



JACOBS
UNIVERSITY

Identification and optimisation of bacterial caseinolytic protease inhibitors as potential novel antibiotics

by

Elisa Sasseti

a Thesis submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy in Biochemistry

Approved Dissertation Committee

Prof. Dr. Mathias Winterhalter, Jacobs University Bremen

Prof. Dr. Ulrich Kleinekathöfer, Jacobs University Bremen

Dr. Björn Windshügel, Fraunhofer IME - ScreeningPort

Dr. Philip Gribbon, Fraunhofer IME - ScreeningPort

Date of Defense: 12th November 2018

Department of Life Sciences & Chemistry

Statutory Declaration

Family Name, Given/First Name	Sassetti, Elisa
Matriculationnumber	20331847
What kind of thesis are you submitting: Bachelor-, Master- or PhD-Thesis	PhD-Thesis

English: Declaration of Authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.

The Thesis has been written independently and has not been submitted at any other university for the conferral of a PhD degree; neither has the thesis been previously published in full.

German: Erklärung der Autorenschaft (Urheberschaft)

Ich erkläre hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprüft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukünftigen Arbeiten verglichen werden zu können. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfältigung.

Diese Arbeit wurde in der vorliegenden Form weder einer anderen Prüfungsbehörde vorgelegt noch wurde das Gesamtdokument bisher veröffentlicht.

.....
Date, Signature

**To my family
To my friends and
To all the amazing people that I met during this journey**

Table of contents

Abstract	XI
Acknowledgments.....	XIII
Funding.....	XIV
Aim of the work	XV
Chapter 1. Introduction	1
1.1 Antibiotics Era and Resistance	1
1.2 Overview of ClpP as a drug target	5
1.2.1 Bacterial proteases.....	5
1.2.2 Caseinolytic protease P	5
1.2.2.1 Prevalence among organisms	5
1.2.2.2 ClpP structure.....	6
1.2.2.3 Gate opening and product realize	7
1.2.3 Clp complex.....	8
1.2.4 ClpP in drug discovery.....	8
1.2.4.1 ClpP activators.....	9
1.2.4.2 ClpP inhibitors.....	9
1.3 Covalent drugs.....	14
1.3.1 Mechanism of action.....	15
1.4 References	16

Chapter 2. α-Amino diphenyl phosphonates as novel inhibitors of <i>Escherichia coli</i> ClpP protease	25
2.1 Abstract.....	26
2.2 Introduction.....	26
2.3 Results	30
2.3.1 Chemical explorations and enzymatic activity screening	30
2.3.2 Biological evaluation of hits.....	45
2.4 Discussion and conclusion	53
2.5 Materials and methods	54
2.5.1 Biological evaluation	54
2.5.2 Chemistry	56
2.6 Abbreviation list.....	57
2.7 References	58
2.8 Supplementary data.....	64
2.8.1 Biological evaluation	64
2.8.2 Synthetic procedure for UAMC library compounds.....	68
2.8.4 References	68
Chapter 3. Identification and characterization of approved drugs and drug-like compounds as covalent <i>Escherichia coli</i> ClpP inhibitors	69
3.1 Abstract.....	70
3.2 Introduction.....	70
3.3 Results	71
3.4 Discussion.....	79
3.5 Materials and methods	80
3.6 Abbreviation list.....	83

3.7 References	84
3.9 Supplementary data.....	88
Chapter 4. Crystallography.....	97
4.1 Introduction.....	97
4.1.1 X-ray crystallography	97
4.1.2 ClpP crystal structures.....	98
4.2 Overview	100
4.3 Status	100
4.4 Comments on the crystal structure	101
4.5 References	104
Chapter 5. Conclusions and outlooks	105
Appendix	109
Part 1. Chemistry procedures - Chapter 2	109
Part 2. Synthetic procedure for UAMC library compounds – Chapter 2.....	140
List of publications	157

Abstract

Antibiotic resistance has become a world-wide problem as the number of resistant and multi-resistant bacteria has dramatically increased over the last years. The situation is even more critical for Gram-negative bacteria where new resistance mechanisms are emerging but the number of newly discovered-drugs is insufficient.

This project is part of INTEGRATE, a multidisciplinary Marie Curie Educational Training Network (ETN), where the common goal is to discover novel alternative approaches to target *Gram*-negative bacteria. The PhD project studies concerned the identification and characterization of small molecule modulators of the enzyme Caseinolytic protease proteolytic subunit (ClpP). The work has been accomplished thanks to a wide network of collaborations, within and outside of the INTEGRATE network. The majority of the study took place at the Fraunhofer IME ScreeningPort, in Hamburg. In addition, specific investigations were carried out at: the Division of Pharmaceutical Biosciences of the University of Helsinki (Finland); the Interfaculty Institute for Microbiology and Infection Medicine of the University of Tuebingen (Germany); the EMBL Grenoble (France) and in collaborations with the Laboratory of Medicinal Chemistry of the University of Antwerp (Belgium).

ClpP is a serine protease, which has been found in all sequenced bacteria to date (except *Mollicutes*), as well as eukaryotes. It has been proposed as an antibacterial target because of its central roles in many essential bacterial cellular processes, in particular, it is involved in protein turnover and homeostasis. A variety of approaches to modulate the activity of ClpP have been proposed as potentially leading to useful antibiotic effects including identification of enzyme activators as well as covalent and non-covalent inhibitors of the enzymatic function.

The main aim of this work is the identification of inhibitors directed against ClpP from Gram-negative bacteria, using *E. coli* as model organism. An *in vitro* target-centric approach was adopted to identify novel hits by screening collections of small molecules, including diverse compounds, approved drugs, and focused subsets of putative serine protease inhibitors. Selected compounds were profiled in biochemical assays (activity and selectivity) and confirmed in biophysical readouts (Surface Plasmon Resonance). Finally, for the most promising compounds, potential cytotoxicity against selected human cell lines was evaluated, and compound efficacy as an anti-bacterial in the presence and absence of selected stress condition was determined. Compounds showing the best combinations of activity, potency and selectivity were studied using *in silico* approaches to defined the hypothetical binding mode. In parallel, an ongoing structural biology effort is ongoing to reveal information on the interaction between selected compounds and ClpP and validate the computational models.

With this PhD thesis, further validation of the existence of phenotypic ClpP-related effects in *E. coli* can be found and new evidence provided for the role of ClpP as a valid target in *Escherichia coli* and Gram-negative bacteria antibiotic research. This thesis reports several compounds, coming from new scaffolds or already

approved drug libraries, active *in vitro* in the low micromolar or sub-micromolar range, with an acceptable safety profile (for the tested mammalian cell lines) and with possible ClpP-related activity in bacteria. This study must be seen as an important starting point for further development inhibitors in follow on studies.

Acknowledgments

A heartfelt thank is for who, in my private life, have always supported me and my choices. This is for my parents, my grandparents and my aunt, for Alessandro and also for my old and new friends, the ones who live close to me and the ones who live all around Europe.

Marie Skłodowska-Curie action financed my doctorate so, thanks to the European community to give me the opportunity to learn and discover a new-for-me science world. To this regard, the INTEGRATE consortium, my network colleagues (in particular Carlos), my University, my academic supervisor and the people from the labs which have hosted me must be warmly acknowledged.

Thanks to the members, past and present, of the "Fraunhofer family": my PhD life would have been much harder without your hints, help, suggestions, etc. your experience often saved me. So, thanks to Gesa, Markus, Mira, Nina, Maria, Carsten, Sheraz, Rashmi, Bernhard, Jennifer, Undine, Oliver, Manfred, Ole, Birte, Denis, Lea, Jeanette and nevertheless, to all the students.

For the same reason and furthermore, a separate mention goes to Adelia, Laura, Andrea, Thales and especially to Alessia. Thanks for convincing me that I could make it and to never let me give up.

Finally, I want to offer a special thanks to my supervisors, Phil and Björn: thanks for having followed my work, advised me and guided me during the ups and downs of the last 3 years!

Funding

This work has been financially supported by Innovative Training Networks INTEGRATE (MSCA-ITN-2014-ETN, grant agreement: 642620, "Interdisciplinary Training Network for Validation of Gram-Negative Antibacterial Targets") as part of European Union's Horizon 2020 research and Innovation programme Marie Skłodowska Skłodowska-Curie and by the Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), department ScreeningPort.

In addition, financial support has been provided by iNEXT, grant number 653706, funded by the Horizon 2020 programme of the European Commission.

Aim of the work

The impact on human health of resistant and multi-resistant bacteria has dramatically increased over the last years, while the number of new classes of antibiotics entering the clinics is decreasing, leading to the shortage of effective treatment options, especially against Gram-negative bacteria.

The goal of the project leading to this thesis is to develop novel antibacterial treatments against Gram-negative bacteria, which are able to avoid the resistance mechanisms common to marketed antibiotics. As part of the INTEGRATE project and through a close collaboration between the project partners, a multidisciplinary approach was taken, engaging a wide spectrum of competences. Specifically, computational chemistry, biochemistry, biophysics, microbiology and medicinal chemistry abilities were combined in the manner which is commonly employed in drug discovery programs.

In particular, the aim of the project was to identify covalent and non-covalent inhibitors directed against *Escherichia coli* caseinolytic protease proteolytic subunit (ClpP) which represents an attractive alternative antibacterial target. Following identification of hits, compounds would be further validated using secondary assays, to determine any off-target effects and the antibacterial efficacy. The project was progressed in close cooperation with our medicinal chemistry counterparts (group of Prof. Augustyns, University of Antwerp) in the INTEGRATE consortium. Here we aim to demonstrate the importance of ClpP as a new target against *E. coli* and Gram-negative infections since this protein has been widely studied during the last decades as a Gram-positive antibacterial target, but very few studies report on ClpP as a new approach in Gram-negatives drug discovery.

Chapter 1

Introduction

1.1 Antibiotics Era and Resistance

With the discovery of penicillin in 1929 by Alexander Fleming, the world entered a period called the antibiotic golden era, during which most antibiotics were discovered. The classic antibiotics can be divided into two categories: bacteriostatic, if their effect is to inhibit bacterial growth and bactericidal, if they are able to kill the microorganism. Representative mechanisms of anti-bacterial action can be to inhibit protein synthesis (e.g. aminoglycosides), cell wall synthesis (e.g. β -lactams), folate synthesis (e.g. sulfa drugs), DNA or RNA synthesis (e.g. fluoroquinolones) and depolarize membrane potential (e.g. daptomycin) ¹.

In the last decades, the rate of antibiotic drug discovery has declined, mostly because the pharmaceutical companies are less focussed on developing those classes of drugs due to economic reasons: the high investments required for developing new antibiotics are not covered by the limited revenues that can be realized ^{2,3}.

With time, antibiotics efficacy is decreasing (see Figure 1). The emergence of resistance towards particular antibiotics is a ubiquitous phenomenon, which in some cases occurs even before the antibiotics themselves become widely used in the clinic (e.g. penicillin) ^{4,1}. The reason behind the evolution of resistant is the selection pressure that the microorganisms face when coming into contact with antibiotics ¹. Furthermore, epidemiological studies confirm the direct correlation between antibiotic consumption, the rise and dissemination of resistant bacteria strains ⁵.

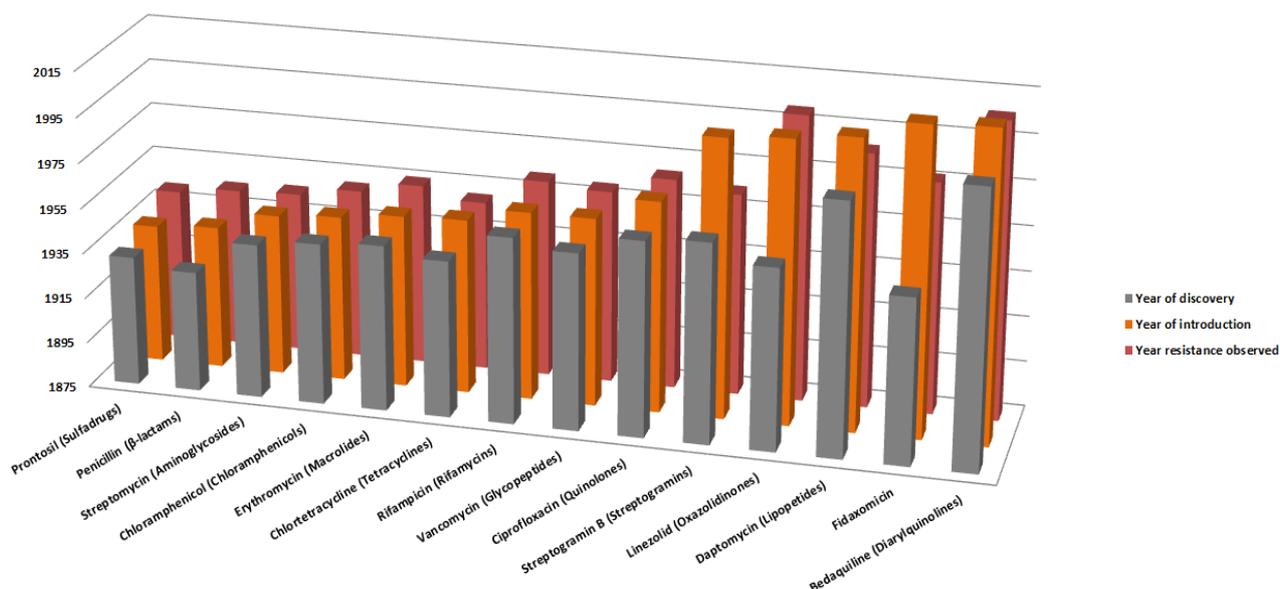


Figure 1. Timeline of antibiotics discovery (grey), market introduction (orange) and when resistance was first observed (red). The data used to build the graph were extracted from Platforms for antibiotic discovery⁶.

The World Health Organization (WHO) report on antimicrobial resistance in 2014⁷ underlined the critical importance of continued research investment to progress antibiotic drug discovery in the future. Antibiotic resistance is a world-wide problem, especially for those countries where the antibiotics were used in the past (and still now) with inadequate controls. For example, a high number of inappropriate prescriptions, incomplete compliance to treatment regimens and widespread indiscriminate use in livestock production. All these factors contribute to the emergence of pathogenic bacteria with resistance towards antibiotics. The WHO has estimated that every year about 2 million worldwide (around 25 000 in Europe) people die due to bacterial infections.

The number of cases of multi-drug resistant bacteria (MDR) is increasing and every year in Europe about 40000 resistant strains are identified whilst very few promising molecules to fight against bacterial infections are to be found in the medicine chest. According to the WHO, in 2013 there were over 480,000 cases of multidrug-resistant tuberculosis (MDR-TB) and even more dangerously, in over 100 countries extensively drug-resistant tuberculosis (XDR-TB) has been reported⁷. Methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant enterococci (VRE), and multidrug-resistant Gram-negative bacteria (all beta-lactam antibiotics) are deadly hospital-acquired infections caused by bacteria with high resistance¹. Among the species that bring severe risks a group of multidrug-resistant bacteria (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species), termed ESKAPE bacteria have been highlighted. These bacteria are particularly dangerous for immunocompromised individuals⁸.

A total of 30 new antibiotics were launched on the market in the period between 2000 and 2015, plus 2 new combinations of β -lactam and β -lactamase inhibitor. Within those, only five were first-in-class antibiotics and all of these were active against Gram-positive bacteria. Only a new diazabicyclooctane (DBO)-type β -lactamase inhibitor (avibactam, launched in 2015) was marketed because of its efficacy against Gram-negative⁹. It is particularly alarming that there are almost no antibiotics against Gram-negative bacteria which do not have resistance issues^{3,10}, whilst at the same time, pathogenic Gram-negative MDR strains (such as extended-spectrum beta-lactamase-producing *Escherichia coli* and *Neisseria gonorrhoeae*) are becoming increasingly prevalent in the community¹⁰.

Bacteria can respond to an antibiotic with different methods⁸⁴. Among the antibiotic resistance strategy of bacteria, a classification based on genetics methods involves intrinsic and acquired resistance. Intrinsic resistance comes from own property of the bacteria to not be affected by the antibacterial molecule. On the other hand, when the resistance derives from mutation or incorporation of external genetic material⁸⁵, for example plasmids, as by conjugation from bacteria that already are resistant, transduction by virus or transformation¹.

Generally, the most common mechanisms to prevent the entrance of a drug are the modification of the cell penetration or the expulsion through efflux pumps. Bacteria can furthermore fight antimicrobial compounds altering the target site or through alteration or destruction of the antibiotic^{1,86}. A representative summary of the main antibiotic targets and the antibiotic resistance mechanisms is shown in Figure 2.

