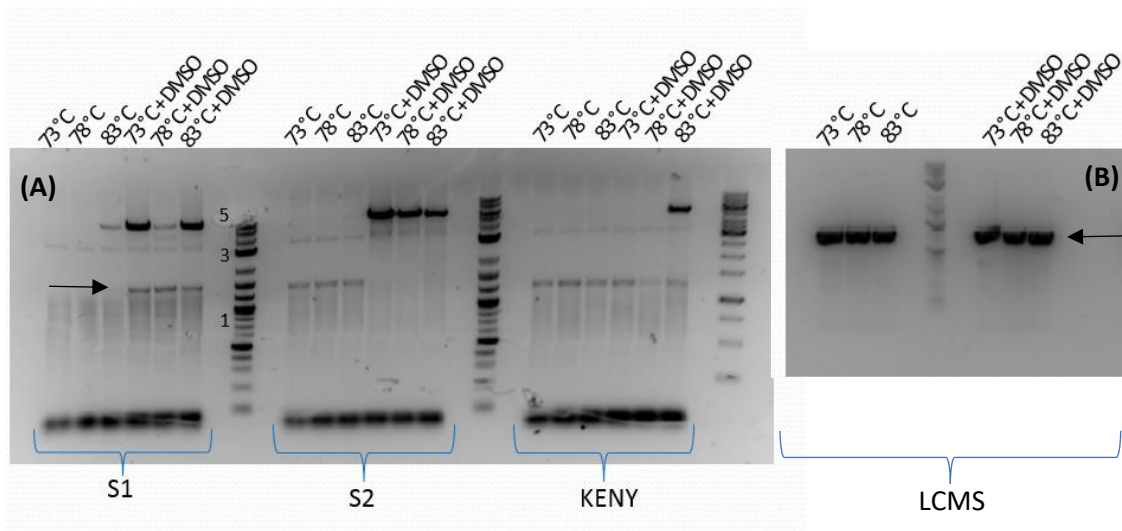
Figure 6.2 (A) Linearization of the shortened Cdu1_S1, Cdu1_S2 and Cdu1_KENY fragments (indicated by arrows). Amplification conditions were optimized by varying the annealing temperature by -5°C (-) and +5°C (+) in the presence and absence of DMSO. **(B)** Amplification of the Cdu1_LCMS gene.
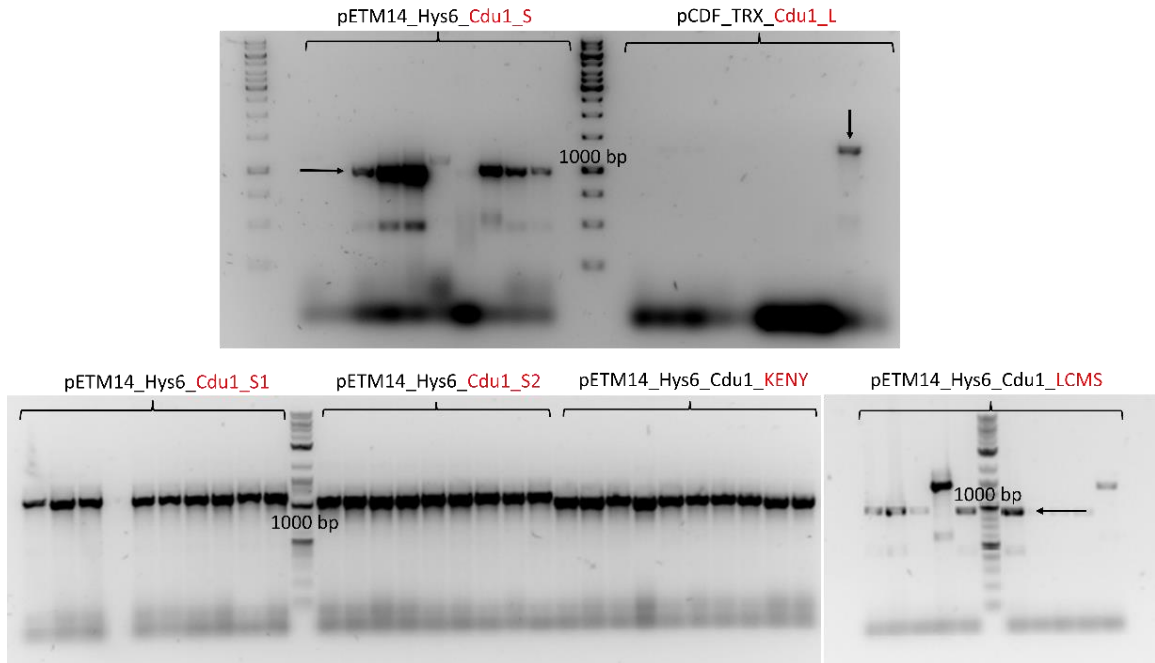
Appendix Figure 6.3. Colony PCR showing positive clones for the Cdu1 fragments in the pETM14 vector. The T7 promoter forward primer (plasmid) and BamHI reverse primer (insert) were used to verify the correct insertion of the Cdu1 gene. Positive clones are marked by a black arrow.
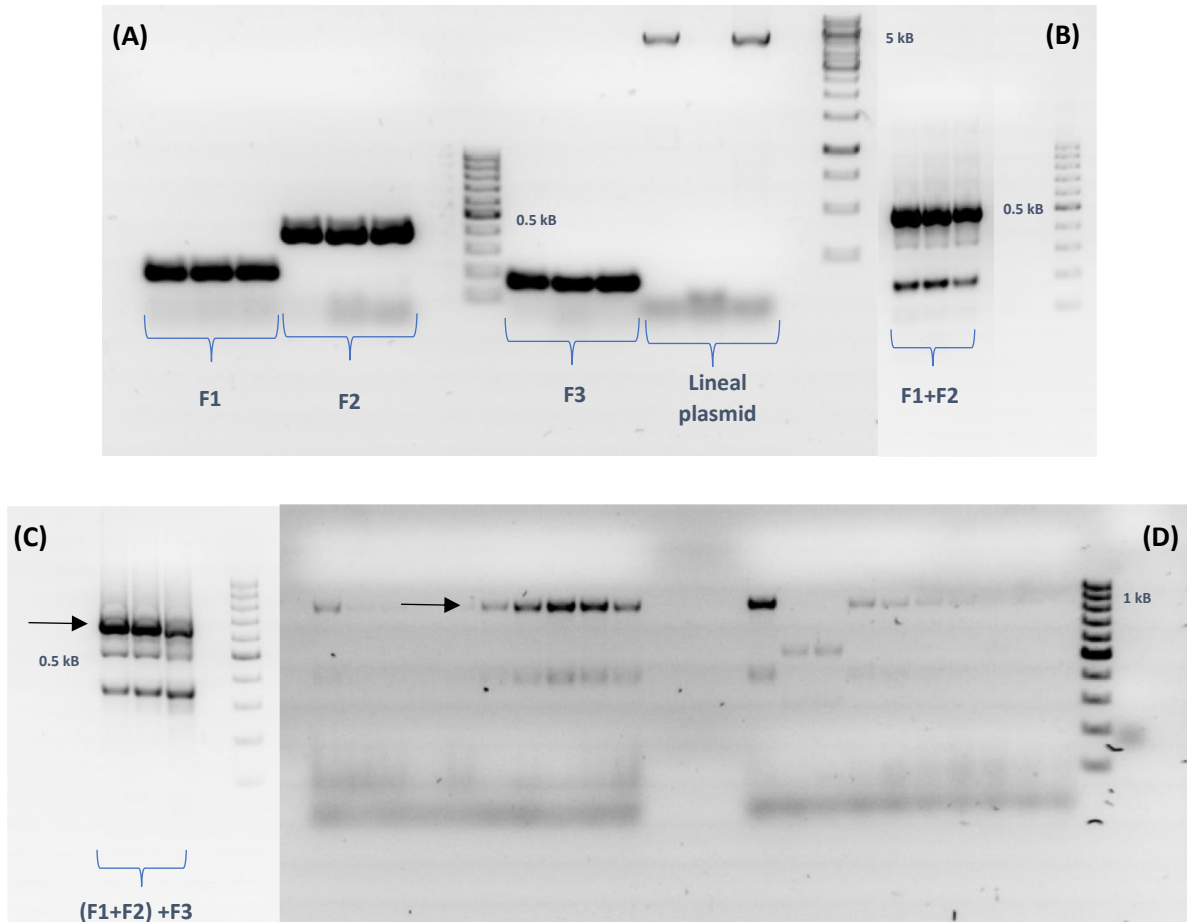
Appendix Figure 6.4. Cloning of the Cdu1_Mut construct