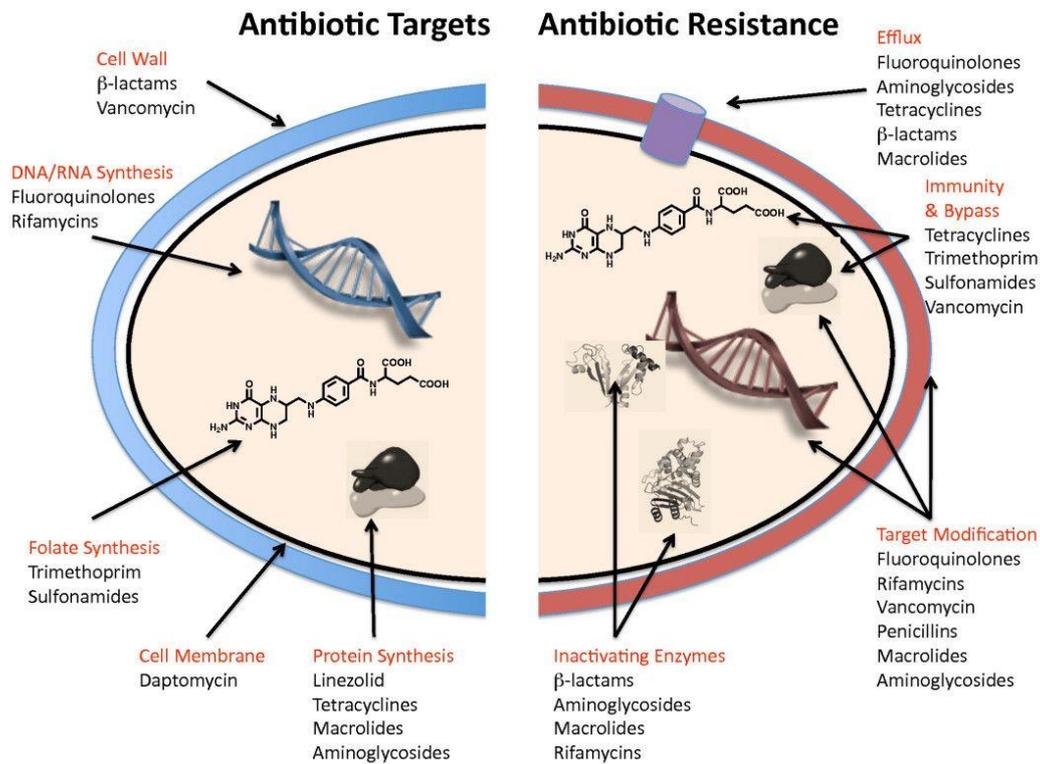


Figure 2. A summary of principal antibiotic targets and mechanisms of bacterial resistance ⁸⁶.

The inactivation of the molecules acting against bacteria happens either by the destruction of those, as in the case of the beta-lactams, where the ring of the molecules is destroyed by the enzymes called beta-lactamases, or the chemical modification by hydrolysis, redox reactions or by the addition of a chemical group (through reaction of phosphorylation, acetylation or adenylation) ^{84,87}

Since most of the targets are inside the cell, reduce the amount of antibiotic inside the intracellular space is a further mechanism of resistance ^{87,88}. This can be achieved decreasing the possibility of the drug to permeate and enter in the cell by modifying the expression of the type or of the number of trans-membrane channels (called porins) or by defecting protein function ⁸⁷. Also, instead of the uptake mechanism, bacteria can implement the mechanism to expel xenobiotics through the expression of efflux pumps, membrane proteins with the function of pumping out the antibiotics ⁸⁸

This mechanism is especially important in Gram-negative bacteria where the presence of an additional inner membrane limits the access of antibacterial ⁸⁷.

The last of the primary resistance mechanisms concerns/involves the modification of the target to reduce or to eliminate the affinity of the drug for it. This can be achieved, depending on the type of ligand/target, in different ways, as changing the peptidoglycan part of the cell wall or altering the protein synthesis, resulting in a modification of the drug binding site, or interfering at structural gene level ⁸⁴.

1.2 Overview of ClpP as a drug target

1.2.1 Bacterial proteases

Proteases play significant roles in many physiological properties of bacteria, such as cell-division and stress response (e.g. heat shock) ^{11,12}. This family of enzymes is also involved in monitoring protein quality and cellular homeostasis by removing damaged or misfolded proteins and short-lived regulatory proteins ^{13,18,19,20}. Typically, the chaperones associated to the proteases are termed “ATPases Associated with diverse cellular Activities” (AAA+), which use ATP hydrolysis to unfold the structure of the protein substrate and translocate it to a proteolytic chamber formed by the protease ²¹. Among the proteases, the important for bacterial organisms are: ClpP (Caseinolytic protease P) ^{11,16,17} and its associated ATPases (e.g. ClpX, ClpA or ClpC) ²¹; Heat Shock Locus V (HslV, known as ClpQ, with the ATPases HslU), and Lon, FtsH and 20S Core Particles (uses by *Actinobacteria* as Mpa–20S proteasome) ^{14,15,21}.

1.2.2 Caseinolytic protease P

ClpP is well studied chymotrypsin-like serine protease which exhibits a variety of different roles and functions within individual bacteria and between different bacteria. The main function of ClpP in many bacteria is to control protein quality by degradation of mistranslated, misfolded or aggregated proteins, which are created as result of stress factors (heat, antibiotics, etc.) ²³. It has been also considered as target, and several studies have focused on the identification of ClpP activity modulators due to the significant role ClpP plays in bacteria cell homeostasis and function ²². It is also involved in proteolytic regulation of transcription factors, and in *E. coli*, ClpP adjusts the level of RecN to control the duration of SOS response (global response to DNA damage, to repair) ²⁶, as well as nitric oxide stress responses. In *Bacillus subtilis*, ClpP regulates motility, exoenzyme synthesis, genetic competence and spore formation ^{28,29}. In *Caulobacter crescentus*, ClpP governs cell-differentiation ³⁰. Also, it has been demonstrated that in some actinomycetales, ClpP, is essential for cell-growth (*in vitro*) ^{23,31}. It has been reported that ClpP has involved modulation of the expression of virulence factors and, consequently, in virulence and stress response in several bacteria, for example *Staphylococcus aureus*, *Listeria monocytogenes* and *Streptococcus pneumoniae* ^{32–35}. In *Listeria monocytogenes* it has been also found to essential for macrophage infections ³².

1.2.2.1 Prevalence among organisms

ClpP is wide conserved among bacteria (including human pathogens) and many eukaryotes, although it is not found in mollicutes, archaea and some fungi ^{13,36}. First discovered in *E. coli*, in most bacteria ClpP appears as a single variant, although exceptions are *actinobacteria*, *chlamidiae*, *cyanobacteria* and some other species, where it is present as multiple variants. One interesting observation is that the usual catalytic triad is not present in the cyanobacteria paralog ClpR ^{37,38}.

ClpP has been also found in human, localized in the mitochondrial matrix and with a function of regulation and degradation of electron transport chain related enzymes and modulating their response to the presence of unfolded proteins^{22, 39}. Encoded by chromosome 15, human ClpP (hClpP), in order to become mature ClpP needs to be processed and lose 56 N-terminal amino acid residues (including the pre-sequence and the tag-sequence)¹³. The human protein showed a high similarity (71%) and sequence identity (56%) with *E. coli* ClpP, but presents 28 adjunctive C-terminal residues, that seem to be a peculiarity of mammalian ClpPs⁴⁰. The catalytic triad, here, is present in the region between the interface of the head and the handle domains.

Human ClpP (hClpP) is typically found as a single heptameric ring. For the formation of the fully functional double ring with ClpX, the presence of ATP is required. There is no evidence of a human analog of ClpA which is found in bacteria. hClpP does not present protease activity and low level of peptidase activity, and targets a different range of substrates compared to those described for the *E. coli* ClpPX complex⁴¹.

In plants, there are 5 ClpR paralogs, may evolve from cyanobacteria, 6 ClpP paralogs plus 10 Clp's AAA+ chaperons, identified in the chloroplast stoma of *Arabidopsis thaliana*¹³.

1.2.2.2 ClpP structure

The self-assembled serine protease ClpP creates a cylindrical structure with barrel shape of 90 Å of diameter, made by 14 subunits, almost always identical, organized in 2 heptameric rings, enclosing a central proteolytic chamber (called also axial pore), of 50 Å of diameter, containing 14 (one for each subunit) proteolytic active sites²³ (see Figure 3 A&B).

Each monomer is composed of three sub-domains (see fig 3 C): handle, head and N-terminal regions¹¹. The head domain represents The main body of ClpP, called the head domain, is globular and is conserved across organisms. This domain forms hydrogen bonds with substrates⁴³. The handle domain is quite flexible and is responsible for the establishment of the interaction between the two rings due to its stand-turn-helix motif¹¹. Last of the three, the N-terminal domain, also preserved in the various species, can be further divided into two parts^{44,16}. Firstly the “axial protrusion”, a loop coming out from the pore of the inner chamber made by charged and hydrophilic residue, and secondly the “axial pore lining”, which is hydrophilic and interacts with the head domain¹⁶.

In most bacterial species, each active site has the typical serine-histidine-aspartic acid catalytic triad, which is physically protected from the cytoplasmic environment, due to its position inside the cylinder^{11,23}. The substrates can bind simultaneously in more than one active site and generates, due to proteolytic activity, peptides that have a size of 7-8 residues^{20,42}.

The active sites are located between the head and the handle domains, and this arrangement is thought to be the reason why the orientation of the catalytic triad can be alternated by the flexibility of the handle region¹³. The necessity for conformational change in the handle domain and the presence of interactions between the

two rings which form ClpP have been confirmed to be essential for the activation of the peptidase activity⁴⁵. This system, together with a narrow entrance of the pore only allows entry to peptides with a maximum 30 residues length, contributes greatly in determining the specificity of proteolysis events^{23,24}.

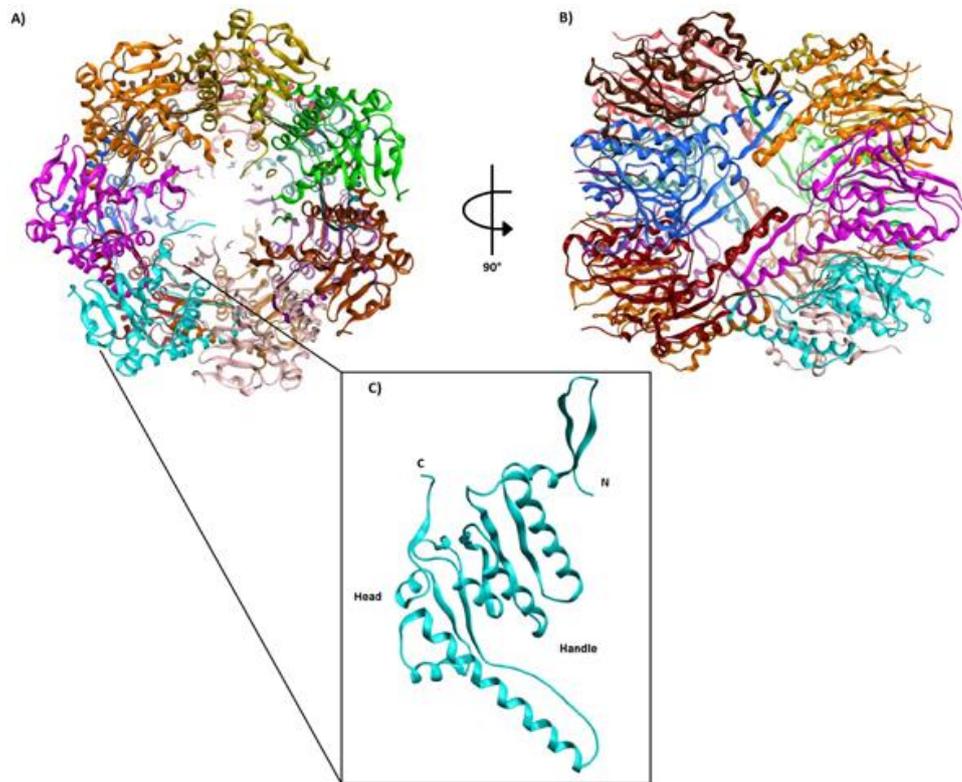


Figure 3. *E. Coli* ClpP structure, dyed by monomers. A) pore view; B) side view; C) single monomer structure.

Three possible ClpP conformation states have been reported in X-ray structural studies: extended (ordered and active), compact and compressed (both inactive because of different orientation of respect catalytic triad and side chain that brings alteration of pore size.)²⁰.

Some examples of organisms where the different conformations have been found are²⁰:

- Extended. *Escherichia coli*⁴⁶, *Helicobacter. pylori*⁴⁷, *Coxiella Burneti*, *Francisella Tularensis*, *Listeria monocytogenes* ClpP2 and human mitochondria²⁰.
- Compact. *Escherichia coli*⁴⁸, *Streptococcus pneumoniae*, *Micobacterium tuberculosis*, *Listeria monocytogenes*, *Plasmodium falciparum*, *Bacillus subtilis*⁴⁹ and *Staphylococcus aureus*⁴⁵.
- Compressed. *Staphylococcus aureus*^{45,50}.

1.2.2.3 Gate opening and product realize

Accordingly to a proposed mechanism, the substrates enter the chamber with the orientation N to C terminal and the substrate entering seems to be regulated by a network of hydrogen bonds associated with the amino

acids located close to the active site⁴³. Moreover, for the release of the products, two hypotheses have been described. The first model proposes the product exits happens *via* passive diffusion from the same entrance pore used from the peptide/protein. This model implies that the substrate binding ATPases then dissociate from ClpP¹³. In the second model, confirmed by studies which involved equatorial region mutated ClpP, transient side pores arising in the interface between the rings are used by the product to exit the chamber^{20,44}.

1.2.3 Clp complex

Although ClpP itself is also able to perform peptidase activity, in order to achieve the full range of proteolytic functions, it requires specific ATPases, called AAA+ partners²³. Through the energy produced by the hydrolysis of ATP, the ATPase, if the refolding process of degraded or misfolded proteins fails, are capable to recognize the substrate and translocate it inside the inner chamber after then unfold the substrate protein structure^{23,51}. Specific degradation tags, called degrons, are used to tag the N or C terminal of to be degraded proteins in the intracellular environment (e.g. the 11 C-terminal amino acid SsrA-tag in *Escherichia coli*) or when the not properly excreted extracellular proteins^{24–27}.

Those, belong to the Hsp100 family, have also hexameric-chaperone structure and, bind one or typically both (when in couple) the axial surfaces of the ClpP, forming a holoenzyme complex with the shape of a barrel, called the Clp complex^{13,11}. The interaction with the ATPase partners via specific loops with containing a specific conserved motif that emerges from the ATPases surface²³ and hooks the ClpP N-terminal loops²⁴.

In the *Escherichia coli* Clp complex, the AAA+ partners are ClpA and ClpX. ClpA is an 83 kDa ATPase that contains two AAA+ domains, while ClpX contains only one AAA+ domain (46 kDa), which is homologous to the second AAA+ domain of ClpA. In other bacteria are present ClpA paralogs, such as ClpC, ClpE, and ClpL¹³. The different types of ATPases define the substrate specificity for the proteolysis²³.

1.2.4 ClpP in drug discovery

Due to its physiological role, as shown in the examples reported below, ClpP and its AAA+ partners, have attracted interest over the last years as antibiotic drug target²³.

ClpP is important for the bacterial virulence for example in *Listeria monocytogenes*, where ClpP2 is required to produce and secrete an exotoxin (listeriolysin) which is essential for intracellular growth of pathogenic agents within macrophages. Consequently, a study on a ClpP deletion mutant has confirmed that the WT form is more infective than the mutant, which is totally a-virulent³³. Another example is *Streptococcus pneumoniae*, where *clpX* and *clpP* deletion gave rise to a down-regulation in several transcription encoding virulence proteins, toxins, extracellular proteases, lipases and hydrolases²³.

A study on ClpP's role in *Haemophilus parasuis* discovered that a *clpP* defective mutant is substantially more susceptible to stress, including high temperature, osmotic pressure and oxidative stress, in comparison to wild

type bacteria. Moreover, the mutant performs a higher rate of auto-agglutination and biofilm formation, compared to the wild type, where *clpP* is present. The role of ClpP in biofilm regulation has been reported but in a species-dependent manner⁵².

Nitric oxide stress has been reported to modulate the levels of ClpP in *Streptococcus pneumoniae*⁵³. In a similar way, mutant *clpP* strains of *Escherichia coli* show a difference in growth compared to the isogenic wild type strains in the presence of the nitric oxide donors⁵⁴.

In the drug discovery context, ClpP has previously been targeted in such a way as to further activate its proteolytic function or inhibiting proteolytic activity by active site interaction or structural modifications.

1.2.4.1 ClpP activators

Recent investigations have focussed on a class of small antibacterial peptides called acyldepsipeptides (ADEPs) and their synthetic derivatives⁵⁵. The ADEPs were identified for the first time in a study in 2005 and shown to be active against ClpP Gram-positive bacteria^{51,80}. ADEPs have been also found to be active against gram-negative bacteria, in particular, *Neisseria meningitidis* and *Neisseria gonorrhoeae*, in a study which involved screening of 26 macrocyclic core residues or N-acyl side chain ADEP-analogues, that show activity against Gram-negative and improved activity against gram-positive bacteria⁵⁶.

The mechanism of action consists in the binding of the molecules in sites where the ATPases bind with ClpP resulting in an activation of the protein and, as a consequence, uncontrolled proteolysis⁷⁹.

Among the ADEPs, for example, ADEP 4 was found to be active against the bacterial growth of a panel of Gram-positive bacteria and be efficient in *S. aureus* mice infections^{80,83}. Although ClpP is not essential in *Staphylococcus aureus*, in combination with ciprofloxacin, rifampicin or linezolid, it significantly reduced the formation of bacterial colonies^{57,81} and eradicated clinical isolates of *Enterococcus faecium* at 0.2 μM ^{82,83}.

1.2.4.2 ClpP inhibitors

Other ClpP drug discovery strategies are focused on the inhibition of proteolytic activity, mainly targeting the active catalytic site of the monomers. Inhibitors in this category are either covalent or non-covalent in terms of their mechanism of action. For the summary of reported inhibitors see Table 1.

The inhibitor F2 has been reported to suppress the ClpXP protease in *Bacillus anthracis* and drug-resistant *Staphylococcus aureus*, in a synergistic manner with bacterial cell envelope antibiotics and cathelicidin antimicrobial peptides. Also, it has been reported that in a pulse-chase assay in the presence of F2 the substrate proteolysis activity was reduced in *Escherichia coli*: this was determined by measuring the half-life of GFP with *Escherichia coli* tmRNA tag, that was found to increase $t_{1/2}$ from approximately 5.4 to 18.7 minutes⁵⁸.

In the case of covalent mechanisms, beta-lactones inhibitors are a class of compounds that have emerged from screening studies as being covalent inhibitors of ClpP. A reaction between the antibiotic and the catalytic serine blocks the active site and produces a catalytically inactive beta-hydroxyacyl-enzyme complex which serves to irreversibly block proteolytic activity⁵⁹. Several studies on a series of beta-lactones have been performed including work on the D3 compound, that was developed by Böttcher and Sieber using a proteomic experiment where ClpP was chemically labeled in situ⁶⁰. Further extensions of this work gave rise to improved compound U1⁶¹, the first active against *Staphylococcus aureus* and in *Listeria monocytogenes* capable of reducing virulence in macrophages^{60,61}. However, stability issues linked to the cyclic ester part of the beta-lactones that is hydrolyzed rapidly in plasma have been reported and have hindered the development of this compound^{22,62}.

In further studies, employing High-throughput screening where *Staphylococcus aureus* ClpP proteolysis activity was evaluated using measuring turnover of a fluorogenic substrate, several inhibitors have been identified. One example, AV170, irreversibly binds the active site⁶², and another, the non-covalent inhibitor AV145, blocks the active site by inducing a conformational change in the protein⁶³. The compound AV170, representing a new class phenyl ester ClpP inhibitors, was found to be active with IC₅₀ estimated around 1 μM, and it not active against human ClpP and showed increased plasma stability compared to beta-lactones⁶². Starting from AV145, a compound with an IC₅₀ below 10 μM and with knowledge of its reversible binding mode (shown in crystal structure), further compounds have been rationally synthesized with improved *in vitro* activity, the best of which is AV280 with an IC₅₀ < 1 μM and a similar binding mode⁶³.

The antimicrobial activity of the cyclopeptide cyclomarin A has been reported in *Mycobacterium tuberculosis*²³. This compound seems to be quite specific to this pathogen since tests in bacteria cultures of several other strains (Gram-positive and negative) did not reveal any efficacy^{23,64}. Affinity chromatography analysis has shown binding of the compound with the ClpC1 ATPases subunit of Clp complex in *Mycobacterium tuberculosis*: cyclomarin A is more a Clp complex inhibitor than ClpP's. It inhibits the protease activity even though does not directly binds to ClpP but interacting with the ClpC1 blocks the unfolding of the substrates into the chamber⁶⁴.

Benzyloxycarbonyl-leucyltyrosine chloromethyl ketone (Z-LY-CMK), a covalent inhibitor of the Gram-negative *Escherichia coli* was discovered by Szyk and Maurizi⁴³. The *in vitro* activity in a fluorescence-based assay addressing the ClpP proteolytic activity, showed an IC₅₀ around 50 μM. The crystal structure of the protein in presence of the inhibitor has been solved and elucidated the covalent binding mode with the catalytic serine⁴³ (see Figure 4).

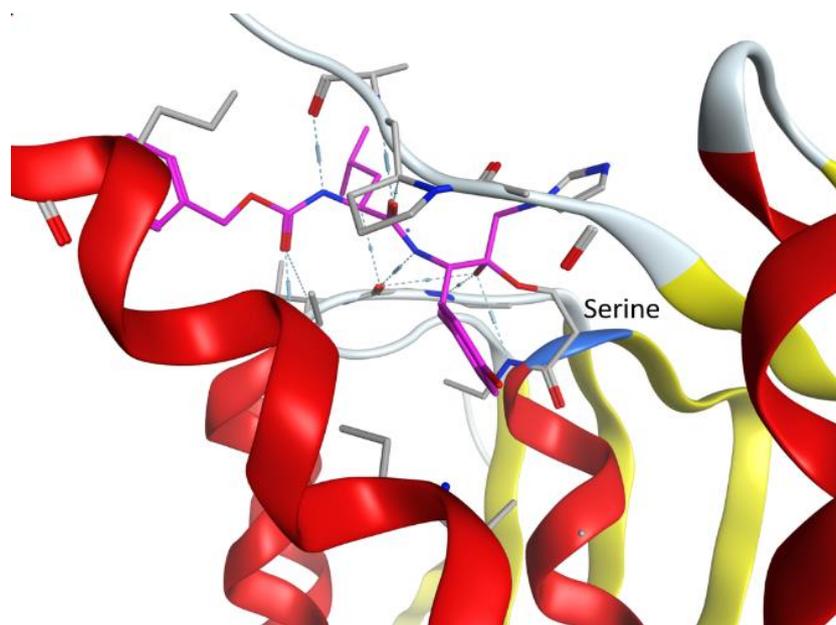
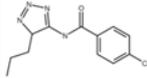
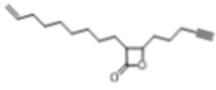
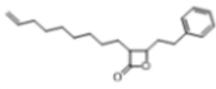
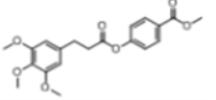
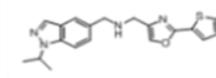
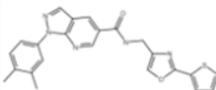
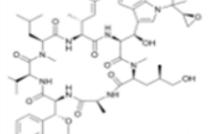


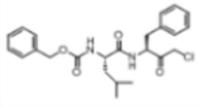
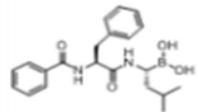
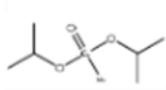
Figure 4. ClpP interacting covalently with Z-LY-CMK inhibitor (PDB ID 2FZS)

Boron derivatives were discovered to inhibit the activity of peptidase and protease activity and growth *Mycobacterium tuberculosis*⁶⁵. A study published in 2015⁶⁶ identified bortezomib as potent *Mycobacterium* inhibitor in a whole-cell screening using a *Mycobacterium* reporter strain. Further tests in bacteria cells confirm that this compound, which is a non-specific proteasome covalent inhibitor, is active in *Mycobacterium smegmatis*, *Mycobacterium bovis* and *Mycobacterium tuberculosis*⁶⁶.

Finally a crystal structure of *Bacillus subtilis*, in compressed form, in complex with diisopropylfluorophosphate (DFP), a serine protease inhibitor, shows an irreversible covalent attachment of the compound to the catalytic serine⁴⁹.

Table 1. Summary of known inhibitors of ClpP (and Clp complex): here each structure is shown and the properties are listed.

<p>F2</p> 	<p>Test in bacteria. Pulse chase assay: from 5.4 to 18.7 minutes (<i>E.coli</i>); Growth: Not detectable ([F2] = 40 μM), reduced by 100.000 fold in combination with LL-37 (1.6 μM), reduction in combination with penicilline (5mg/ml) and daptomycin (2-4 μg/ml) (<i>B. anthracis</i> and <i>S. aureus</i>) Cytotoxicity against mammalian cells. Not detectable at concentrations of 500 and 1000 μM Mode of action. Inhibition of proteolysis activity (general)</p>
<p>D3</p> 	<p>In vitro activity. Acyl-enzyme stability: $t_{1/2} = 5.0 \pm 0.4$ hours (Human) Test in bacteria. Chemical label of ClpP: with concentration of 1.3 μM (<i>S. aureus</i>) Hemolysis: EC_{50} 34 nM (<i>S. aureus</i>) and 80 (<i>L. monocytogenes</i>); Proteolysis: EC_{50} 11 nM (<i>S. aureus</i>) Mode of action. Inhibition (irreversible) of the active site with formation of beta-hydroxyacyl-enzyme complex</p>
<p>U1</p> 	<p>Test in bacteria. Hemolysis: EC_{50} 30 μM (<i>S. aureus</i>) Cytotoxicity against mammalian cells. Not detectable at concentration of 1000 μM Mode of action. Inhibition (irreversible) of the active site with formation of beta-hydroxyacyl-enzyme complex</p>
<p>AV170</p> 	<p>Test in bacteria. Peptidase activity: IC_{50} 0.3-1.3 μM (<i>S. aureus</i>); Proteolytic activity (ClpXP): $IC_{50} < 10$ μM (<i>S. aureus</i>) acyl-enzyme stability: $t_{1/2} = 8.2 \pm 0.8$ hours (Human) Mode of action. Inhibition (irreversible) of the active site with formation of acyl-enzyme complex and release of phenol</p>
<p>AV145</p> 	<p>In vitro activity. Proteolytic activity: $IC_{50} < 10$ μM (<i>S. aureus</i>); Crystal structure (<i>S. aureus</i>) Mode of action. Inhibition (reversible) with conformational change due to rotation of residue close to the active site (<i>S. aureus</i>)</p>
<p>AV280</p> 	<p>In vitro activity. Proteolytic activity: $IC_{50} < 1.5$ μM (<i>S. aureus</i>); Mode of action. Inhibition (reversible) with conformational change due to rotation of residue close to the active site (<i>S. aureus</i>)</p>
<p>Cyclomarine A</p> 	<p>In vitro activity. Affinity chromatography with ClpC1: $IC_{50} = 10$ nM (<i>M. tuberculosis</i>) Test in bacteria. Growth: Inhibited around 2.5 μM (<i>M. tuberculosis</i>) Mode of action. Inhibition of ClpC1</p>

<p>Z-LY-CMK</p>  <p>The chemical structure of Z-LY-CMK is a dipeptide derivative. It features a benzyl group attached to the nitrogen of a lysine residue, which is linked to a tyrosine residue. The tyrosine residue has a chlorine atom on its para position.</p>	<p>In vitro activity. Proteolytic activity: estimated IC_{50} around 50 μM (<i>E. coli</i>); Crystal structure (<i>E. coli</i>)</p> <p>Mode of action. Inhibition (irreversible) of the active site with catalytic serine and histidine</p>
<p>Bortezomib</p>  <p>The chemical structure of Bortezomib is a boronic acid derivative. It consists of a central boron atom bonded to a hydroxyl group and a hydroxymethyl group. The boron atom is also bonded to a phenyl ring and a side chain containing a benzamide group and a phenyl ring.</p>	<p>Test in bacteria. Growth: $MIC_{50} = 4$ (<i>M. smegmatis</i>), $MIC_{50} = 0.3$ (<i>M. bovis</i>), $MIC_{50} = 0.8$ (<i>M. tuberculosis</i>)</p> <p>Mode of action. Inhibition (irreversible) of the active site with catalytic serine</p>
<p>DFP</p>  <p>The chemical structure of DFP (diisopropyl fluorophosphate) shows a central phosphorus atom double-bonded to an oxygen atom and single-bonded to a fluorine atom. The phosphorus atom is also bonded to two isopropyl groups.</p>	<p>In vitro activity. Crystal structure (<i>Bacillus subtilis</i>)</p> <p>Mode of action. Inhibition (irreversible) of the active site with catalytic serine</p>

1.3 Covalent drugs

Molecules which form covalent (irreversible) interaction with their targets proteins are very important therapeutics agents especially in the anti-infective field ⁶⁸. In fact, almost a third of the drugs on the market act in a covalent associated manner ^{69,70}. The first covalent drug in common use was aspirin which acetylates a serine residue in the active site of the cyclooxygenases COX-1 and COX-2 ⁷². Other examples are β -lactam antibiotics which kill bacteria by inhibiting the synthesis of the peptidoglycan layer of bacterial cell walls. This is achieved by acetylation of a serine residue in the active site of penicillin's binding protein that catalyzes the crosslinks of the nascent peptidoglycan layers or by inhibiting the active site of proteases ⁷¹.

Table 2. List of covalent drugs approved by the US Food and Drug Administration (FDA). The data used to build this table are from a paper published in 2015 by Kumalo and collaborators ⁷⁰.

Drug Name	Target	Therapeutic Domain
Amoxicillin	PBP	Anti-infective
Ceftriaxone/Rocephin		
Cefaclor/Ceclor		
Cefuroxime axetil/ceftin		
Cephalexin/keflex		
Meropenem		
Omnicef PBP		
Penicillin V		
D-cycloserine/seromycin	Alanine racemase	Cardiovascular
Fosfomycin/monurol	UDP-N-acetylglucosamine-3-enolpyruvyl-transferase	
Isoniazid	Enol-acyl carrier protein reductase	
Exemestane/Aromasin	Aromatase	
Floxuridine	Thymidylate synthase	
Gemcitabine/gemzar	Ribonucleoside reductatas	Gastro-intestinal
Proscar/finasteride	5- α -Reductase	
Warfarin	Vitamin K reductase	
Phenoxy-benzamine hydrochloride	α -Adrenoceptor	
Prevacid/lansoprazole	H+/K+ ATPase	Gastro-intestinal
Nexium/esomeprazole		
Prilosec/omeprazole		
Protonix/pantoprazole		
Aciphex/rabeprazol		
Orlistat	Lipase	Cancer
Azacytidine	Methyltransferase	
Decitabine/azadC		
Bortezomib	Protesome	
Dutasteride/avodart	5- α -Reductase	
mercaptopurine/purinthol	Purine-nucleotide synthesis	Anti-diabetic
Saxagliptin/Onglyza	DPP-IV	
Vildagliptin/Eugreas		

Rasagiline	MAO-B	Parkinson's disease
Selegiline		
Disulfiram/antabuse	Aldehyde dehydrogenase	Chronic alcoholism
Carbidopa/lodosyn	DOPA decarboxylase	CNS
Vigabatrin/sabril	GABA-Aminotransferase	Anti-epileptic
Eflornithine	Ornithine decarboxylase	Hirsutism
Propylthiouracil/procasil	Thyroxine-5-deiodinase	Hyperthyroidism
Aspirin	Cyclooxygenase	Inflammation

According to the data published by Kumalo and Table 2 ⁷⁰ the majority of the approved drugs in the US market are molecules acting as anti-infective (around 21% of the total of covalent drug), like antibiotics, followed by drugs interacting with cardiovascular targets and treating gastrointestinal disorders (around 16%). Another huge slice of the covalent drugs on the market is represented by the anti-cancer drugs, with a percentage of approx. 13%. ⁷⁰

Several are the advantages of covalent drugs, the most important are listed below:

- High potency (due to the target blockage) ⁷⁴;
- Low dose (a consequence of high potency) ⁷¹;
- Extended duration of action (thanks to the limited turnover) ⁷³;
- High selectivity ⁷³.

Notwithstanding these advantages, there were safety concerns in the past about covalent drugs, especially correlated with potential immunogenicity of the resulting target adduct (could cause allergic response and drug hypersensitivity reaction) ⁷⁵ and drug-toxicity due to hyper reactive warheads ⁸⁹.

A number of approaches may be adopted to reduce the risk of adverse reaction to this class of enzyme inhibitors ⁷⁶. At the high level, this means fully profiling the compound selectivity for the target protein in order to decrease the risk of toxicity correlated with off-target interactions ^{73,77}. To accomplish this, the initial reversible target-inhibitor association can be optimized or, using bioinformatics techniques, it is may be possible to improve the position of the inhibitor warhead in its interaction with the nucleophilic center ⁷⁶. Another alternative is maintaining a safe low dosage of the drugs, and selecting the candidate's molecules with good oral availability. For some drugs, for example, 10 mg/day can lead to with hepatic injury or other cytotoxicity issues while dosing more than 50 mg/day in specific case increase the possibility to have serious adverse effects ⁷⁸.

1.3.1 Mechanism of action

The general mechanism of action (see Eq. 1) is at the first reaction step the formation of an initial reversible association covalent drug/inhibitor (AB) with the target protein. After this intermediate step, occurs the creation of a covalent bond between the ligand electrophile part and the protein nucleophilic center (A-B) ⁷³.



Equation 1. Mechanism of covalent bond formation between a ligand and a protein, where A and B represent the enzyme and covalent drug.

This reaction describes the different inhibition which could take place, in accordance with the nature of the molecule. k_{-2} and k_2 are essentially the parameters that establish the strength of the bond and inhibitory potency of the compound. In fact, when k_{-2} is equal to zero, the irreversible inhibition happens, with a permanent and complete block of the protein. In this case, the dissociation time will be higher than the time for the protein turnover. On the other hand, if k_{-2} and k_2 have infinite values, the situation faced is of a formation of covalent but reversible inhibition. Lastly, the protein turnover time the dissociation time of the complex is lower compared with the dissociation time the reaction appears to be completely reversible with k_{-2} is equal to zero. Here the covalent complex AB will never be formed ⁷¹.

1.4 References

1. Chellat, M. F., Raguž, L. & Riedl, R. Targeting Antibiotic Resistance. *Angewandte Chemie - International Edition*. 55 (23), 6600-26 (2016).
2. Bartlett, J. G., Gilbert, D. N. & Spellberg, B. Seven ways to preserve the Miracle of antibiotics. *Clinical Infectious Diseases*. 56 (10), 1445-5 (2013).
3. Golkar, Z., Bagasra, O. & Gene Pace, D. Bacteriophage therapy: A potential solution for the antibiotic resistance crisis. *Journal of Infection in Developing Countries*. 8 (2), 129-36 (2014).
4. Palumbi, S. R. Humans as the world's greatest evolutionary force. *Science*. 293 (5536), 1786-90 (2001).
5. Ventola, C. L. The Antibiotic Resistance Crisis Part 1: Causes and Threats. *P&T*. 40 (4), 277-283 (2015).
6. Lewis, K. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov*. 12 (5), 371-87 (2013).
7. WHO. Antimicrobial resistance Global Report on Surveillance. (2014).
8. Santajit, S. & Indrawattana, N. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. *Biomed Res. Int*. 2016 (2), 1-8 (2016).
9. Butler, M. S., Blaskovich, M. A. T. & Cooper, M. A. Antibiotics in the clinical pipeline at the end of 2015. *J Antibiot (Tokyo)*. 70 (1), 3-24 (2016).