A. Amplification of mutated Cdu1 fragments.

B. Fusion of fragment 1 with fragment 2

C. Fusion of fragment 1 and fragment 2 with fragment 3

D. Colony PCR after SLIC

Docking and Scoring

Taken from Ramirez et al, 2018 with permission.

- **GOLD docking:** GOLD run using ChemScore and writing out the 50 best poses

  non-default settings used in the gold.conf file: flip_free_corners = 1, match_ring_templates = 1, flip_planar_n = 1 rot_ring_NRR flip_ring_NRR rot_ring_NHR flip_ring_NHR,

  flip_amide_bonds = 1, flip_pyramidal_n = 1, rotate_carboxylic_oh = flip, use_tordist = 1, postprocess_bonds = 1

  pocket definition: 20 A radius around sulfur atom of Cys 345.

- noteworthy **Scoring details:**

  DSX_CSD includes torsion, clash and SAS potentials: i.e., the dsx call used the flags: -T1 1 -T2 1 -T3 1

  The 50 pose geometries generated by GOLD in the initial ChemScore docking run were rescored by DSX_PDB and DSX_CSD.

- **Details of the consensus ranking scheme and how the representative docking pose was chosen:**

  Definitions:

  - **Rank_ =** rank of a pose obtained according to the value of a specific scoring function: here: **ChemScore**, **DSX_PDB** or **DSX_CSD**.

  - **Consensus Rank = (1/3) * (Rank_ChemScore + Rank_DSX_PDB + Rank_DSX_CSD)**

  - **EFFECTIVE Consensus Rank**: **-** translates the fractional 'Consensus Rank' into a more convenient integer rank again

    - considers not only the ascending order of the fractional 'Consensus Rank" values, but also whether the pose belongs to a

      certain cluster of poses. The same 'Effective Consensus Rank' value is given to all poses belonging to the same cluster: The

      lowest possible value is chosen.

- cluster A cluster of poses contains ligand poses with similar atomic positions. Cluster affiliation can be quantified by RMSD calculations

and a cutoff definition or often just by visual inspection of overlaid structures.

It is reasonable to work with a cluster concept in docking since the accuracy of scoring functions is limited and does not

allow for a final decision which pose in a cluster of poses is better than another one.

Details regarding cyano-pyrimidine 5:

☐ indicates the representative docking pose chosen for comparison with the crystal pose in the main text of the paper.

☐ best scored ChemScore pose

Table 6.1: Scores, ranks and root mean square deviations from the crystal structure (PDB 6FDQ) of cyano-pyrimidine **5**. The data of the 10 best poses according to the

'Consensus Rank' are listed.

PoseNr (GOLD)	RMSD to crystal pose (Å)	ChemScore	Rank ChemScore	DSX_PDB	Rank DSX_PDB	DSX_CSD	Rank DSX_CSD	Consensus Rank	EFFECTIVE Consensus Rank	Comments to EFFECTIVE Consensus Rank attribution
22	0.93	34.07	10	-78.19	2	-104.35	5	5.67	1	**1**, since best Consensus Rank value
35	1.16	34.54	1	-77.12	6	-102.78	11	6.00	2	**2**, since second best ConsensusRank pose and geometry different from #22
19	0.97	33.78	21	-78.84	1	-105.54	1	7.67	1	**1**, since atomic positions very similar to pose #22, i.e. #22 and #19 belong to the same cluster of pose geometries
17	1.17	34.52	2	-76.82	8	-102.16	15	8.33	…	
39	0.98	34.02	13	-76.75	9	-104.75	3	8.33	…	
23	0.95	33.94	16	-76.91	7	-102.87	10	11.00	…	
9	0.91	33.82	19	-77.27	4	-102.52	13	12.00	…	
34	0.93	33.86	17	-76.01	19	-105.23	2	12.67	…	
14	1.16	34.35	7	-76.60	11	-101.10	23	13.67	…	
5	1.18	34.44	4	-76.30	14	-100.58	25	14.33	…	

Details regarding cyano-pyrimidine 3:

☐ indicates the representative docking pose chosen for comparison with the crystal pose in the main text of the paper
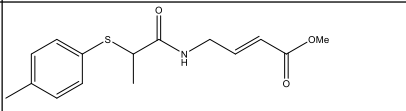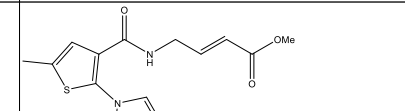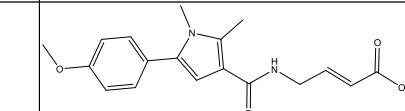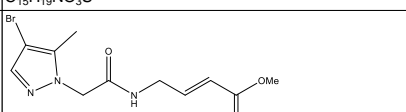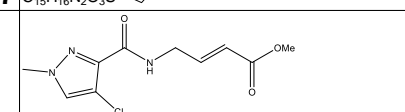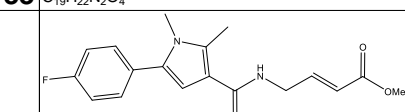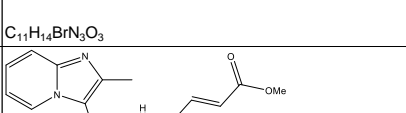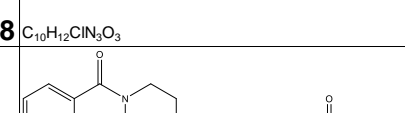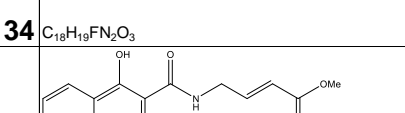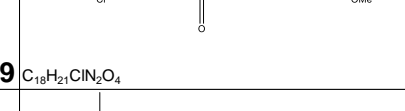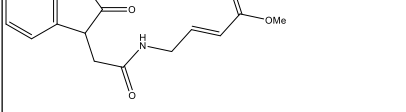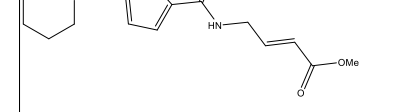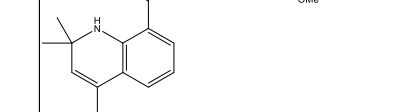
🟩 best scored ChemScore pose
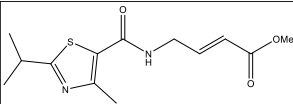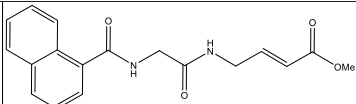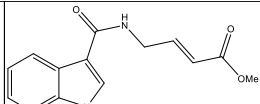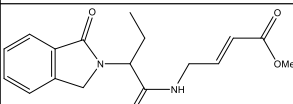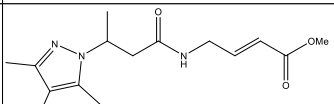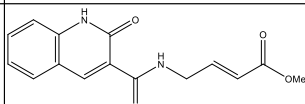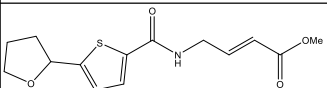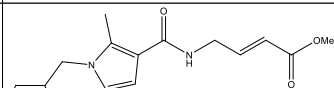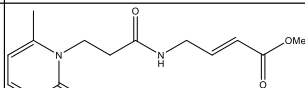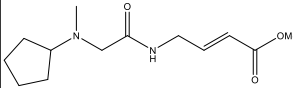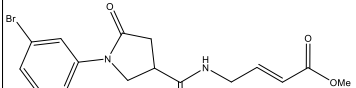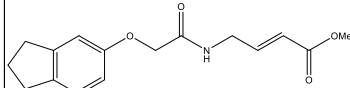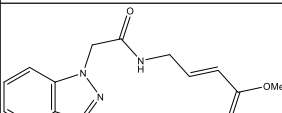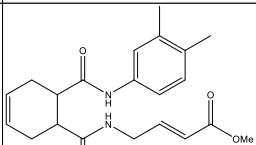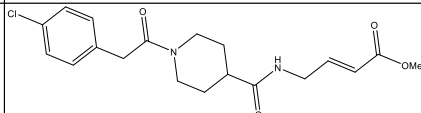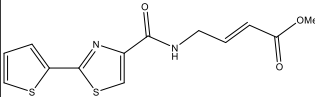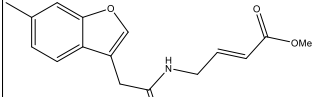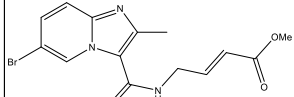
Table 6.2: Scores, ranks and root mean square deviations from the crystal structure (PDB 6FDU) of cyano-pyrimidine **3**. The data of the 10 best poses according to the 'Consensus Rank' are listed.
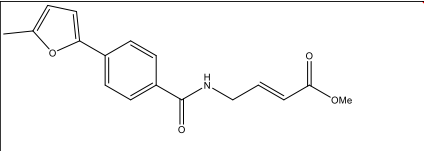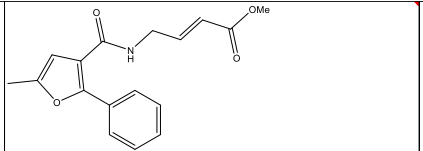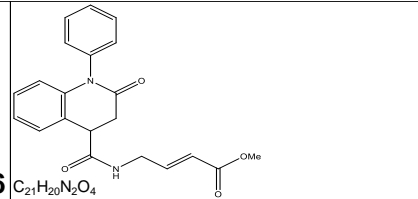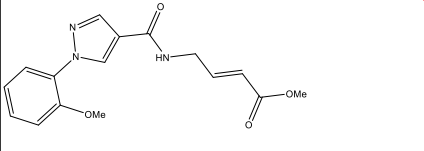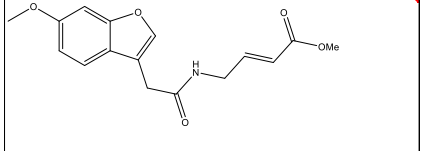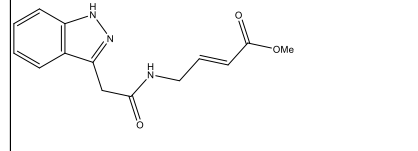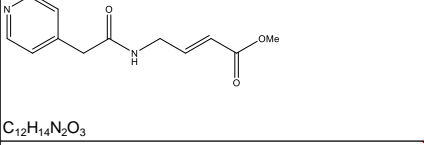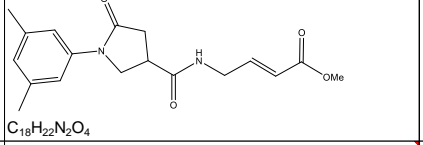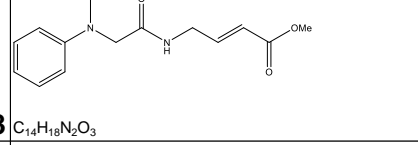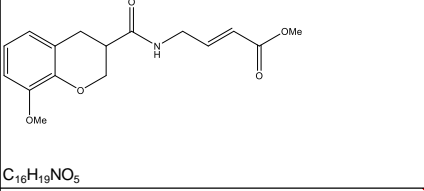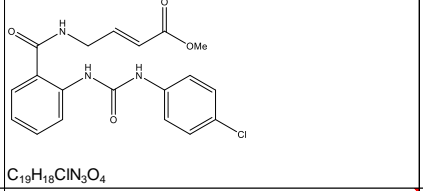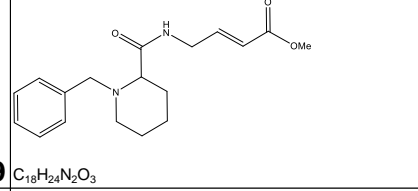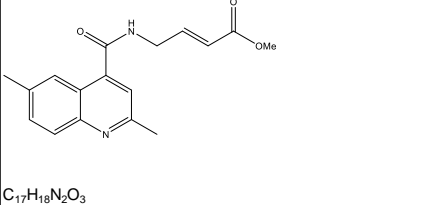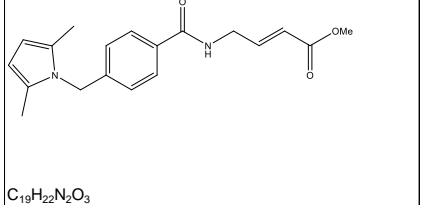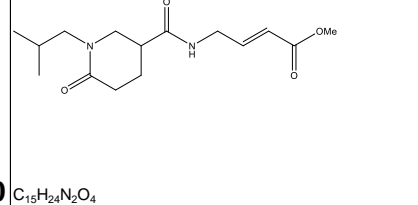
PoseNr (GOLD)	RMSD to crystal pose (Å)	ChemScore	Rank ChemScore	DSX_PDB	Rank DSX_PDB	DSX_CSD	Rank DSX_CSD	Consensus Rank	EFFECTIVE Consensus Rank	Comments to EFFECTIVE Consensus Rank attribution
32	3.31	36.50	3	-88.56	14	-108.72	1	6.00	1	**1**, since the pose has the best Consensus Rank value. However, the dichloro-benzene ring points towards the back wall of the pocket in disagreement with the X-ray experiment.
50	3.25	35.05	5	-96.30	5	-100.91	9	6.33	1	**1**, since pose #50 belongs to same cluster as pose #32. However, the di-chloro-benzene ring points again towards the back wall of the pocket in disagreement with the X-ray experiment.
2	1.04	32.97	10	-93.49	9	-102.03	5	8.00	2	This is the representative docking pose chosen for the comparison with the crystal pose in the main text discussions.
					Pose #**2** is effectively the second best ranked pose (not the third best!), since poses #32 and #50 belong to the same cluster with Effective Consensus Rank **1**. Pose #2 is geometrically distinct from above cluster **1** (RMSD to pose #32: 2.6 Å), since the dichloro-benzene ring points out of the pocket. This is <u>in agreement</u> with the experiment.					
10	1.47	32.31	12	-103.60	1	-96.95	14	9.00	…	
33	3.22	32.13	14	-89.47	11	-105.83	2	9.00	…	