10. Rossolini, G. M., Arena, F., Pecile, P. & Pollini, S. Update on the antibiotic resistance crisis. *Curr. Opin. Pharmacol.* 18, 56-60 (2014).
11. Alexopoulos, J. A., Guarne, A. & Ortega, J. ClpP: a structurally dynamic protease regulated by AAA+ proteins. *J. Struct. Biol.* 179 (2), 202-210 (2012).
12. Gottesman, S., Wickner, S. & Maurizi, M. R. Protein quality control: Triage by chaperones and proteases. *Genes and Development.* 11(7), 815-23 (1997).
13. Yu, A. Y. H. & Houry, W. A. ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Letters.* 581 (19), 3749-57 (2007).
14. Humbard, M. A. & Maupin-Furlow, J. A. Prokaryotic proteasomes: Nanocompartments of degradation. *J. Mol. Microbiol. Biotechnol.* 23 (4-5), 321-34 (2013).
15. Inobe, T. & Matouschek, A. Paradigms of protein degradation by the proteasome. *Current Opinion in Structural Biology.* 24 (1), 156-164 (2014).
16. Liu, K., Ologbenla, A. & Houry, W. A. Dynamics of the ClpP serine protease: A model for self-compartmentalized proteases. *Critical Reviews in Biochemistry and Molecular Biology.* 49 (5), 400-412 (2014).
17. Wickner, S., Maurizi, M. R. & Gottesman, S. Posttranslational quality control folding, refolding, and degrading proteins. *Science.* 286 (5446), 1888-93 (1999).
18. De Mot, R., Nagy, I., J. W. & W. B. Intracellular proteases in prokaryotic cells perform many tasks, including cleavage of signal peptides during protein export, timely inactivation of regulatory proteins, and removal of aberrant nonfunctional proteins. *Trends in Microbiology.* 88 (9), 88 (1999).
19. Pickart, C. M. & Cohen, R. E. Proteasomes and their kin: proteases in the machine age. *Nat. Rev. Mol. Cell Biol.* 5 (3), 177 (2004).
20. Liu, K., Ologbenla, A. & Houry, W. A. Dynamics of the ClpP serine protease: A model for self-compartmentalized proteases. *Critical Reviews in Biochemistry and Molecular Biology.* 49 (5), 400-412 (2014).
21. Sauer, R. T. & Baker, T. A. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. *Annu. Rev. Biochem* 80, 587–612 (2011).
22. Bhandari, V., Wong, K. S., Zhou, J. L., Mabanglo, M. F., Batey, R. A. & Houry, W. A. The Role of ClpP Protease in Bacterial Pathogenesis and Human Diseases. *ACS Chem. Biol.* 13 (6), 1413–1425

(2018).

23. Brotz-Oesterhelt, H. & Sass, P. Bacterial caseinolytic proteases as novel targets for antibacterial treatment. *Int. J. Med. Microbiol.* 304 (1), 23-30 (2014).
24. Baker, T. A. & Sauer, R. T. ClpXP, an ATP-powered unfolding and protein-degradation machine ☆. *BBA - Mol. Cell Res.* 1823 (1), 15–28 (2012).
25. Flynn, J. M., Neher, S. B., Kim, Y.-I., Sauer, R. T. & Baker, T. A. Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell.* 11, 671–683 (2003).
26. Neher, S. B., Villen, J., Oakes, E. C., Bakalarski, C. E., Sauer, R. T., Gygi, S. P. & Baker, T. A. Proteomic Profiling of ClpXP Substrates after DNA Damage Reveals Extensive Instability within SOS Regulon. *Mol. Cell.* 22, 193–204 (2006).
27. Gottesman, S., Roche, E., Zhou, Y. & Sauer, R. T. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev.* 12 (9):1338-47 (1998).
28. Msadek, T., Dartois, V., Kunst, F., Herbaud, M. L., Denizot, F. & Rapoport, G. ClpP of *Bacillus subtilis* is required for competence development, motility, degradative enzyme synthesis, growth at high temperature and sporulation. *Mol. Microbiol.* 27 (5), 899–914 (1998).
29. Pummi, T., Leskelä, S., Wahlström, E., Gerth, U., Tjalsma, H., Hecker, M., *et al.* ClpXP protease regulates the signal peptide cleavage of secretory preproteins in *Bacillus subtilis* with a mechanism distinct from that of the *ecs* ABC transporter. *J. Bacteriol.* 184 (4), 1010–1018 (2002).
30. Jenal, U. & Fuchs, T. An essential protease involved in bacterial cell-cycle control. *EMBO J.* 17 (19), 5658-69 (1998).
31. Sassetti, M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* 48 (1), 77–84 (2003).
32. Frees, D., Sørensen, K. & Ingmer, H. Global Virulence Regulation in *Staphylococcus aureus*: Pinpointing the Roles of ClpP and ClpX in the *sar / agr* Regulatory Network. *Society.* 73 (12), 8100–8108 (2005).
33. Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. & Berche, P. The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol. Microbiol.* 35 (6), 1286-1294 (2000).

34. Gailliot, O., Bregenholt, S., Jaubert, F., Di Santo, J. P. & Berche, P. Stress-induced ClpP serine protease of *Listeria monocytogenes* is essential for induction of listeriolysin O-dependent protective immunity. *Infect Immun.* 69 (8), 4938-4943 (2001).
35. Kwon, H. Y., Ogunniyi, A. D., Choi, M. H., Pyo, S. N., Rhee, D. K. & Paton, J. C. The ClpP protease of *Streptococcus pneumoniae* modulates virulence gene expression and protects against fatal pneumococcal challenge. *Infect. Immun.* 72 (10), 5646-5653 (2004).
36. Wong, P. & Houry, W. A. Chaperone networks in bacteria: Analysis of protein homeostasis in minimal cells. in *Journal of Structural Biology.* 146 (1-2), 79-89 (2004).
37. Stanne, T. M., Pojidaeva, E., Andersson, F. I. & Clarke, A. K. Distinctive types of ATP-dependent Clp proteases in cyanobacteria. *J. Biol. Chem.* 282 (19), 14394-402 (2007).
38. Schelin, J., Lindmark, F. & Clarke, A. K. The clpP multigene family for the ATP- dependent Clp protease in the cyanobacterium *Synechococcus*. *Amino Acids.* 148 (7), 2255-65 (2002).
39. Szczepanowska, K., Maiti, P., Kukat, A., Hofsetz, E., Nolte, H., Senft, K. *et al.* CLPP coordinates mitoribosomal assembly through the regulation of ERAL1 levels. *EMBO J.* 35 (23), 2566–2583 (2016).
40. Kang, S. G., Maurizi, M. R., Thompson, M., Mueser, T. & Ahvazi, B. Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *J. Struct. Biol.* 148 (3), 338-52 (2004).
41. Corydon, T. J., Wilsbech, M., Jespersgaard, C., Andresen, B. S., Borglum, A. D., Pedersen, S., *et al.* Human and mouse mitochondrial orthologs of bacterial ClpX. *Mamm. Genome.* 11 (10), 899–905 (2000).
42. Jennings, L. D., Lun, D. S., Médard, M. & Licht, S. ClpP hydrolyzes a protein substrate processively in the absence of the ClpA ATPase: Mechanistic studies of ATP-independent proteolysis. *Biochemistry.* 47 (44), 11536–11546 (2008).
43. Szyk, A. & Maurizi, M. R. Crystal structure at 1.9Å of *E. coli* ClpP with a peptide covalently bound at the active site. *J. Struct. Biol.* 156 (1), 165-174 (2006).
44. Sprangers, R., Gribun, A., Hwang, P. M., Houry, W. A. & Kay, L. E. Quantitative NMR spectroscopy of supramolecular complexes: Dynamic side pores in ClpP are important for product release. *Pnas.* 102 (46), 16678–16683 (2005).

-
45. Geiger, S. R., Böttcher, T., Sieber, S. A. & Cramer, P. A conformational switch underlies ClpP protease function. *Angew. Chemie - Int. Ed.* 50 (25), 5749-52 (2011).
 46. Wang, J., Hartling, J. A. & Flanagan, J. M. Crystal Structure Determination of Escherichia coli ClpP Starting from an EM-Derived Mask. *J. Struct. Biol.* 124 (5), 151–163 (1998).
 47. Kim, D. Y. & Kim, K. K. The Structural Basis for the Activation and Peptide Recognition of Bacterial ClpP. *J. Mol. Biol.* 379 (4), 760-71 (2008).
 48. Kimber, M. S., Yu, A. Y., Borg, M., Leung, E., Chan, H. S. & Houry, W. A. Structural and Theoretical Studies Indicate that the Cylindrical Protease ClpP Samples Extended and Compact Conformations. *Structure.* 18 (7), 798-808 (2010).
 49. Lee, B. G., Kim, M. K. & Song, H. K. Structural insights into the conformational diversity of ClpP from Bacillus subtilis. *Mol. Cells.* 32 (6), 589–595 (2011).
 50. Zhang, J., Ye, F., Lan, L., Jiang, H., Luo, C. & Yang, C. G. Structural switching of Staphylococcus aureus Clp protease: A key to understanding protease dynamics. *J. Biol. Chem.* 286 (43), 37590-601 (2011).
 51. Olivares, A. O., Baker, T. A. & Sauer, R. T. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat. Rev. Microbiol.* 14 (1), 33-44 (2016).
 52. Huang, J., Wang, X., Cao, Q., Feng, F., Xu, X. & Cai, X. ClpP participates in stress tolerance and negatively regulates biofilm formation in Haemophilus parasuis. *Vet. Microbiol.* 182, 141-9 (2016).
 53. Park, C., Kim, E. H., Choi, S. Y., Tran, T. D., Kim, I. H., Kim, S. N., *et al.* Virulence Attenuation of Streptococcus pneumoniae clpP mutant by sensitivity to oxidative stress in macrophages via a NO-mediated pathway. *J Microbiol.* 48 (2), 229–235 (2010).
 54. Robinson, J. L. & Brynildsen, M. P. An ensemble-guided approach identifies ClpP as a major regulator of transcript levels in nitric oxide-stressed Escherichia coli. *Metab. Eng.* 31, 22–34 (2015).
 55. Hinzen, B., Raddatz, S., Paulsen, H., Lampe, T., Schumacher, A., Häbich, D., *et al.* Medicinal chemistry optimization of acyldepsipeptides of the enopeptin class antibiotics. *ChemMedChem.* 1 (7), 689-93 (2006).
 56. Goodreid, J. D., Janetzko, J., Santa Maria, J. P. Jr., Wong, K. S., Leung, E., Eger, B. T., *et al.* Development and characterization of potent cyclic acyldepsipeptide analogues with increased antimicrobial activity. *J. Med. Chem.* 59 (2), 624-646 (2016).

-
57. Olivares, A. O., Baker, T. A. & Sauer, R. T. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat. Rev. Microbiol.* 14 (1), 33-44 (2016).
58. McGillivray, S. M., Tran, D. N., Ramadoss, N. S., Alumasa, J. N., Okumura, C. Y., Sakoulas, G., *et al.* Pharmacological inhibition of the ClpXP protease increases bacterial susceptibility to host cathelicidin antimicrobial peptides and cell envelope-active antibiotics. *Antimicrob. Agents Chemother.* 56 (4), 1854-61 (2012).
59. Gersch, M., Gut, F., Korotkov, V. S., Lehmann, J., Böttcher, T., Rusch, M., *et al.* The mechanism of caseinolytic protease (ClpP) inhibition. *Angew. Chemie - Int. Ed.* 52 (10), 3009-14 (2013).
60. Böttcher T. & Sieber, S. A. β -Lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of *Staphylococcus aureus*. *J. Am. Chem. Soc.* 130 (44), 14400-14401 (2008).
61. Böttcher, T. & Sieber, S. A. β -Lactones decrease the Intracellular virulence of *Listeria monocytogenes* in macrophages. *Chemmedchem.* 4 (8), 1260-1263 (2009).
62. Hackl, M. W., Lakemeyer, M., Dahmen, M., Glaser, M., Pahl, A., Lorenz-Baath, K., *et al.* Phenyl esters are potent inhibitors of caseinolytic protease P and reveal a stereogenic switch for deoligomerization. *J. Am. Chem. Soc.* 137 (26), 8475-8483 (2015).
63. Pahl, A., Lakemeyer, M., Vielberg, M. T., Hackl, M. W., Vomacka, J., Korotkov, V. S., *et al.* Reversible Inhibitors Arrest ClpP in a defined conformational state that can be revoked by ClpX association. *Angew. Chem. Int. Ed. Engl.* 54 (52), 15892-15899 (2015).
64. Schmitt, E. K., Riwanto, M., Sambandamurthy, V., Roggo, S., Miault, C., Zwingelstein, C., *et al.* The Natural Product Cyclomarin Kills *Mycobacterium Tuberculosis* by Targeting the ClpC1 Subunit of the Caseinolytic Protease. *Angew. Chemie Int. Ed.* 50 (26), 5889–5891 (2011).
65. Akopian, T., Kandror, O., Tsu, C., Lai, J. H., Wu, W. G., Liu, Y. X., *et al.* Cleavage specificity of *Mycobacterium tuberculosis* ClpP1P2 protease and identification of novel peptide substrates and boronate inhibitors with anti-bacterial activity. *J. Biol. Chem.* 290 (17), 11008-11020 (2015).
66. Moreira, W., Ngan, G. J., Low, J. L., Poulsen, A., Chia, B. C., Ang, M. J., *et al.* Target mechanism-based whole-cell screening identifies bortezomib as an inhibitor of caseinolytic protease in mycobacteria. *mBio.* 6 (3), e00253-15 (2015).
67. Gersch, M., Kolb, R., Alte, F., Groll, M. & Sieber, S. A. Disruption of Oligomerization and Dehydroalanine Formation as Mechanisms for ClpP Protease Inhibition. *J. Am. Chem. Soc.* 136 (4),

- 1360–1366 (2014).
68. Zhu, K., Borrelli, K. W., Greenwood, J. R., Day, T., Abel, R., Farid, R. S., *et al.* Docking covalent inhibitors: A parameter free approach to pose prediction and scoring. *J. Chem. Inf. Model.* *54* (7), 1932–40 (2014).
69. Robertson, J. G. Enzymes as a special class of therapeutic target: clinical drugs and modes of action. *Current Opinion in Structural Biology.* *17*(6), 674–9 (2007).
70. Kumalo, H. M., Bhakat, S. & Soliman, M. E. S. Theory and Applications of Covalent Docking in Drug Discovery: Merits and Pitfalls. *Molecules.* *20* (2), 1984–2000 (2015).
71. Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat. Rev.* *10* (4), 307–17 (2011).
72. Alfonso, L., Ai, G., Spitale, R. C. & Bhat, G. J. Molecular targets of aspirin and cancer prevention. *Br. J. Cancer* *111*, 61–67 (2014).
73. Baillie, T. A. Covalent Drugs Targeted Covalent Inhibitors for Drug Design. *Angew.Chem.Int.* *551* (43), 3408–13421 (2016).
74. Smith, A. J. T., Zhang, X., Leach, A. G. & Houk, K. N. Beyond picomolar affinities: Quantitative aspects of noncovalent and covalent binding of drugs to proteins. *Journal of Medicinal Chemistry.* *52* (2), 225–33 (2009).
75. González-Bello, C. Designing Irreversible Inhibitors--Worth the Effort? *ChemMedChem.* *11* (1), 22–30 (2015).
76. Johnson, D. S., Weerapana, E. & Cravatt, B. F. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors Brief history & examples of covalent inhibitors. *Futur. Med Chem.* *2* (6), 949–964 (2010).
77. Lanning, B. R., Whitby, L. R., Dix, M. M., Douhan, J., Gilbert, A. M., Hett, E. C., *et al.* A roadmap to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat Chem Biol.* *10* (9), 760–767 (2014).
78. Lammert, C., Einarsson, S., Saha, C., Niklasson, A., Bjornsson, E. & Chalasani, N. Relationship between daily dose of oral medications and idiosyncratic drug-induced liver injury: Search for signals. *Hepatology.* *47* (6), 2003–9 (2008).
79. Kirstein, J., Hoffmann, J., Lilie, H., Schmidt, R., Rubsamen-Waigmann, H., Brotz-Oesterhelt, *et al.* The

- antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO Molecular Medicine*. 1 (1), 37-49 (2009).
80. Brotz-Oesterhelt, H., Beyer, D., Kroll, H.-P., Endermann, R., Ladel, C., Schroeder, W., *et al.* Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat. Med.* 11 (12), 1082–1087 (2005).
81. Conlon, B. P., Nakayasu, E. S., Fleck, L. E., LaFleur, M. D., Isabella, V. M., Coleman, K., *et al.* Activated ClpP kills persisters and eradicates a chronic biofilm infection. *Nature*. 503 (7476), 365-70 (2013).
82. Honsa, E. S., Cooper, V.S., Mhaisien, M. N., Frank, M., Shaker, J., Iverson, A., *et al.* RelA mutant *Enterococcus faecium* with multiantibiotic tolerance arising in an immunocompromised host. *mBio*. 8 (1), e02124-16 (2017).
83. Malik, I. T. & Brotz-Oesterhelt, H. Conformational control of the bacterial Clp protease by natural product antibiotics. *Nat. Prod. Rep.* 34 (7),815, (2017).
84. Dzidic, S., Suskovic, J. & Kos, B. Antibiotic Resistance Mechanisms in Bacteria: Biochemical and Genetic Aspects. *Food Technol. Biotechnol.* 46 (1), 11–21 (2008).
85. Fernández, L., Breidenstein, E. B. M., Song, D. & Hancock, R. E. W. Role of Intracellular Proteases in the Antibiotic Resistance , Motility , and Biofilm Formation of *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother.* 56(2), 1128–1132 (2012).
86. Wright, G. D. Q&A: Antibiotic resistance: Where does it come from and what can we do about it? *BMC Biol.* 8, 123 (2010).
87. Munita, J. M., Arias, C. A., Unit, A. R. & Santiago, A. De. Mechanisms of Antibiotic Resistance. *Mech. Antibiob. Resist. HHS Public Access.* 4 (2), 1–37 (2016).
88. Santajit, S. & Indrawattana, N. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. *Biomed Res. Int.* 2016, 2475067, (2016).
89. Bauer, R.A. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. *Drug Discov Today.* 20 (9), 1061-73 (2015).