36	3.58	37.01	1	-84.83	20	-101.42	6	9.00	…	Best ChemScore pose. The direction of the dichloro-benzene ring is similar to pose #32, i.e. points inwards into the pocket.
					Our analysis of pose #**36**'s scoring function contributions ChemScore.Lipo and ChemScore.DEClash and comparison with pose #**2** the experimental geometry shows that ChemScore overestimates the favorable lipophilic interactions of the inwards pointing dichloro-benzene ring with the alpha-helix D and the back wall of the pocket and strongly underestimates the van der Waals clashes. The direction of the dichloro-benzene ring in pose #**36** is <u>in disagreement</u> with the experiment. The RMSD to pose #**2** is 2.4 Å, i.e., #**36** belongs to a different cluster than pose #**2**. But the conformation of the main skeleton scaffold alone is not so different: RMSD without dichloro-benzene and NMe2 sidechain 1.4Å, RMSD without dichloro-benzene, NMe2 and iso-pentyl sidechain: 0.5 Å. Whereas ChemScore remains reliable for highly ranking a pose with a main scaffold conformation in agreement with experiment, the Consensus Ranking scheme is slightly more successful to get the dichloro-benzene direction right, too.					
46	3.34	30.11	22	-100.86	2	-104.35	3	9.00	…	
31	1.45	30.84	18	-95.41	7	-101.08	7	10.67	…	
25	3.49	36.72	2	-81.69	28	-102.05	4	11.33	…	
28	3.33	34.64	6	-90.26	10	-94.39	18	11.33	…	
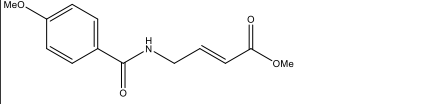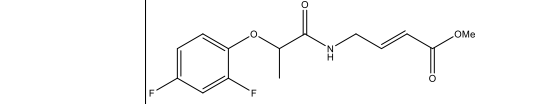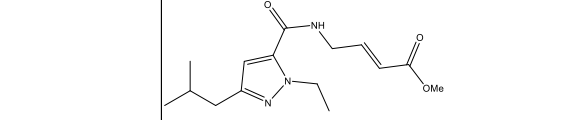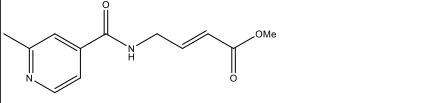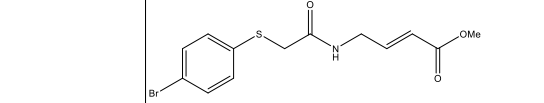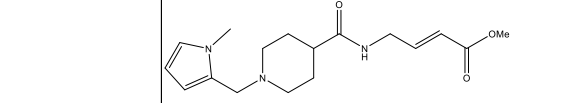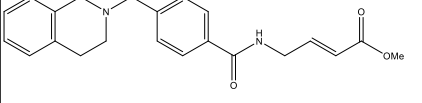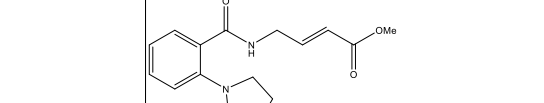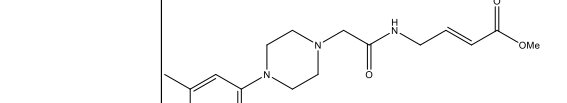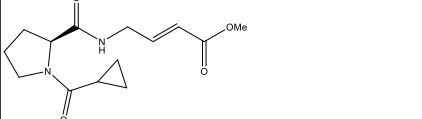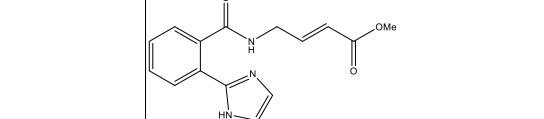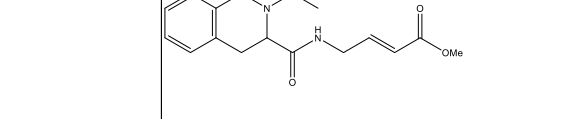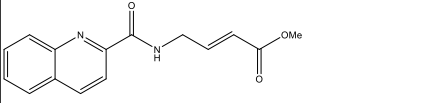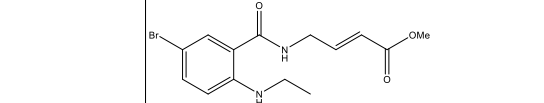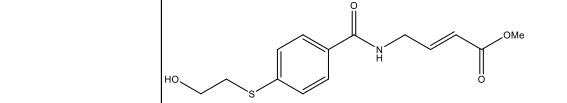
It is no surprise that one finds highest-ranked poses differing from the crystal pose if mobile sidechains (e.g. NMe2, iso-pentyl) or rings (e.g., dichloro-benzene) are present in a ligand. Scoring functions are often biased by design towards final docking poses with high shape-complementarity: they often tend to overweigh favorable lipophilic contacts with the pocket surface and underestimate van der Waals clash contributions instead. Since this was the case here for cyano-pyrimidine **3** and the best Chemscore pose #**36** for example, we decided not only to rely on ChemScore, but also to include the predictive power of the DrugscoreX scoring functions by rescoring the 50 best GOLD poses and using the ranks obtained with all three scoring functions to calculate a Consensus Rank. Our Consensus Rank approach gave us more confidence to decide on a final representative docking pose to be discussed in comparison with experiment.
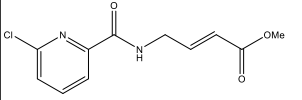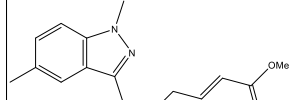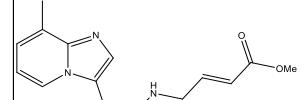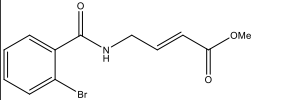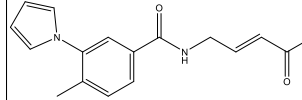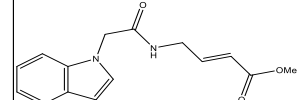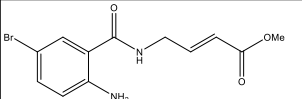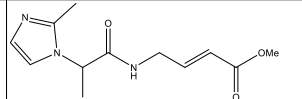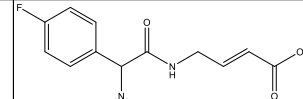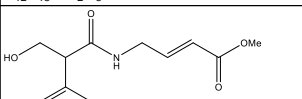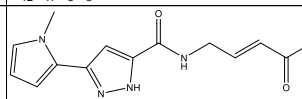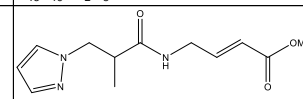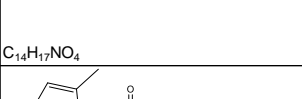
Table 6.3: Structures of the 180 VME-based compounds screened against Cdu1