Chapter 2

α -Amino diphenyl phosphonates as novel inhibitors of *Escherichia coli* ClpP protease

The content of this chapter is partially based on the paper “ **α -Amino diphenyl phosphonates as novel inhibitors of *Escherichia coli* ClpP protease**”

Reprinted with permission from α -Amino Diphenyl Phosphonates as Novel Inhibitors of *Escherichia coli* ClpP Protease

Carlos Moreno-Cinos, Elisa Sassetti, Irene G. Salado, Gesa Witt, Siham Benramdane, Laura Reinhardt, Cristina D. Cruz, Jurgen Joossens, Pieter Van der Veken, Heike Brötz-Oesterhelt, Päivi Tammela, Mathias Winterhalter, Philip Gribbon, Björn Windshügel, and Koen Augustyns

Journal of Medicinal Chemistry **2019** 62 (2), 774-797

DOI: 10.1021/acs.jmedchem.8b01466

Copyright 2019 American Chemical Society

Carlos Moreno-Cinos,^{1,§} Elisa Sassetti,^{2,5,§} Irene G. Salado,¹ Gesa Witt,² Siham Benramdane,¹ Laura Reinhardt,³ Cristina Durante Cruz,⁴ Jurgen Joossens,¹ Pieter Van der Veken,¹ Heike Brötz-Oesterhelt,³ Päivi Tammela,⁴ Mathias Winterhalter,⁵ Philip Gribbon,² Björn Windshügel,² and Koen Augustyns¹.

¹ *Laboratory of Medicinal Chemistry, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium.*

² *Fraunhofer Institute for Molecular Biology and Applied Ecology, ScreeningPort, Schnackenburgallee 114, 22525, Hamburg, Germany.*

³ *Interfaculty Institute for Microbiology and Infection Medicine, University of Tübingen, Auf der Morgenstelle 28, 72076, Tübingen, Germany.*

⁴ *Drug Research Program, Division of Pharmaceutical Biosciences, University of Helsinki, Viikinkaari 5E, FI-00014 Helsinki, Finland.*

⁵ *Department of Life Sciences and Chemistry, Jacobs University Bremen gGmbH, Campus Ring 1, 28759 Bremen, Germany.*

[§] *Shared first author.*

Contribution to the work. I have designed, performed (except two repetitions of a bacterial growth assay, conducted by Dr. Cristina Durante Cruz) and analyzed personally all the biological experiments and the computational tests included in this part of the work. I also gave a substantial contribution in writing and reviewing the manuscript. The rest of the work has been performed by the co-author of the paper.

2.1 Abstract

Increased Gram-negative bacteria resistance to antibiotics is becoming a global problem and new classes of antibiotics with novel mechanisms of action are required. The caseinolytic protease subunit P (ClpP) is a serine protease conserved among bacteria that is considered as an interesting drug target. ClpP function is involved in protein turnover and homeostasis, stress-response and virulence among other processes. The focus of this study was to identify new inhibitors of *Escherichia coli* ClpP and to understand their mode of action. A focused library of serine protease inhibitors based on diaryl phosphonate warheads was tested for ClpP inhibition and a chemical exploration around the hit compounds was conducted. Altogether 14 new potent inhibitors of *E. coli* ClpP were identified. Compounds 85 and 92 emerged as most interesting compounds from this study due to their potency and, respectively, to its moderate but consistent antibacterial properties as well as the favourable cytotoxicity profile.

2.2 Introduction

Antibiotic resistance is a major global problem in both developed and developing countries.¹ The selection pressures on microorganisms when in contact with antibacterial agents underlies the emergence of resistance,² and the efficacy of first and second line antibiotics is decreasing at an alarming rate.³ The importance of antimicrobial drug discovery was underlined by the World Health Organization's (WHO) first global report on antibiotic resistance which attributed 25,000 deaths in Europe and 2 million worldwide per year to bacterial infections.⁴ Of particular concern are Gram-negative multidrug-resistant bacteria (MDR) which are becoming more prevalent.⁵ Among the antibiotic drugs launched since the year 2000, only five new classes were introduced and only one was directed against Gram-negative bacteria in combination with β -lactams.⁶ To avoid key resistance mechanisms to pre-existing antibiotics, drug discovery research has focussed on addressing alternative targets with novel mechanisms of antibacterial action.⁷

The antibacterial drug target caseinolytic protease proteolytic subunit (ClpP) is a widely conserved protein which is present in bacteria, in many eukaryotes (including humans, localised in mitochondria), but is absent in archaea and mollicutes.⁸⁻⁹ ClpP, a chymotrypsin-like serine protease,¹⁰ is thought to play an important role in determining virulence and stress response by modulating virulence factor expression in several bacteria including *Staphylococcus aureus*, *Listeria monocytogenes* and *Streptococcus pneumoniae*.¹¹⁻¹⁴ ClpP degrades mistranslated, misfolded or aggregated proteins, arising as a result of stress factors (*e.g.* heat stress and antibiotics).⁸ In *Listeria monocytogenes* ClpP was found to be essential for bacterial survival in macrophages.¹² In *Streptococcus pneumoniae* the levels of ClpP were demonstrated to be correlated with nitric oxide stress.¹⁴

ClpP proteases in Gram-positive bacteria have been more thoroughly studied as drug targets, but also advances in Gram-negative ClpPs were recently reported. Robinson *et al.*¹⁶ identified ClpP as potential target for antivirulence therapies by showing differences between growth curves of wild-type and *clpP*-defective *E. coli* under nitric oxide stress conditions. It has also been demonstrated that in *E. coli* is responsible for the cleavage of proteins involved in metabolism, transcription factors, as well as in oxidative stress response and starvation.¹⁷ Furthermore, *clpP*-deficient *Legionella pneumophila* showed impaired virulence and reduced translocation of effector proteins in the studies from Zhao *et al.*¹⁸ and ClpX and ClpP2 were identified by Qiu *et al.*¹⁹ as part of the proteolytic network of the exopolysaccharide alginate biosynthesis in *Pseudomonas aeruginosa*, a marker for the onset of chronic lung infection in cystic fibrosis.

ClpP is a tetradecamer with a cylindrical shape. The 14 subunits are arranged in two heptameric rings and a central chamber which contains the active sites of each subunit.⁸ Each active site comprises the canonical Ser-His-Asp catalytic triad (for most bacteria).²⁰ The peptidase activity, a characteristic of a chymotrypsin-like serine protease,²¹ typically results in peptides of 7-8 residues length,²² with cuts occurring after non-polar residues.²³ ClpP proteolytic activity requires the presence of specific ATPases (ClpX and ClpA in the case of *E. coli*),²³ of the AAA+ enzyme superfamily, whose function is to recognize, unfold and then transfer the substrates into the chamber, thus forming the Clp complex together with ClpP.²⁴ The interface between ClpP and the AAA+ partners has been investigated as a drug targeting site and several antibacterial peptides were identified, which activate and deregulate ClpP.²⁴⁻²⁵ These acyldepsipeptides (ADEPs) prevent ATPases binding to the heptameric rings of ClpP, resulting in uncontrolled proteolysis of essential bacterial proteins and eventually in bacterial cell death.^{8, 26}

A promising approach to target the virulence-related functions of ClpP is to develop enzyme inhibitors used in combination with existing antibiotics. The pioneering efforts of Böttcher and Sieber to target ClpP led to the development of a series of β -lactone inhibitors (among them D3, Figure 1) for *S. aureus* ClpP.²⁷ These inhibitors bind covalently to the catalytic serine, leading to irreversible inhibition of proteolytic activity. Further characterization proved their ability to reduce bacterial virulence expression not only in *S. aureus* but also in *L. monocytogenes*.²⁸⁻²⁹ The potency of this inhibitor was improved 3- to 5-fold with the optimized β -lactone U1 (Figure 1).³⁰ However, the reduced plasma stability of these compounds, due to the fast hydrolysis of the cyclic ester, impeded further clinical development.³¹

Thereafter, a new class of potent ClpP inhibitors with better plasma stability was discovered by the Sieber group.³¹ The phenyl esters (AV170, Figure 1) irreversibly inhibited *S. aureus* ClpP, and triggered deoligomerization of the ClpP tetradecamer into inactive heptamers. Their higher potency, inhibition kinetics and plasma lifetime, compared to the β -lactone series, were countered by their lower anti-virulence activity. Furthermore, attempts to further improve their acyl-enzyme complex stability unfortunately led to a loss of

ClpP reactivity.³¹ A non-covalent inhibitor against *S. aureus* ClpP has been also identified in high-throughput screening (HTS) campaign.³² The inhibitor (AV145, Figure 1) bound to the handle region near the active site, locking *S. aureus* ClpP in a novel and inactive conformation. However, binding of ClpX to ClpP revoked the inhibitory effect of AV145 and its analogs in bacteria.³²

Boron derived compounds have also shown evidence of successfully inhibiting ClpP in *Mycobacterium tuberculosis* as demonstrated for bortezomib by Moreira *et al.* or the substrate-based peptide boronate inhibitors by Akopian *et al.* (Figure 1).^{33,34} Nevertheless, proteasome inhibition, short half-life, poor pharmacokinetics and its high cost limited the direct use of bortezomib, the most potent *in cellulo* of previously described compounds at *M. tuberculosis* treatment.³³

Recently, also pyrimidines have been shown to inhibit ClpP.³⁵ Compounds (P33, Figure 1) targeting *Plasmodium falciparum* ClpP achieved inhibition of growth and segregation of the apicoplast during the cell cycle, leading to parasite death.

Although ClpPs have been investigated in several organisms, inhibition of Gram-negative bacteria ClpP remains untapped and the chloromethyl ketone (Z-LY-CMK, Figure 1) co-crystallized by Szyk and Maurizi was the only reported inhibitor for *E. coli* ClpP reported so far.³⁶

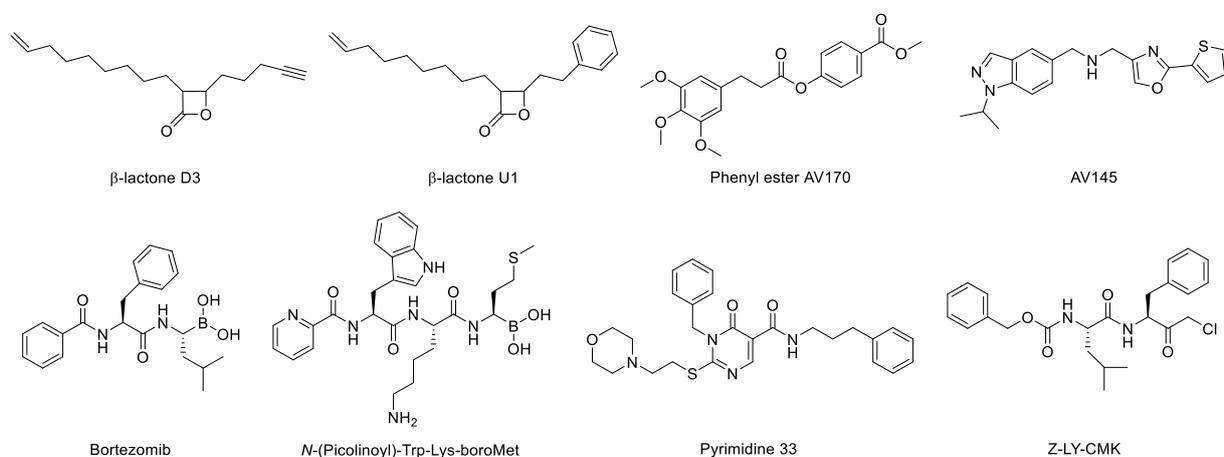


Figure 1. Examples of ClpP reported inhibitors.

Despite the demonstration of the potential of irreversible inhibitors on different ClpPs, inhibitors with a classical α -amino diaryl phosphonate warhead remained unexplored.³⁷ Thus far, several diaryl phosphonate compounds have been identified as potent, irreversible serine protease inhibitors. Some illustrative examples are a urokinase plasminogen activator (uPA) inhibitor reported by Joossens *et al.*,³⁸⁻³⁹ a dipeptidyl peptidase 8 (DPP8) inhibitor by Van der Veken *et al.*,⁴⁰ an elastase inhibitor by Winiarski *et al.*,⁴¹ a subtilisin inhibitor by Pietruszewicz *et al.*,⁴² and the GluC and SplA inhibitors by Burchacka *et al.*⁴³⁻⁴⁴

The mode of action for this class of inhibitors (Figure 2) involves a nucleophilic attack by the hydroxyl of the active site serine on the electrophilic phosphorus atom, leading to the formation of a phosphonate ester. The initial enzyme-inhibitor complex is unstable. Therefore, hydrolysis of the aryl ester (with a half-life ranging from few hours to few days) leads to the formation of the “aged complex”.

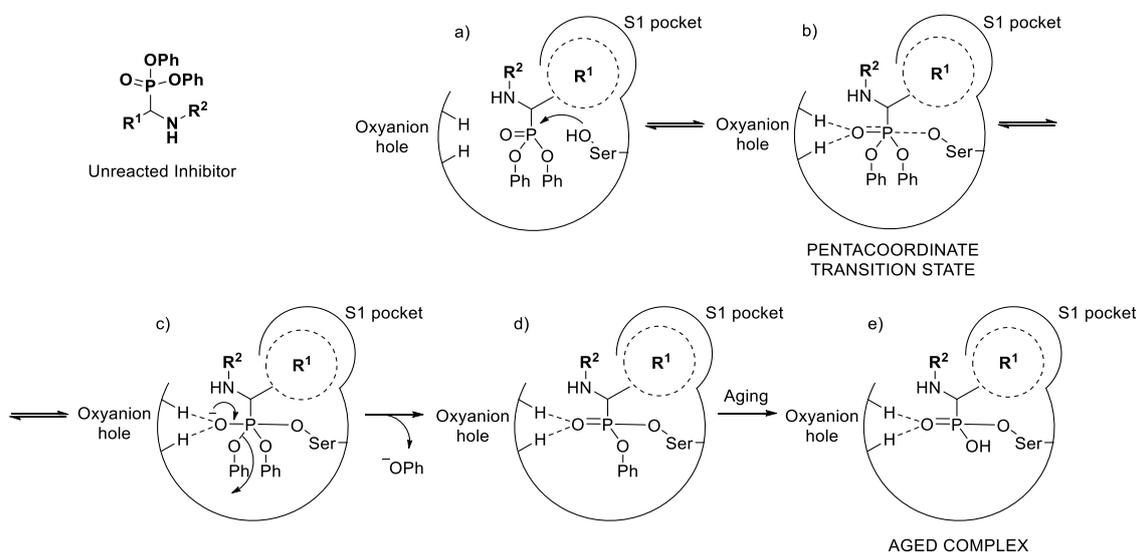


Figure 2. Binding mechanism of diphenyl phosphonates with serine proteases. a) The unreacted inhibitor enters the active site, with the R^1 moiety filling the S1 pocket while the phosphonate sits at a reachable distance from the oxyanion hole and the catalytic serine. b) Nucleophilic attack of the serine to the phosphonate facilitated by the hydrogen bonds of this group with the oxyanion hole residues to form the pentacoordinate transition state. c) Formalized bonds between the serine oxygen and phosphorus of the phosphonate lead to a negative charge on the oxygen that, when recovering the tetrahedral geometry, leads to the release of the phenol group. d) Stabilized configuration after covalent bonding between ligand and serine protease. e) Slow hydrolysis of the remaining phenolate leads to the formation of the aged complex.