Compound #	Structure	Compound #	Structure	Compound #	Structure
1 C₁₅H₁₉NO₃S		**67** C₁₅H₁₆N₂O₃S		**133** C₁₉H₂₂N₂O₄	
2 C₁₁H₁₄BrN₃O₃		**68** C₁₀H₁₂ClN₃O₃		**134** C₁₈H₁₉FN₂O₃	
3 C₁₅H₁₇N₃O₃		**69** C₁₈H₂₁ClN₂O₄		**135** C₁₆H₁₄ClNO₄	
4 C₁₆H₁₈N₂O₄		**70** C₁₇H₂₄N₂O₄		**136** C₁₈H₂₂N₂O₃	
5 C₁₅H₁₅FN₂O₄		**71** C₁₃H₂₁NO₃		**137** C₁₄H₁₇NO₄	

8 C$_{13}$H$_{18}$N$_2$O$_3$S	**74** C$_{18}$H$_{18}$N$_2$O$_4$	**140** C$_{13}$H$_{13}$N$_3$O$_3$
9 C$_{17}$H$_{20}$N$_2$O$_4$	**75** C$_{14}$H$_{20}$BrN$_3$O$_3$	**141** C$_{15}$H$_{14}$N$_2$O$_4$
10 C$_{14}$H$_{17}$NO$_4$S	**76** C$_{17}$H$_{20}$N$_2$O$_4$	**142** C$_{14}$H$_{18}$N$_2$O$_4$
11 C$_{13}$H$_{22}$N$_2$O$_3$	**77** C$_{16}$H$_{17}$BrN$_2$O$_4$	**143** C$_{16}$H$_{19}$NO$_4$
12 C$_{14}$H$_{15}$N$_3$O$_3$	**78** C$_{21}$H$_{26}$N$_2$O$_4$	**144** C$_{19}$H$_{23}$ClN$_2$O$_4$
13 C$_{13}$H$_{12}$N$_2$O$_3$S$_2$	**79** C$_{16}$H$_{17}$NO$_4$	**145** C$_{14}$H$_{14}$BrN$_3$O$_3$
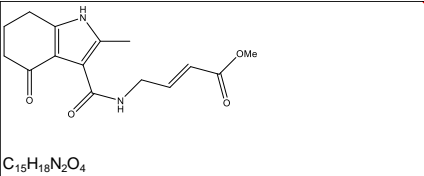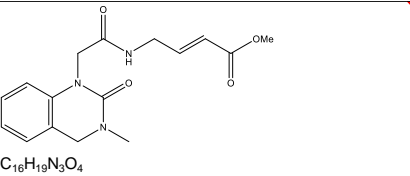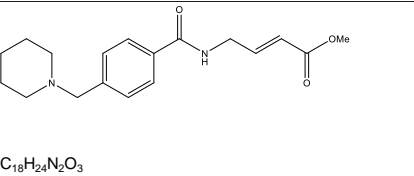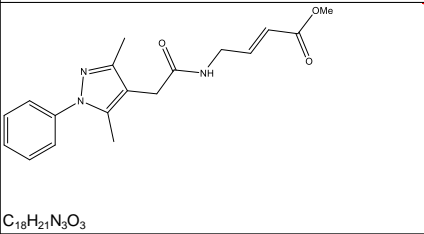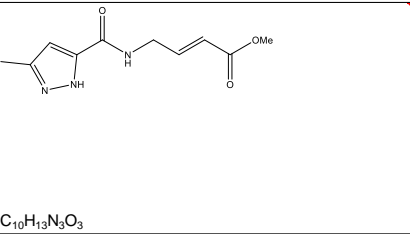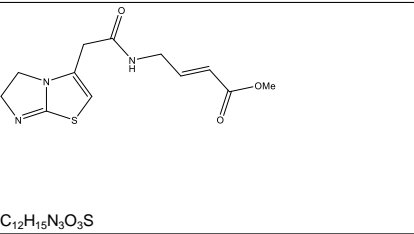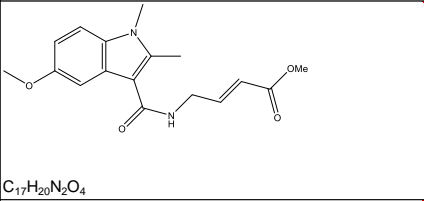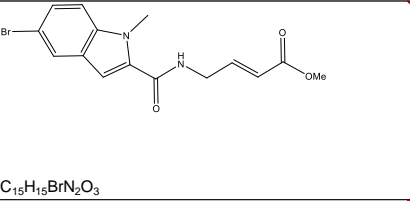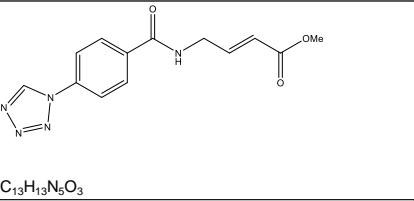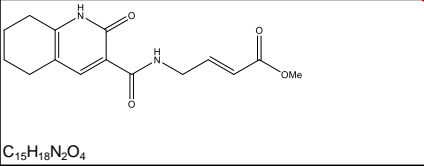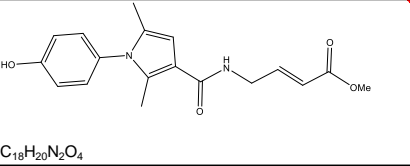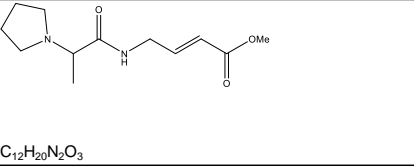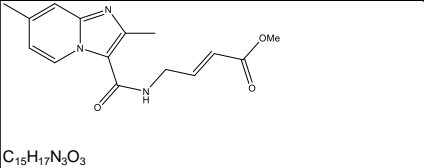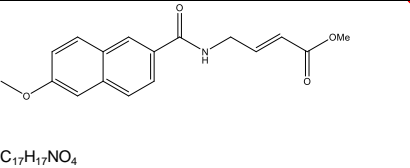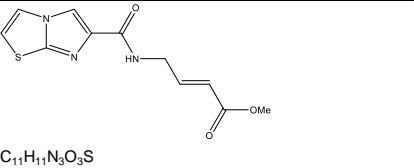
14 C$_{17}$H$_{17}$NO$_4$

80 C$_{17}$H$_{17}$NO$_4$

146 C$_{21}$H$_{20}$N$_2$O$_4$

15 C$_{16}$H$_{17}$N$_3$O$_4$

81 C$_{16}$H$_{17}$NO$_5$

147 C$_{14}$H$_{15}$N$_3$O$_3$

16 C$_{12}$H$_{14}$N$_2$O$_3$

82 C$_{18}$H$_{22}$N$_2$O$_4$

148 C$_{14}$H$_{18}$N$_2$O$_3$

17 C$_{16}$H$_{19}$NO$_5$

83 C$_{19}$H$_{18}$ClN$_3$O$_4$

149 C$_{18}$H$_{24}$N$_2$O$_3$

18 C$_{17}$H$_{18}$N$_2$O$_3$

84 C$_{19}$H$_{22}$N$_2$O$_3$

150 C$_{15}$H$_{24}$N$_2$O$_4$
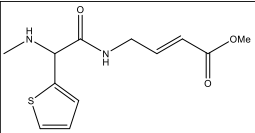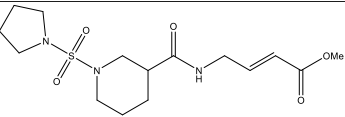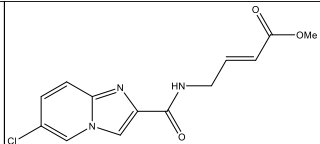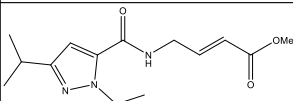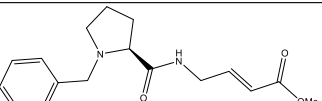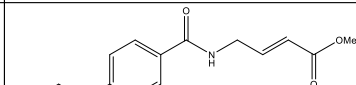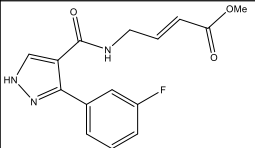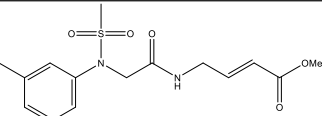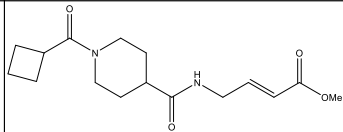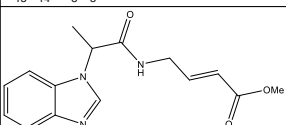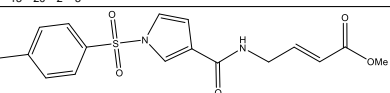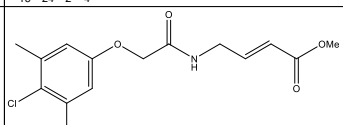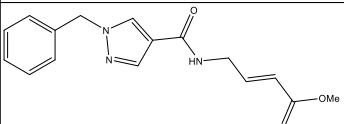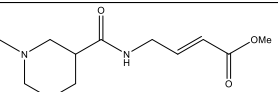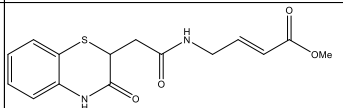
19 C$_{13}$H$_{15}$NO$_4$