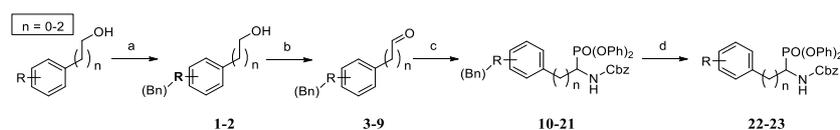
The aim of this work was to identify new classes of compounds as inhibitors of ClpP activity and investigate their mechanisms of action. We describe a series of α -amino diphenyl phosphonate esters as the first potent inhibitors of *E. coli* ClpP, using this species as a model organism for Gram-negative bacteria, encouraged by the availability of a crystal structure and by the previous studies where a *clpP*-defective strain showed a decreased growth under nitric oxide stress conditions.^{16, 36}

2.3 Results

2.3.1 Chemical explorations and enzymatic activity screening

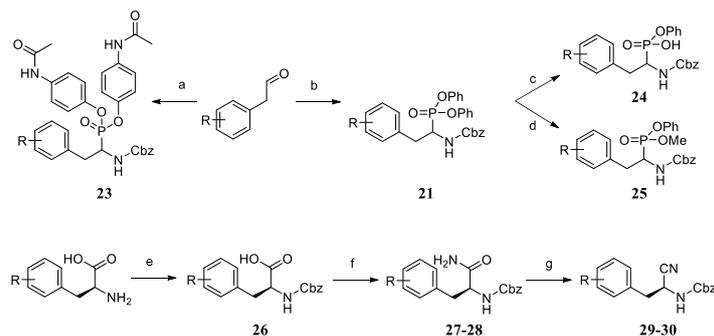
The existing diarylphosphonate library of the Medicinal Chemistry group of the University of Antwerp (UAMC) was highly enriched in hydrophilic and polar residues in R¹ position, since it was mainly focused on targeting trypsin-like serine proteases. Based on the specificity of chymotrypsin-like serine proteases for lipophilic residues in the S1 pocket, a library of hydrophobic moieties in R¹ was designed. Some analogs of the previously described inhibitors (Z-LY-CMK and Lys-boroMet in Figure 1) were also included, together with variations on the warhead (diversity of arylphosphonates and nitriles).^{34, 36}

Synthesis of the analogues with -Cbz in R² position and a diphenyl phosphonate as warhead (10-23) was carried out following the general synthesis described in Scheme 1, where protection of the hydroxyl groups on some of the aromatic rings was carried out in order to improve the yield of the following steps: Dess-Martin oxidation⁴⁵ and a modified alternative of the Birum-Oleksyszyn reaction previously reported by Van der Veken *et al.*⁴⁶ Those protected compounds, were finally debenzylated following the conditions of Okano *et al.*⁴⁷



Scheme 1. Reagents and conditions. a) K₂CO₃, BnBr, DMF, rt, 4 h. b) Dess-Martin periodinane, DCM, 0-25 °C, 2 h. c) CbzNH₂, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h. d) Pentamethylbenzene, BCl₃, DCM, - 78 °C, 15 min.

Synthesis of the compounds with modifications on the phosphonate warhead or its substitution by a nitrile (23-30) were undertaken as described in Scheme 2, while dipeptidic diphenyl phosphonates (32-41) were obtained by prior cleavage of -Cbz and subsequent peptidic coupling. These protocols can be found in the experimental section.



Scheme 2. Reagents and conditions. a) CbzNH₂, tris(4-acetamidophenyl) phosphite, Cu(OTf)₂, DCM, rt, 16 h. b) CbzNH₂, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h. c) KOH, H₂O:dioxane (1:1), rt, 16 h. d) NH₃, NH₄Cl, MeOH, rt, 72 h. e) CbzCl, NaOH, H₂O, 0-25 °C, 2 h. f) Isobutylchloroformate, *N*-methylmorpholine, NH₃, DCM, 0-25 °C, 16 h. g) Burgess reagent, DCM, rt, 16 h.

The compounds 10-41 were evaluated for ClpP inhibition together with a subset of diaryl phosphonates from the UAMC library (42-74), selected in order to expand the variety of R¹ and R² residues (Table 1). ClpP inhibition was assessed by a high-throughput screen in 384-well format using a fluorescence assay with Suc-LY-AMC as fluorogenic substrate. The compounds were screened at 200 μM concentration. Compounds were considered as active if the percentage of inhibition (compared to a control without compounds) was higher or equal to 75 % (or ≤ 25% remaining activity).

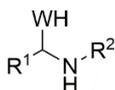


Figure 3. Generic compound structure.

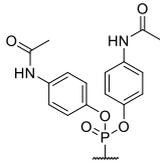
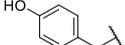
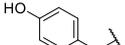
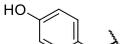
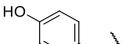
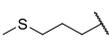
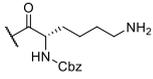
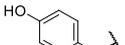
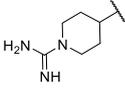
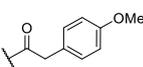
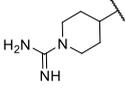
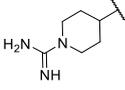
Table 1. Enzymatic inhibition of the apolar exploration and first library screening.

	%I (200 μM)	IC ₅₀ (μM)

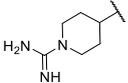
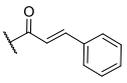
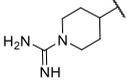
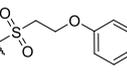
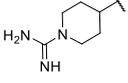
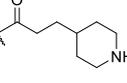
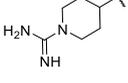
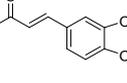
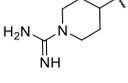
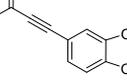
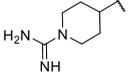
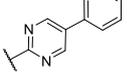
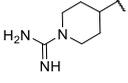
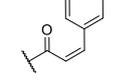
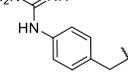
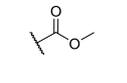
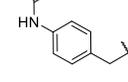
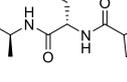
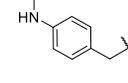
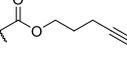
Chapter 2

Z-LY-CMK				100	14.4
10		-Cbz	-PO(OPh) ₂	7	ND
11		-Cbz	-PO(OPh) ₂	7	ND
12		-Cbz	-PO(OPh) ₂	6	ND
13		-Cbz	-PO(OPh) ₂	6	ND
14		-Cbz	-PO(OPh) ₂	<1	ND
15		-Cbz	-PO(OPh) ₂	13	ND
16		-Cbz	-PO(OPh) ₂	100	14.2 ± 1.3
17		-Cbz	PO(OPh) ₂	9	ND
18		-Cbz	-PO(OPh) ₂	<1	ND
21		-Cbz	-PO(OPh) ₂	<1	ND
22		-Cbz	-PO(OPh) ₂	3	ND

Chapter 2

23		-Cbz		9	ND
24		-Cbz		11	ND
25		-Cbz		14	ND
29		-Cbz	-CN	<1	ND
30		-Cbz	-CN	<1	ND
32			-PO(OPh) ₂	<1	ND
36			-PO(OPh) ₂	4	ND
41			-CN	13	ND
42			-PO(OPh) ₂	13	ND
43			-PO(OPh) ₂	13	ND
44			-PO(OPh) ₂	24	ND

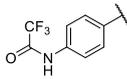
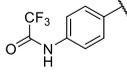
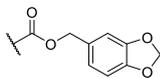
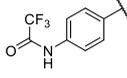
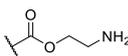
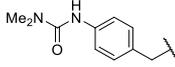
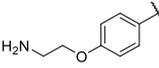
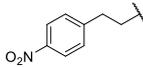
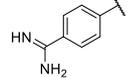
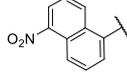
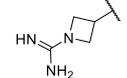
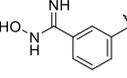
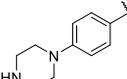
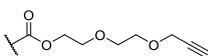
Chapter 2

45			-PO(OPh) ₂	18	ND
46			-PO(OPh) ₂	27	ND
47			-PO(OPh) ₂	6	ND
48			-PO(OPh) ₂	4	ND
49			-PO(OPh) ₂	24	ND
50			-PO(OPh) ₂	12	ND
51			-PO(OPh) ₂	11	ND
52			-PO(OPh) ₂	27	ND
53			-PO(OPh) ₂	90	49.5 ± 0.5
54			-PO(OPh) ₂	23	ND

Chapter 2

55			-PO(OPh) ₂	68	ND
56			-PO(OPh) ₂	25	ND
57			-PO(OPh) ₂	55	ND
58		-Cbz	-PO(OPh) ₂	96	39.8 ± 2.9
59			-PO(OPh) ₂	1	ND
60			-PO(OPh) ₂	<1	ND
61			-PO(OPh) ₂	28	ND
62			-PO(OPh) ₂	27	ND
63			-PO(OPh) ₂	29	ND
64			-PO(OPh) ₂	48	ND

Chapter 2

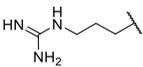
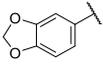
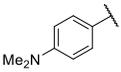
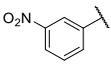
65			-PO(OPh) ₂	64	ND
66			-PO(OPh) ₂	93	8.2 ± 0.8
67			-PO(OPh) ₂	73	ND
68		-Cbz	-PO(OPh) ₂	9	ND
69		-Cbz	-PO(OPh) ₂	15	ND
70		-Cbz	-PO(OPh) ₂	100	13.1 ± 1.2
71			-PO(OPh) ₂	93	48.1 ± 1.7
72			-PO(OPh) ₂	17	ND
73			-PO(OPh) ₂	12	ND
74			-PO(OPh) ₂	23	ND
75			-PO(OPh) ₂	25	ND

Six compounds emerged as active in the primary screen (16, 53, 58, 66, 70 and 71). Dose-response experiments confirmed all initial hits and the IC₅₀ values ranged between 8.2 and 49.5 μM (Table 1).

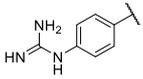
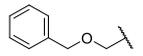
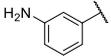
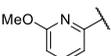
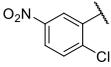
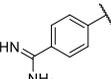
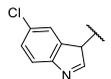
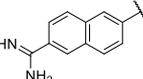
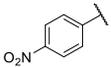
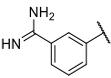
The biochemical tests revealed that the S1 pocket showed a preference for hydrophilic moieties, while 16 represents the only active compound with a lipophilic R¹ moiety. Regarding the R² substitution, a variety of simple carbamates were tolerated, -Cbz being the most common and chemically accessible. However, methyl carbamates and benzodioxol carbamates were also taken into account for future investigations.

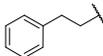
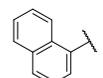
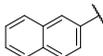
After learning that the S1 pocket accepted a wider range of side chains, every remaining compound from the library with -Cbz in R² position (76-95) was submitted to a second round of experimental testing (Table 2).

Table 2. Enzymatic inhibition of second library screening.

				%I (200 μM)	IC ₅₀ (μM)
76		-Cbz	-PO(OPh) ₂	15	ND
77		-Cbz	-PO(OPh) ₂	6	ND
78		-Cbz	-PO(OPh) ₂	100	0.6 ± 0.1
79		-Cbz	-PO(OPh) ₂	27	ND
80		-Cbz	-PO(OPh) ₂	6	ND

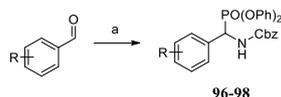
Chapter 2

81		-Cbz	-PO(OPh) ₂	3	ND
82		-Cbz	-PO(OPh) ₂	4	ND
83		-Cbz	PO(OPh) ₂	29	ND
84		-Cbz	-PO(OPh) ₂	4	ND
85		-Cbz	-PO(OPh) ₂	100	0.5 ± 0.0
86		-Cbz	-PO(OPh) ₂	100	38.0 ± 2.4
87		-Cbz	-PO(OPh) ₂	6	ND
88		-Cbz	-PO(OPh) ₂	30	ND
89		-Cbz	-PO(OPh) ₂	88	100.5 ± 8.0
90		-Cbz	-PO(OPh) ₂	81	79.7 ± 7.2
91		-Cbz	-PO(OPh) ₂	71	ND
92		-Cbz	-PO(OPh) ₂	100	0.5 ± 0.0

93		-Cbz	PO(OPh) ₂	<1	ND
94		-Cbz	-PO(OPh) ₂	6	ND
95		-Cbz	-PO(OPh) ₂	51	ND

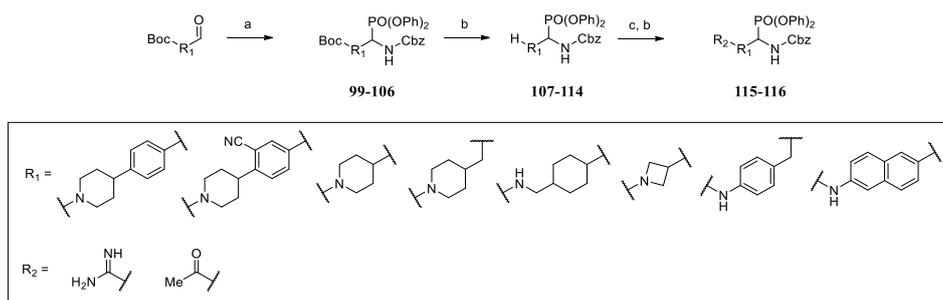
The biochemical tests resulted in the identification of six additional ClpP inhibitors (78, 85, 86, 89, 90 and 92), of which three inhibited the enzyme with a sub-micromolar IC₅₀ value. From this and the previous screening, we concluded that hydrophilicity in the S1 pocket is preferred, and we, therefore, continued with a chemical exploration of polar groups in R¹ together with further modifications around three selected R¹ residues.

First, based on the polarity of the most active compounds identified so far, the scope of hydrophilic moieties for R¹ was enlarged, leaving the rest of the structure unchanged (96-127). Some of these compounds (96-98) were directly obtained from the commercial aldehydes after a Birum-Oleksyszyn reaction as stated in Scheme 3.



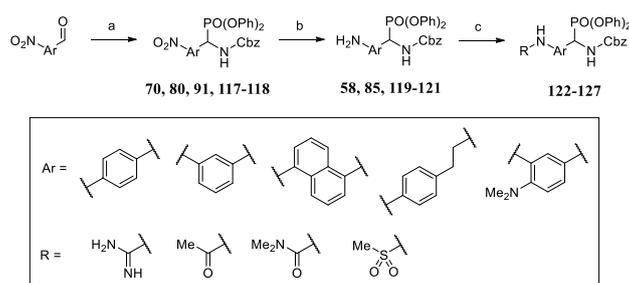
Scheme 3. Reagents and conditions. a) CbzNH₂, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h.

Still, most of them required a higher synthetic effort. The remaining compounds can be summarized in two synthetic schemes. For the first group (107-116), the Birum-Oleksyszyn reaction was conducted on the selected commercial aldehydes with Boc-protected amine, with the subsequent deprotection and guanylation for 115 and 116. This group comprises a variety of aniline and piperidine related moieties in the R¹ position (Scheme 4).



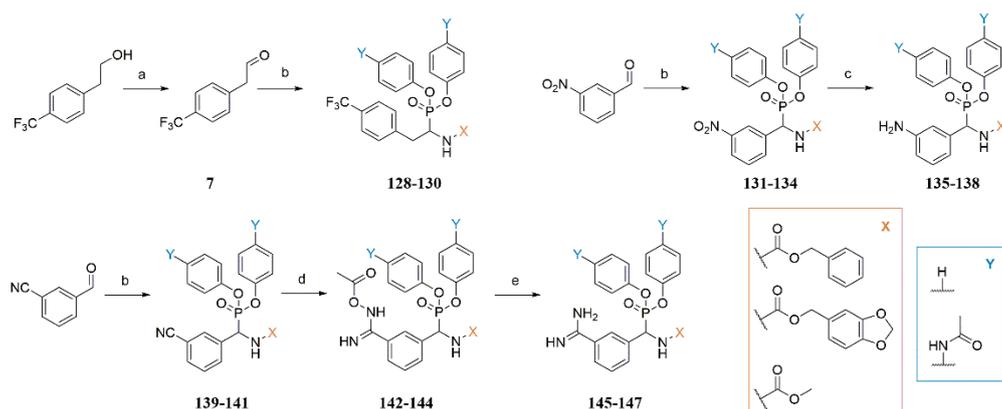
Scheme 4. Reagents and conditions. a) CbzNH₂, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h. b) TFA, DCM, rt, 1 h. c) *N,N'*-bis-Boc-1-guanylpyrazole, Et₃N, DCM, rt, 48 h.

For a second group of aniline-related compounds and aromatic guanidines (117-127), the starting materials were a variety of commercial nitroaryl aldehydes that, after a Birum-Oleksyszyn reaction, were reduced and subsequently substituted in some cases to generate methyl sulphonylamines, guanydines, dimethylarylureas and methyl amides (Scheme 5). The biochemical tests for ClpP inhibition of these compounds revealed two inhibitors with IC₅₀ values of 0.6 and 71.3, respectively (Table 3).

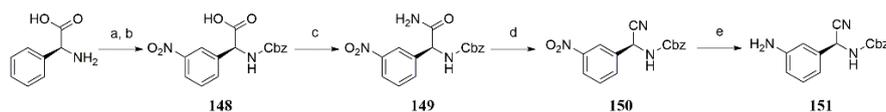


Scheme 5. Reagents and conditions. a) CbzNH₂, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h. b) Zn, THF:NH₄Cl (sat. sol.) (1:1), 0 °C, 1 h. c) 124-125: *N,N'*-bis-Boc-1-guanylpyrazole, Et₃N, DCM, rt, 48 h, then TFA, DCM, rt, 1 h; 122-123, 126-127: RCl, DIPEA, DCM, rt, 2 h.

Finally, further investigation around the two most potent R¹ moieties (aniline 85 and amidine 92) and the only lipophilic structure with activity (4-(trifluoromethyl)benzyl 16) was undertaken, with substitution of the -Cbz by other active substituents from the first screening, together with some warhead alternatives (paracetamol-like phosphonates and nitriles). The chemistry regarding this exploration can be found in Scheme 6 and Scheme 7. None of the 10 tested compounds revealed any pronounced ClpP inhibition (Table 4). A summary of all the compounds initial screening is reported in Figure S1 in the supporting information.



Scheme 6. Reagents and conditions. a) Dess-Martin periodinane, DCM, 0-25 °C, 2 h. b) CbzNH₂ (129, 133, 140)/methyl carbamate (128, 130-132, 134, 139, 141)/benzo[*d*][1,3]dioxol-5-ylmethyl carbamate (132), P(OPh)₃ (128, 131-132, 139)/ tris(4-acetamidophenyl) phosphite (129-130, 133-134, 140-141), Cu(OTf)₂, DCM, rt, 16 h. c) Zn, THF:NH₄Cl (aq. sat. sol.) (2:1), 0-25 °C, 16 h. d) NH₂OH·HCl, DIPEA, EtOH, 95 °C, 30-72 h, then acetic anhydride, MeCN, rt, 1 h; e) Pd(II)/C 10%, H₂ gas, AcOH, rt, 30 h.

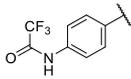
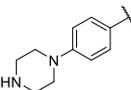
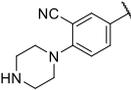
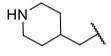
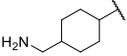
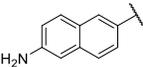
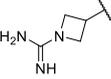
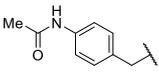
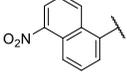
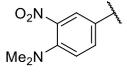


Scheme 7. Reagents and conditions. a) HNO₃, H₂SO₄, rt, 2 h. b) CbzCl, NaOH, H₂O, 0-25 °C, 2 h. c) Isobutylchloroformate, *N*-methylmorpholine, NH₃, DCM, 0-25 °C, 16 h. d) Burgess reagent, DCM, rt, 16 h. e) Zn, THF:NH₄Cl (aq. sat. sol.) (2:1), 0-25 °C, 16 h.

Table 3. Enzymatic inhibition of the hydrophilic exploration.

				%I (200 μM)	IC ₅₀ (μM)
96		-Cbz	-PO(OPh) ₂	15	ND
97		-Cbz	-PO(OPh) ₂	<1	ND

Chapter 2

98		-Cbz	-PO(OPh) ₂	15	ND
107		-Cbz	-PO(OPh) ₂	100	0.6 ± 0.0
108		-Cbz	-PO(OPh) ₂	10	ND
109		-Cbz	-PO(OPh) ₂	30	ND
110		-Cbz	PO(OPh) ₂	16	ND
111		-Cbz	-PO(OPh) ₂	15	ND
114		-Cbz	-PO(OPh) ₂	<1	ND
115		-Cbz	-PO(OPh) ₂	100	71.3 ± 2.4
116		-Cbz	-PO(OPh) ₂	18	ND
117		-Cbz	-PO(OPh) ₂	49	ND
118		-Cbz	-PO(OPh) ₂	39	ND

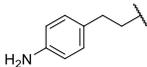
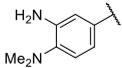
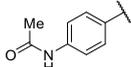
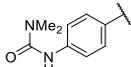
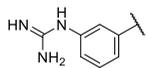
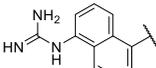
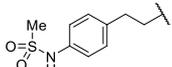
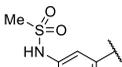
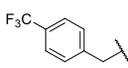
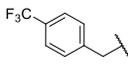
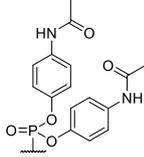
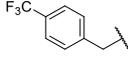
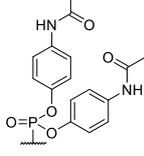
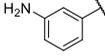
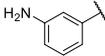
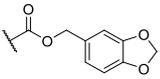
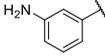
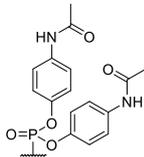
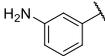
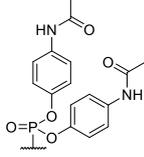
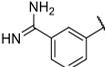
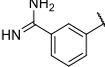
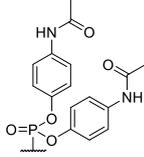
120		-Cbz	-PO(OPh) ₂	13	ND
121		-Cbz	PO(OPh) ₂	22	ND
122		-Cbz	-PO(OPh) ₂	8	ND
123		-Cbz	-PO(OPh) ₂	22	ND
124		-Cbz	-PO(OPh) ₂	15	ND
125		-Cbz	-PO(OPh) ₂	38	ND
126		-Cbz	-PO(OPh) ₂	10	ND
127		-Cbz	-PO(OPh) ₂	46	ND

Table 4. Enzymatic inhibition of the exploration around 16, 85 and 92.

%I (200 μM)	IC ₅₀ (μM)
-------------	-----------------------

Chapter 2

128			-PO(OPh) ₂	4	ND
129		-Cbz		1	ND
130				20	ND
135			-PO(OPh) ₂	7	ND
136			-PO(OPh) ₂	20	ND
137		-Cbz		21	ND
138				12	ND
145			-PO(OPh) ₂	30	ND
146		Cbz		29	ND

147				12	ND
151		-Cbz	-CN	8	ND

The hydrophilic exploration resulted in two additional (107 and 115), with 107 having an IC_{50} in the sub-micromolar range. Unfortunately, every alteration on the structure of our reference compounds (16, 85 and 92) led to loss of activity.

2.3.2 Biological evaluation of hits

The 14 hits identified after the different stages were submitted to further *in vitro* profiling. The selectivity properties *versus* chymotrypsin-like serine proteases were evaluated by screening the 14 compounds against α -chymotrypsin (bovine) at 200 μ M concentration. Most compounds showed no significant inhibition of chymotrypsin with the exception of 85, which resulted in a residual enzyme activity of 18 % (Table 5).

Cytotoxicity was tested against the human cell lines A549 (lung), HepG2 (liver) and HeLa (cervical cancer) in dose response (Table 5). Compounds 89, 90 and 107 were toxic for all cell lines while compounds 66 and 58 exerted high cytotoxicity (<10-fold of compound IC_{50}) against the lung cell line A549. Compounds 78 and 85 showed moderate toxicity against A549 and HepG2 cell lines compared with their *in vitro* potency against the target (IC_{50}). The cytotoxicity effects reported here against human cell lines could in principle be caused by the interaction with ClpP present in the human mitochondria as well by ClpP unrelated mechanisms.

Table 5. Enzymatic activity, cytotoxicity and activity against chymotrypsin of the selected hits.

Compd.	<i>E. coli</i> ClpP	Cytotoxicity EC_{50} (μ M)			Chymotrypsin
	IC_{50} (μ M)	HeLa	HepG2	A549	% of remaining

					activity (200 μM)
16	14.2 ± 1.3	≥ 100	≥ 100	≥ 100	93.9 ± 0.8
53	49.5 ± 0.5	≥ 100	≥ 100	≥ 100	≥ 100
58	39.8 ± 2.9	≥ 100	≥ 100	57.8 ± 6.7	≥ 100
66	8.2 ± 0.8	≥ 100	≥ 100	41.5 ± 3.8	46.0 ± 2.9
70	13.1 ± 1.2	≥ 100	≥ 100	≥ 100	≥ 100
71	48.1 ± 1.7	≥ 100	≥ 100	≥ 100	≥ 100
78	0.6 ± 0.1	≥ 100	28.4 ± 3.5	65.6 ± 8.3	≥ 100
85	0.5 ± 0.0	≥ 100	23.8 ± 2.8	27.5 ± 2.3	17.9 ± 1.6
86	38.0 ± 2.4	≥ 100	≥ 100	≥ 100	≥ 100
89	100.5 ± 8.0	19.9 ± 1.8	10.5 ± 1.4	5.9 ± 8.5	≥ 100
90	79.7 ± 7.2	19.9 ± 2.3	28.3 ± 2.5	25.6 ± 2.7	≥ 100
92	0.5 ± 0.0	≥ 100	≥ 100	≥ 100	≥ 100
107	0.6 ± 0.0	8.6 ± 1.2	1.1 ± 0.1	0.4 ± 0.0	≥ 100
115	71.3 ± 2.4	≥ 100	≥ 100	≥ 100	≥ 100

In order to investigate the mode of interaction between ClpP and selected compounds, surface plasmon resonance measurements were conducted. The known covalently binding compound chloromethyl ketone (Z-LY-CMK)³⁶ was used as positive control for irreversible binding.

Compounds with IC₅₀ values <10 μM were tested in a range of concentrations. In addition, the known covalent inhibitor Z-LY-CMK was tested as positive control. Z-LY-CMK clearly showed irreversible binding to ClpP, since the compound signal in the sensogram did not return to the baseline (0 RU), even after stop of the compound injection (at ~350 seconds in all experiments) (Figure 4A). In contrast, all compounds from this study (Figure 4B-F) bound reversibly to the protein, as shown by the signal drop to the baseline after stopping the injection. Moreover, the sensorgrams revealed rapid on- and off-rates for all newly identified ClpP inhibitors.

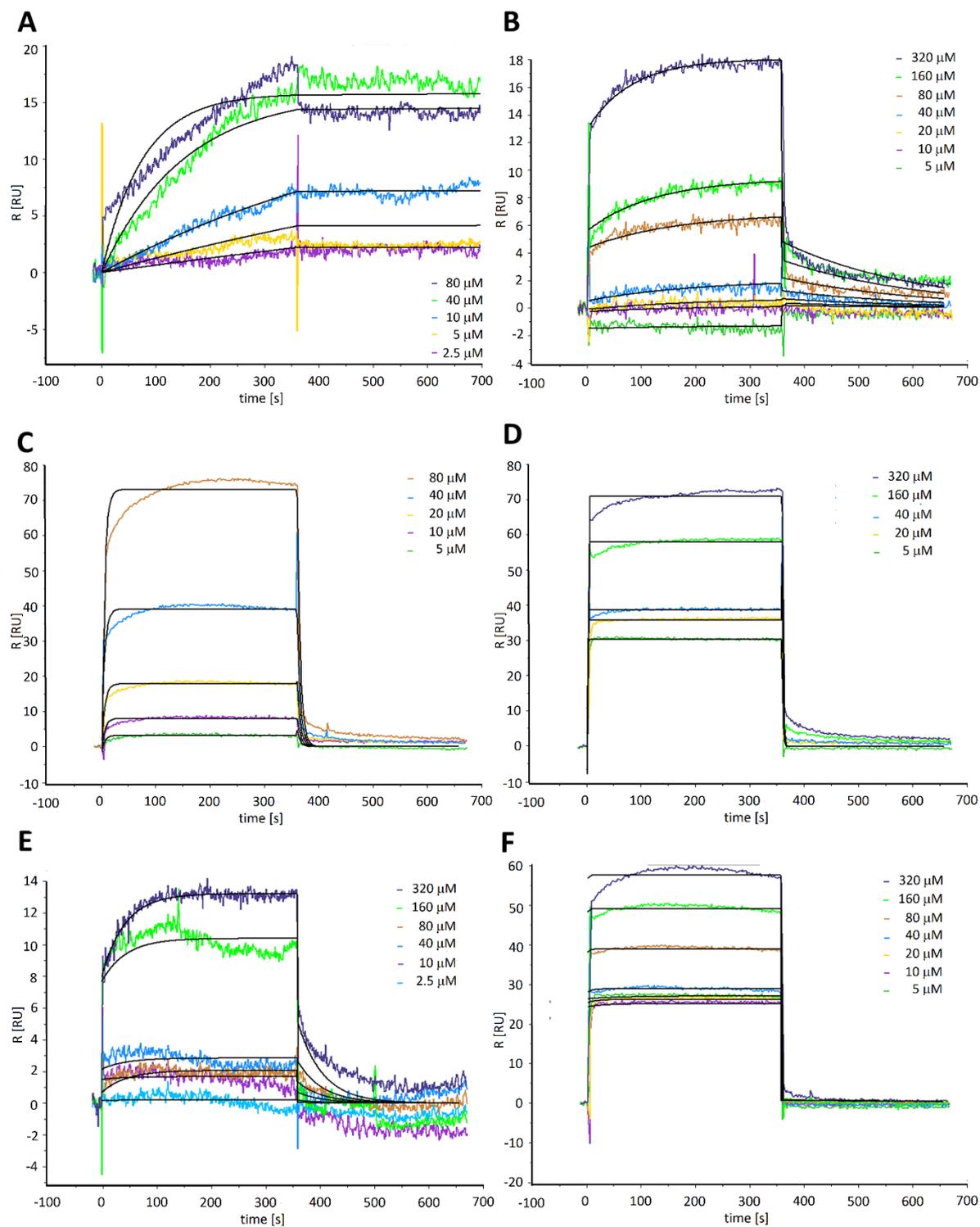


Figure 4. Surface plasmon resonance sensorgrams. A) Known covalent inhibitor Z-LY-CMK as control for irreversible binding.³⁶ B-F) Selected screening hits, B) 78, C) 107, D) 85, E) 66, F) 92.

Antibacterial assays. Since ClpP is not an essential protease in *E. coli*, an assay was required to investigate the influence of the ClpP inhibitors on bacterial growth rates. We utilized a method reported by Robinson *et al.*,¹⁶ who observed that a ClpP deletion mutant recovered more slowly from nitric oxide stress than the corresponding wild type, and adapted this assay to an HTS format. Nitric oxide stress was induced by addition of DPTA NONOate (2 mM) to the *E. coli* WT and the isogenic *E. coli* ClpP deletion strain ($\Delta clpP$). Although $\Delta clpP$ strain grew less well under our assay conditions (M9 minimal medium, 96-well format) compared to the wild type, we observed a small but significant difference in time to growth recovery after nitric oxide stress for the $\Delta clpP$ strain compared to the WT strain (Figure 5). Statistical analysis indicated that the ClpP deletion strain required approximately one hour longer than the wild type for growth recovery (see Figure S2 in the supporting information).

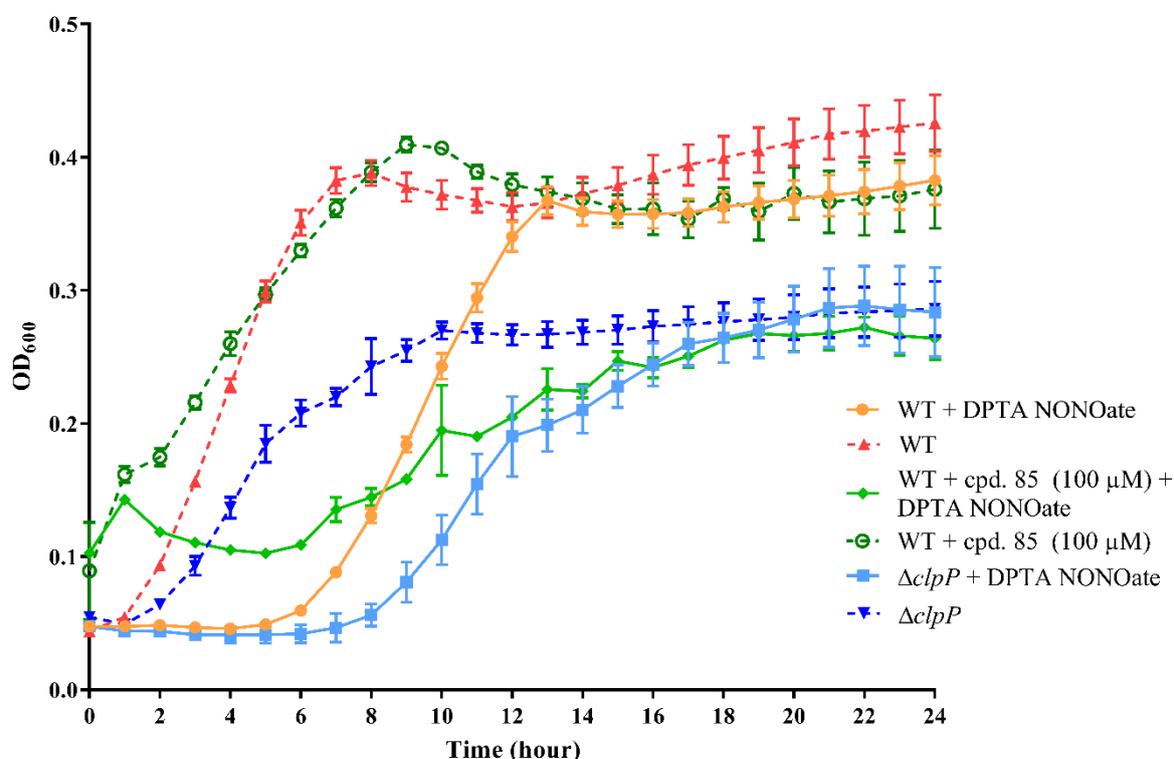


Figure 5. Comparison of the growth curves, depicted by OD_{600} , of *E. coli* WT and the isogenic mutant *E. coli* JW0427-1 (*clpP*-defective mutant), both in absence and presence of the •NO chemical donor DPTA NONOate (2 mM) in minimal M9 medium supplemented with 10 mM of glucose. Effect of compound 85 on bacterial growth is quantified by OD_{600} of *E. coli* BW25-113 (WT) in presence and in absence of DPTA NONOate (2 mM). Each value represents the mean of three independent experiments \pm standard deviation.

The fourteen hits identified in this study were tested in WT strain in presence of nitric oxide stress. Only compound 85 showed a remarkable effect in the WT growth (Figure 5), opposite to all the other compounds

which were also tested but without showing any effect. Therefore, while not every hit achieved the desired inhibition, the effects of compound 85 on the stressed WT strain showed a comparable growth delay observed to the $\Delta clpP$ strain exposed to nitric oxide stress conditions, consistent with a potential ClpP protease inhibition mediated effect of compound 85.