85 C$_{14}$H$_{15}$F$_2$NO$_4$

151 C$_{15}$H$_{23}$N$_3$O$_3$

20 C$_{12}$H$_{14}$N$_2$O$_3$

86 C$_{13}$H$_{14}$BrNO$_3$S

152 C$_{17}$H$_{25}$N$_3$O$_3$

21 C$_{22}$H$_{24}$N$_2$O$_3$

87 C$_{16}$H$_{20}$N$_2$O$_3$

153 C$_{18}$H$_{25}$N$_3$O$_3$

22 C$_{14}$H$_{20}$N$_2$O$_4$

88 C$_{15}$H$_{15}$N$_3$O$_3$

154 C$_{17}$H$_{22}$N$_2$O$_3$

23 C$_{15}$H$_{14}$N$_2$O$_3$

89 C$_{14}$H$_{17}$BrN$_2$O$_3$

155 C$_{14}$H$_{17}$NO$_4$S
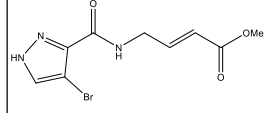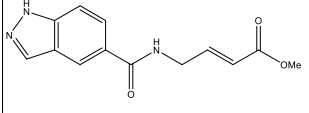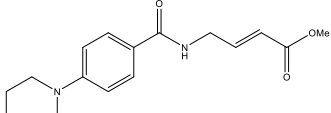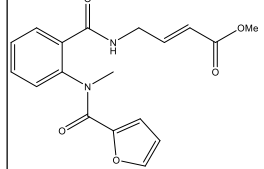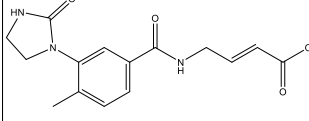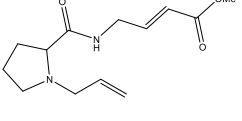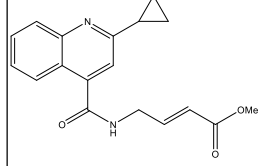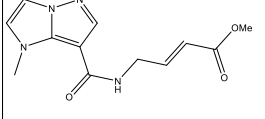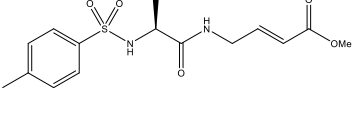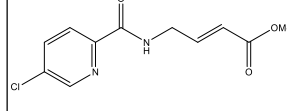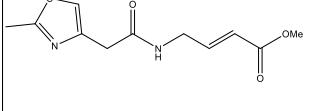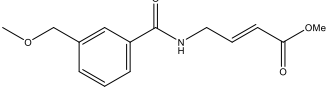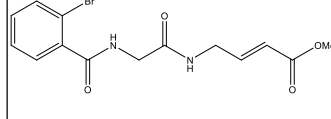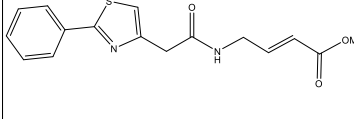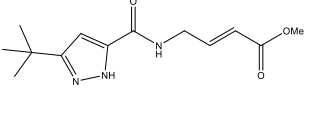
24 C$_{11}$H$_{11}$ClN$_2$O$_3$		**90** C$_{15}$H$_{17}$N$_3$O$_3$		**156** C$_{15}$H$_{17}$N$_3$O$_3$	
25 C$_{12}$H$_{12}$BrNO$_3$		**91** C$_{17}$H$_{18}$N$_2$O$_3$		**157** C$_{19}$H$_{20}$N$_2$O$_4$	
26 C$_{12}$H$_{13}$BrN$_2$O$_3$		**92** C$_{12}$H$_{17}$N$_3$O$_3$		**158** C$_{15}$H$_{19}$FN$_2$O$_3$	
27 C$_{14}$H$_{17}$NO$_4$		**93** C$_{14}$H$_{16}$N$_4$O$_3$		**159** C$_{12}$H$_{17}$N$_3$O$_3$	
28 C$_{12}$H$_{17}$N$_3$O$_3$		**94** C$_{15}$H$_{14}$ClNO$_5$		**160** C$_{18}$H$_{19}$NO$_4$	
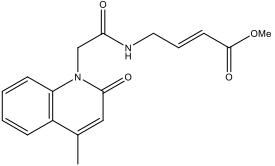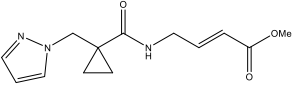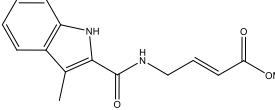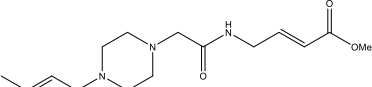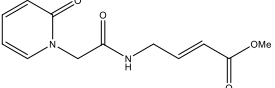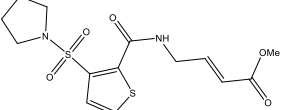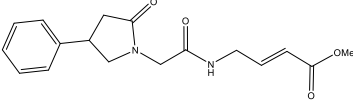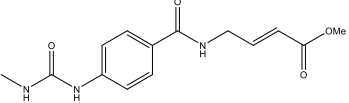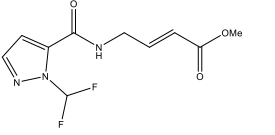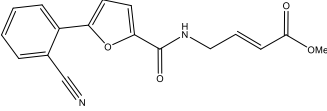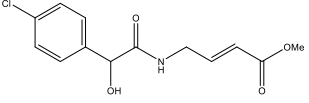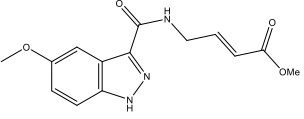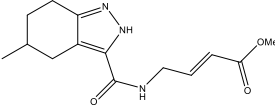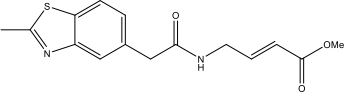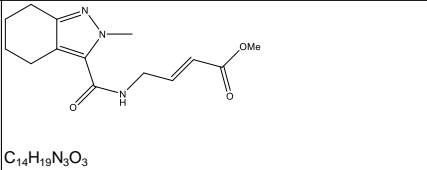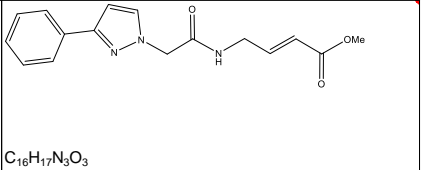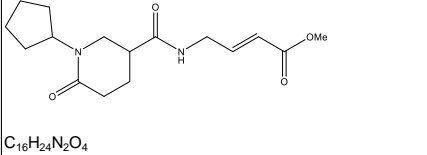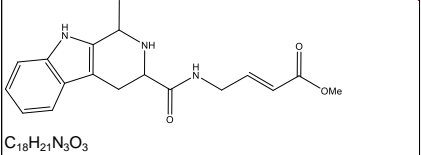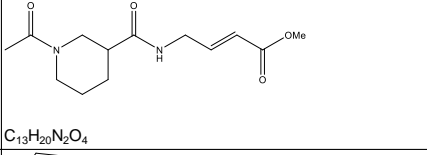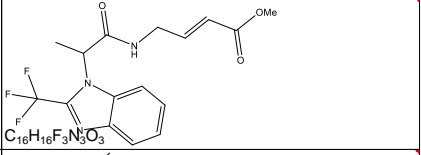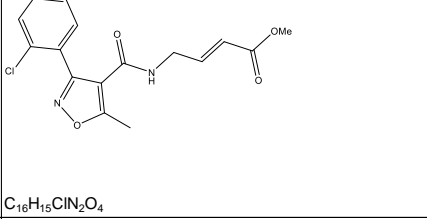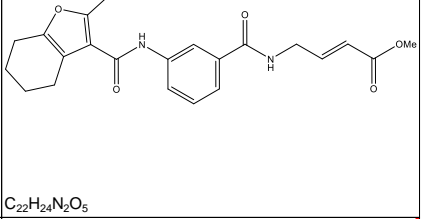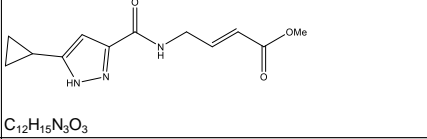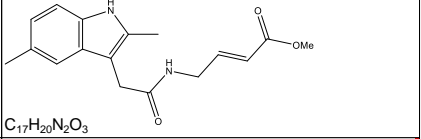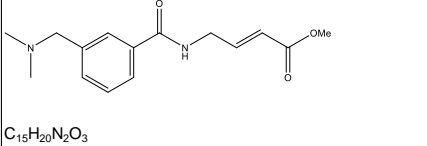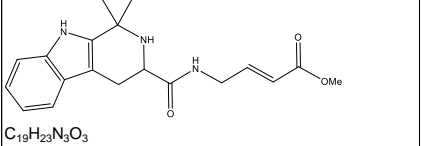
29 C$_{13}$H$_{14}$BrNO$_4$	**95** C$_{11}$H$_{12}$F$_3$N$_3$O$_3$	**161** C$_{15}$H$_{19}$FN$_2$O$_3$
30 C$_{16}$H$_{16}$N$_2$O$_4$	**96** C$_{14}$H$_{15}$NO$_3$S	**162** C$_{14}$H$_{22}$N$_2$O$_4$
31 C$_{12}$H$_{17}$N$_3$O$_3$	**97** C$_{15}$H$_{18}$N$_2$O$_4$	**163** C$_{16}$H$_{17}$N$_3$O$_3$
32 C$_{16}$H$_{26}$N$_2$O$_4$	**98** C$_{17}$H$_{24}$N$_2$O$_4$	**164** C$_{17}$H$_{19}$N$_3$O$_4$
33 C$_{13}$H$_{18}$N$_2$O$_4$	**99** C$_{12}$H$_{15}$NO$_4$S	**165** C$_{14}$H$_{18}$N$_2$O$_4$S
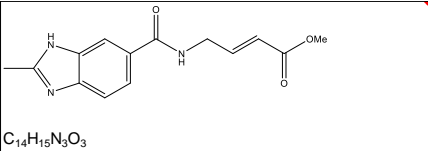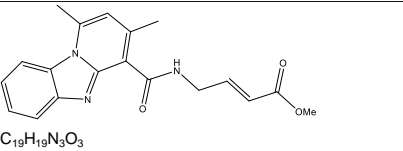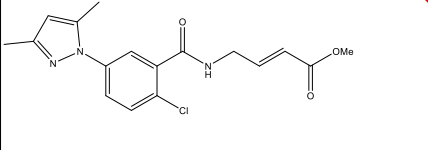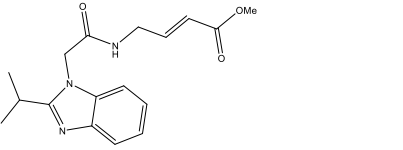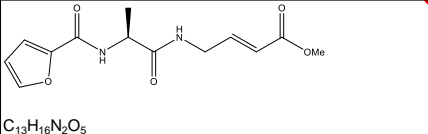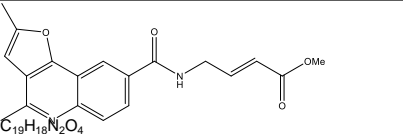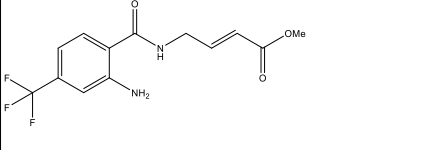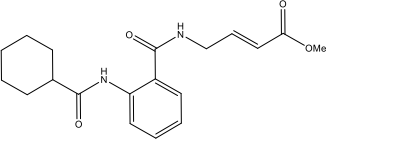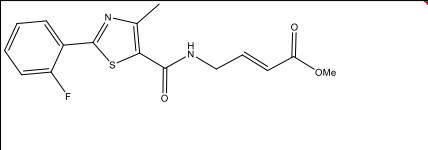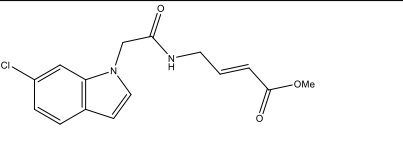
34 C$_{15}$H$_{18}$N$_2$O$_4$

100 C$_{16}$H$_{19}$N$_3$O$_4$

166 C$_{18}$H$_{24}$N$_2$O$_3$

35 C$_{18}$H$_{21}$N$_3$O$_3$

101 C$_{10}$H$_{13}$N$_3$O$_3$

167 C$_{12}$H$_{15}$N$_3$O$_3$S

36 C$_{17}$H$_{20}$N$_2$O$_4$

102 C$_{15}$H$_{15}$BrN$_2$O$_3$

168 C$_{13}$H$_{13}$N$_5$O$_3$

37 C$_{15}$H$_{18}$N$_2$O$_4$

103 C$_{18}$H$_{20}$N$_2$O$_4$

169 C$_{12}$H$_{20}$N$_2$O$_3$

38 C$_{15}$H$_{17}$N$_3$O$_3$

104 C$_{17}$H$_{17}$NO$_4$

170 C$_{11}$H$_{11}$N$_3$O$_3$S

39 C$_{12}$H$_{16}$N$_2$O$_3$S	**105** C$_{15}$H$_{25}$N$_3$O$_5$S	**171** C$_{13}$H$_{12}$ClN$_3$O$_3$
40 C$_{14}$H$_{21}$N$_3$O$_3$	**106** C$_{17}$H$_{22}$N$_2$O$_3$	**172** C$_{17}$H$_{23}$N$_3$O$_3$
41 C$_{15}$H$_{14}$FN$_3$O$_3$	**107** C$_{15}$H$_{20}$N$_2$O$_5$S	**173** C$_{16}$H$_{24}$N$_2$O$_4$
42 C$_{15}$H$_{17}$N$_3$O$_3$	**108** C$_{17}$H$_{18}$N$_2$O$_5$S	**174** C$_{15}$H$_{18}$ClNO$_4$
43 C$_{16}$H$_{17}$N$_3$O$_3$	**109** C$_{12}$H$_{20}$N$_2$O$_3$	**175** C$_{15}$H$_{16}$N$_2$O$_4$S

44 C$_9$H$_{10}$BrN$_3$O$_3$	**110** C$_{13}$H$_{13}$N$_3$O$_3$	**176** C$_{16}$H$_{20}$N$_2$O$_4$
45 C$_{18}$H$_{18}$N$_2$O$_5$	**111** C$_{16}$H$_{19}$N$_3$O$_4$	**177** C$_{13}$H$_{20}$N$_2$O$_3$
46 C$_{18}$H$_{18}$N$_2$O$_3$	**112** C$_{12}$H$_{14}$N$_4$O$_3$	**178** C$_{15}$H$_{20}$N$_2$O$_5$S
47 C$_{11}$H$_{11}$ClN$_2$O$_3$	**113** C$_{11}$H$_{14}$N$_2$O$_3$S	**179** C$_{14}$H$_{17}$NO$_4$
48 C$_{14}$H$_{15}$BrN$_2$O$_4$	**114** C$_{16}$H$_{16}$N$_2$O$_3$S	**180** C$_{13}$H$_{19}$N$_3$O$_3$

49 C₁₇H₁₈N₂O₄		**115** C₁₃H₁₇N₃O₃	
50 C₁₅H₁₆N₂O₃		**116** C₁₈H₂₅N₃O₃	
51 C₁₂H₁₄N₂O₄		**117** C₁₄H₁₈N₂O₅S₂	
52 C₁₇H₂₀N₂O₄		**118** C₁₄H₁₇N₃O₄	
53 C₁₀H₁₁F₂N₃O₃		**119** C₁₇H₁₄N₂O₄	
54 C₁₃H₁₄ClNO₄		**120** C₁₄H₁₅N₃O₄	
55 C₁₄H₁₉N₃O₃		**121** C₁₅H₁₆N₂O₃S	

	56 C$_{14}$H$_{19}$N$_3$O$_3$		**122** C$_{16}$H$_{17}$N$_3$O$_3$
	57 C$_{16}$H$_{24}$N$_2$O$_4$		**123** C$_{18}$H$_{21}$N$_3$O$_3$
	58 C$_{13}$H$_{20}$N$_2$O$_4$		**124** C$_{16}$H$_{16}$F$_3$N$_3$O$_3$
	59 C$_{16}$H$_{15}$ClN$_2$O$_4$		**125** C$_{22}$H$_{24}$N$_2$O$_5$
	60 C$_{12}$H$_{15}$N$_3$O$_3$		**126** C$_{17}$H$_{20}$N$_2$O$_3$
	61 C$_{15}$H$_{20}$N$_2$O$_3$		**127** C$_{19}$H$_{23}$N$_3$O$_3$

	62 C$_{14}$H$_{15}$N$_3$O$_3$		**128** C$_{19}$H$_{19}$N$_3$O$_3$
	63 C$_{17}$H$_{18}$ClN$_3$O$_3$		**129** C$_{17}$H$_{21}$N$_3$O$_3$
	64 C$_{13}$H$_{16}$N$_2$O$_5$		**130** C$_{19}$H$_{18}$N$_2$O$_4$
	65 C$_{13}$H$_{13}$F$_3$N$_2$O$_3$		**131** C$_{19}$H$_{24}$N$_2$O$_4$
	66 C$_{16}$H$_{15}$FN$_2$O$_3$S		**132** C$_{15}$H$_{15}$ClN$_2$O$_3$

LIST OF PUBLICATIONS

Y. Ramirez, T. Adler, E. Altmann, C. Tiesmeyer, T. Klemm, F. Sauer, S. Kathman, A. Statsyuk, C. Sotriffer and C. Kisker. *"Structural basis of substrate recognition and covalent inhibition of Cdu1 from Chlamydia trachomatis"* ChemMedChem 2018. **[https://doi.org/10.1002/cmdc.201800364]**

A. Fischer, K. S. Harrison, **Y. Ramirez**, D. Auer, S. R. Chowdhury, B. K. Prusty, F. Sauer, Z. Dimond, C. Kisker, P. Scott Hefty, T. Rudel. *"Chlamydia trachomatis-containing vacuole serves as deubiquitination platform to stabilize Mcl-1 and to interfere with host defense"* eLife 2017, *6*, e21465. **[https://doi.org/10.7554/eLife.21465.001]**

ACKNOWLEDGEMENTS

Affidavit

I hereby confirm that my thesis entitled

Structural basis of ubiquitin recognition and rational design of novel covalent inhibitors targeting Cdu1 from Chlamydia Trachomatis

is the result of my own work. I did not receive any help or support from commercial consultants.
All sources and / or materials applied are listed and specified in the thesis.
Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Wurzburg, ……………………………………								………………………………………….
(Date)															(Signature)

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation

Strukturelle Grundlage der Ubiquitin-Erkennung und rationales Design neuer kovalenter Inhibitoren gegen die Deubiquitinylase Cdu1 aus *Chlamydia trachomatis*

eigenständig, d. h. insbesondere selbstständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und
Hilfsmittel verwendet zu haben.
Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in
einem anderen Prüfungsverfahren vorgelegen hat.

Wurzburg, ……………………………………								………………………………………

(Date) (Signature)