Figure 5 shows that in the presence of nitric oxide stress the growth rate of the WT strain is reduced and requires an additional 4 h to reach maximum absorbance (orange circle versus red triangle). The growth rate of the $\Delta clpP$ strain compared to the WT strain is reduced and is further reduced in the presence of nitric oxide stress, taking an additional 6 h to reach maximum absorbance (light blue square versus dark blue inverted triangle). The addition of compound 85 at 100 μM does not affect the growth rate of the WT strain (green ring versus red triangle). However, under nitric oxide stress conditions, the WT strain growth is reduced in presence of compound 85 compared to nitric oxide stress only (green diamond versus orange circle) and, interestingly, the growth of $\Delta clpP$ is similar to that of the WT strain in the presence of compound 85 (blue square versus green diamond). Moreover, the effect of compound 85 in $\Delta clpP$ growth was tested in presence and absence of nitric oxide stress to ensure that the compound-mediated effect on the WT bacteria growth was due to ClpP inhibition and not due to an off-target effect. As shown in Figure S3 (in the supporting information), the compound did not significantly affect the growth of the $\Delta clpP$ bacteria either in the absence or presence of nitric oxide stress conditions. This confirms that the effect of 85 on WT bacteria under nitric oxide stress conditions is most likely mediated through its inhibition of ClpP.

Selected compounds were also screened against the wild-type strains of *S. aureus*, *Acinetobacter baumannii*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *E. coli* in a standard bacterial growth assay. Two *E. coli* mutants with either *lpxC* defect (impaired in lipidA synthesis) or *tolC* defect (efflux pump defect) were also included. At 100 μM compound concentration, only 115 inhibited the growth of *S. aureus* WT (% inhibition of growth 99.9 ± 0.04), while 90 (98.9 ± 0.05), 107 (69.9 ± 5.6) and 115 (99.0 ± 0.5) inhibited the efflux pump deficient *E. coli* strain. The mode of action underlying this growth inhibition remains elusive and could be caused by mechanisms unrelated to *E. coli* ClpP. In order to investigate whether the compounds are efflux pump substrates, the growth of *E. coli* wild type was examined in presence of test compounds (concentration 50 μM) and 25 $\mu\text{g/ml}$ of phenylalanine-arginine beta-naphthylamide (PA β N), a known efflux pump substrate. At 24 h compound 90 inhibited bacterial growth (98.5 ± 1.2). In order to verify whether the observed effect of compound 90 was due to ClpP inhibition, the assay was repeated using the *E. coli* $\Delta clpP$ strain. The same output of the assay with the wild type strain was obtained (growth inhibition 99.4 ± 0.7), we can, therefore, assert that compound 90 addresses a different target that is influencing bacterial growth in presence of the efflux pump substrate.

A summary of the compounds active in bacteria can be found in the supporting information (Table S1).

Molecular docking. Potential binding modes of the most potent inhibitors (92 and 85) within the active site of ClpP were investigated by molecular docking of the compounds into the X-ray crystal structure of *E. coli* ClpP (PDB ID 2FZS) using GOLD.

Clustering of the docking poses of 92 revealed two preferred binding modes (Figure 6A-D). The top-ranked pose of the first cluster (Figure 6A&B) shows the benzamidine group to be positioned deeply within the S1 pocket, while a hydrogen bond network between the phosphonate and residues Gly68 (a constituent of the oxanion hole) and Leu125 is well established. However, the distance between the side chain oxygen atom of Ser97 and the phosphorus atom of the ligand is larger than required for the expected nucleophilic attack (3.35 Å). The second predominant binding mode revealed the benzamidine group to be solvent-exposed and the docked ligand shares several hydrogen bonds with the protein (Figure 6C&D). The interaction energy between docked ligand and ClpP was calculated using the Amber10:EHT force field. The top-ranked docking pose of cluster 1 (R¹ moiety placed in the S1 pocket) revealed more favorable interaction energy (-62.5 kcal mol⁻¹) compared to the top-ranked pose of cluster 2 (-54.6 kcal mol⁻¹).

The predicted binding mode of 85 is shown in Figure 6E&F. Only one cluster was identified and the binding mode revealed the phenyl groups of the diarylphosphonate to be solvent-exposed, while the aniline moiety is positioned inside the S1 pocket. Again, several hydrogen bonds are formed between the phosphonate group and the protein, but Ser97 did not display a favorable position for the nucleophilic attack.

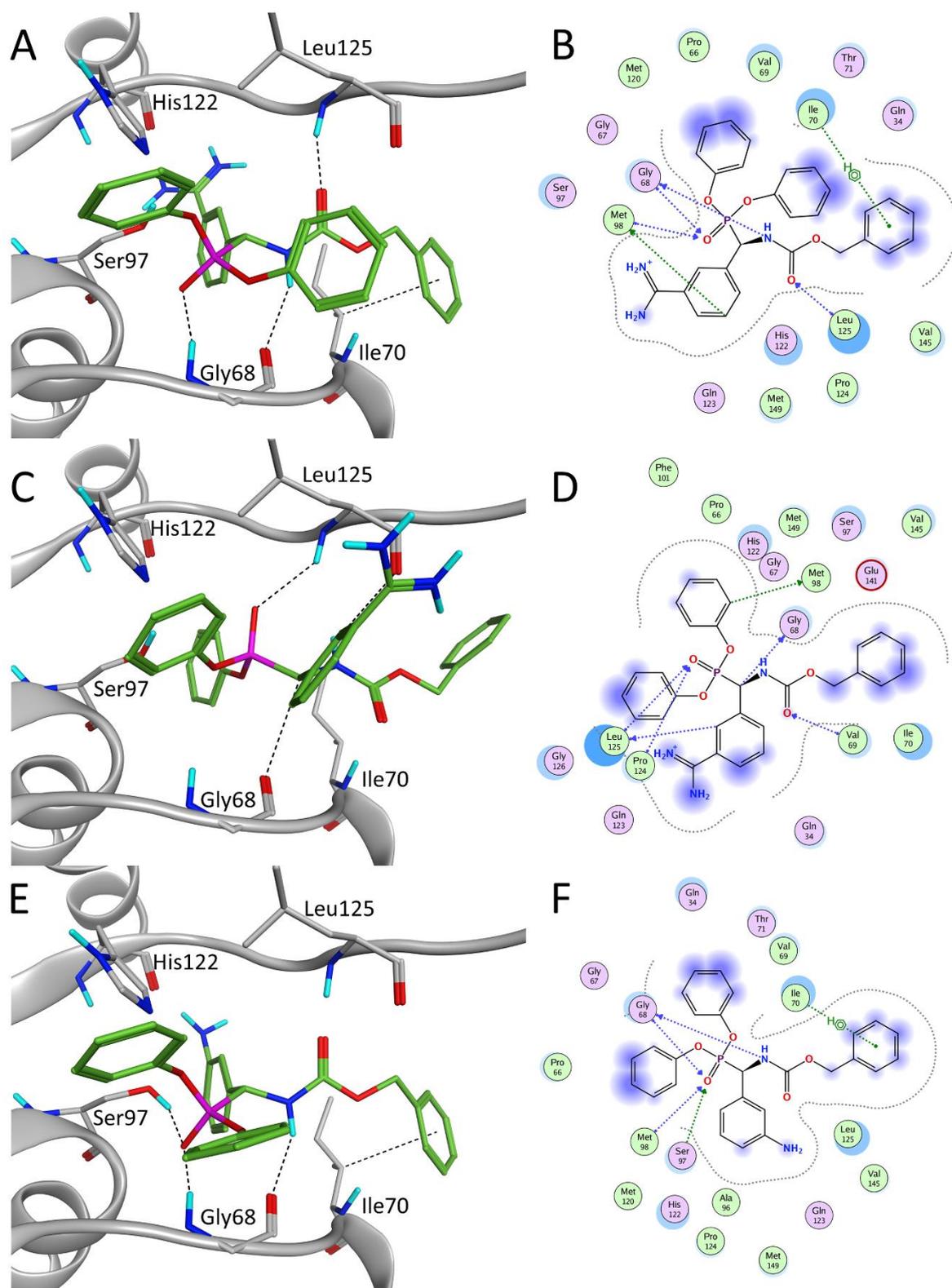


Figure 6. Computational prediction of potential binding modes for 92 and 85 within the *E. coli* ClpP crystal structure (PDB ID: 2FZS). A and C) Docking poses of the two main docking clusters of compound 92. E) Pose of compound 85. Black dotted lines indicate hydrogen bonds between the ligands and the protein. B, D and F)

2D interaction plots between protein and 92 of cluster 1 (B), cluster 2 (D) and 85 (F) where the polar residues are indicated in purple, the non-polar or charged residues in green, and the solvent exposure by blue shadow. The black dotted line designates the proximity contour. The green dotted arrows indicate hydrogen bonds involving amino acid side chain atoms (donors and acceptors) while blue dotted arrows indicate hydrogen bonds accepted or donated by protein backbone atoms. Moreover, arene-H interactions are shown as green dotted line.

2.4 Discussion and conclusion

An extensive chemical exploration and enzymatic screening identified 14 compounds inhibiting *E. coli* ClpP *in vitro* with sub-micromolar IC₅₀ values for 78, 85, 92 and 107. Despite the expected hydrophobicity of the protease recognition pocket, compounds containing polar residues in R¹ position displayed the highest inhibitory activity. Molecular docking analysis of compound 85 and 92 revealed that the most favorable poses had the aniline or benzamidine group deeply positioned within the recognition pocket. The diphenyl phosphonate warhead was crucial, with none of the replacements or small modifications attempted maintaining the inhibitory activity.

Surface plasmon resonance demonstrated a reversible binding for all tested compounds. With help of the docking studies of 85 and 92, it can be hypothesized that the inhibitory poses do not allocate the phosphonate esters of the ligands in a favorable position to form the pentacoordinate transition state (Figure 2) after the attack by Ser97. Albeit unexpected, this reversible binding of the chemical family has a precedent in the KLK4 inhibitors reported by Van Soom *et al.*⁴⁸

Benzamidine 92 emerges as the safest option for further optimization due to its potent enzymatic inhibition, the absence of activity against chymotrypsin and lack of toxicity against the tested eukaryotic cell lines. Even though the latter could be related to a compound's incapability to enter eukaryotic cells or to the possibility of being a substrate of an efflux pump and would also explain the lack of activity in the nitric oxide stress assay. At the same time, aniline 85 showed a reduction of growth in *E. coli* WT under nitric oxide stress conditions, consistent with a ClpP-mediated effect. However, it requires further improvement in terms of limiting its toxicity against human cell lines and decreasing activity against chymotrypsin. Both compounds, as well as 16, 66, 70, 78 and 107 are significantly more potent ClpP inhibitors compared to the so far only known inhibitor Z-LY-CMK. In order to enlarge the scope and to understand the lack of effect of some of our inhibitors in the nitric oxide stress assay, further exploration of chemical alternatives is needed. Given the already known potential of covalent binding compounds as antimicrobial agents, for example the huge success of β -lactam antibiotics (*e.g.* penems, cephalosporins, carbapenems, monobactams), the development of a covalent binder for *E. coli* ClpP should focus on the replacement of the diaryl phosphonate by a different covalent warhead for serine proteases. The comprehensive R¹ moiety library developed in this study may guide future work in the

field combining them with warheads such as boronates, based on the success of bortezomib with *M. tuberculosis* ClpP and its inhibition of human 26S proteasome.³³

2.5 Materials and methods

2.5.1 Biological evaluation

Protein production. *E. coli* ClpP protein carrying a C-terminal His₆ affinity tag was produced starting from pETDclpPec (ORF ECK0431).⁶⁵ pETDclpPec was transformed into *E. coli* SG1146a strain (Δ *clpP*) for overexpression. Overnight cultures were grown at 37 °C on an orbital shaker at 130 rpm, then diluted in fresh LB medium including 100 µg/mL ampicillin, and grown until an OD₆₀₀ of 0.6 was reached. Induction was carried out in 5 L flasks by 1 mM isopropyl β-D-1-thiogalactopyranoside at 30 °C at 180 rpm over 5 h. Cell lysis of cooled samples was conducted using Precellys Evolution (Bertin Technologies, France). Supernatant was applied to Ni-NTA (Sigma Aldrich, USA) followed by batchwise washing and elution steps, sequentially applying buffer A (pH 7.6, 50 mM Tris-HCl buffer, 150 mM NaCl, 10 mM imidazole), buffer B (pH 7.6, 50 mM Tris-HCl buffer, 150 mM NaCl, 20 mM imidazole), buffer C (pH 7.6, 50 mM Tris-HCl buffer, 150 mM NaCl, 500 mM Imidazole) and buffer D (pH 7.6, 20 mM Tris-HCl buffer, 100 mM NaCl, 5 mM MgCl, 10% (v/v) glycerol). Purified ClpP protein was frozen in liquid nitrogen and stored at -80 °C. Protein purity and concentration were determined by SDS-PAGE and the Bradford assay, respectively.

Assay Development and Screening. A microplate screening assay with a fluorescence based readout was developed to measure ClpP proteolytic activity. The ClpP activity assay was performed with the fluorogenic substrate Suc-LY-AMC (Enzo Life Sciences, Germany) at 75 µM and *E. coli* ClpP at 625 nM in 100 mM NaCl and 100 mM Hepes pH 7.5, 0.05% Brij[®] 35 (#P1254, Sigma Aldrich, USA). Compound selectivity for ClpP was assessed by measuring the level of activity of the compounds in the presence of 40 nM of alpha-chymotrypsin (#C4129, Sigma-Aldrich, USA) and 100 µM of Suc-LY-AMC substrate, in a buffer containing 150 mM NaCl, 10 mM CaCl₂, 50 mM Tris-HCl and 0.05% Brij[®] 35. Assays were performed in 384-well black, flat-bottom microtiter plates (#3820, Corning Inc., Corning, USA). All compounds were dissolved in 99.8% DMSO (ROTIPURAN[®] CAS No.[67-68-5], Carl Roth GmbH, Germany), at a stock concentration of 10 mM. Compound serial dilutions were carried out in 1:3 or 1:2 dilutions and stored at -20 °C. Screening of compounds was conducted at a final compound concentration of 200 µM in triplicates, with compound transfer carried out using acoustic liquid handling (Echo 550, Labcyte, USA). ClpP protein and compounds were incubated for 10 min at 30 °C, followed by addition of the fluorescent substrate Suc-LY-AMC. The reaction was monitored by following the increase of fluorescence (excitation 350 nm, emission 435 nm) at 30 °C over 1h. Vehicle controls contained the same DMSO concentration (2% v/v) without compound and chloromethyl ketone (Z-LY-CMK) (#4016342, Bachem, Switzerland) was used as a positive control (200 µM). Assays were performed under automated conditions (Fluent[®] 780, Tecan, Switzerland) equipped with a microplate reader

(Infinite M1000 Pro, Tecan, Switzerland). Calculation of Z prime (Z') for validation was performed according to Zhang *et al.*⁶⁶ and plates were considered valid for further analyses where Z' was >0.6 . Data analysis was conducted using Prism 7.02 (GraphPad Software, USA).

Surface Plasmon Resonance Spectroscopy (SPR). Measurements were conducted on a flow based SPR instrument (Sierra spr-16, Bruker Daltonics, USA). *E. coli* ClpP was immobilized to an amino coupling chip at a concentration of 80 $\mu\text{g}/\text{mL}$, in 10 mM sodium acetate pH 4, according to the manufacturer's protocols, with a buffer containing 150 mM NaCl, 10 mM Hepes, 3 mM EDTA and 0.05% (v/v) Tween-20. The flow rate for protein immobilization was 10 $\mu\text{L}/\text{min}$. The binding assay was performed in immobilization buffer with added DMSO (3.2% (v/v) final) at a flow rate of 20 $\mu\text{L}/\text{min}$, 6 min injection and up to 300 s of dissociation time. The compounds were tested in a range of concentration between 2.5 and 320 μM . To compensate for non-specific interactions and solvent effects, signals were subjected to reference channel subtraction, DMSO and bulk-shift correction and further analyzed with Analyzer 3 (Bruker Daltonics, USA).

Cytotoxicity and cell-viability assays. Potential toxicity of the tested compounds was evaluated by ATP quantification using the CellTiter-Glo® viability assay kit (Promega, USA) and human cell lines A549, HepG2, HeLa. Cells were cultured in 95% air incubator at 5% CO_2 at 37 °C (Heracell™ 240, Thermo Fisher Scientific, USA). The assays were performed in 96 white, flat bottom, sterile plates (# 781073, Greiner Bio-One, Germany), with an assay volume of 20 μL . Cells were seeded at day zero at concentration of 2000 cells/well, except for A549 (500 cells/well), and placed in 95% air incubator, 5% CO_2 at 37 °C. After 24 h, 200 nL of test compounds were transferred into the plate using an Echo® 550 liquid handler (Labcyte, USA), resulting in a final DMSO concentration of 1%, and were further incubated for 48 h. Luminescence was quantified by EnVision plate reader (PerkinElmer, Germany) and compared to DMSO-treated cells. The compounds were tested in dose-response at 1:3 dilutions starting from 100 μM . DMSO (1%) and valinomycin (10 μM) were used as negative and positive controls respectively. Data analysis was conducted using Prism 7.02 (GraphPad Software, USA). Plates with $Z' > 0.5$ were accepted.

Antibacterial assays. Compounds were tested for antimicrobial activity using seven different strains (Table S1, supporting information). All assays were conducted in 96-well, flat bottom, sterile plates (#167008, Nunc, VWR, USA). 5 mL of fresh sterile saline solution was inoculated with a single colony from a Mueller Hinton Agar (MHA) plate of the bacteria strain (not older than 24 h). Bacterial suspension was adjusted to contain 1×10^6 CFU/mL in fresh sterile Mueller Hilton Broth (MHB) media.

Compounds to be tested were transferred into the assay plate in triplicate for screening (at 100 μM), along with controls of 2% DMSO or ciprofloxacin (the latter employed at MIC concentration for each strain) and 100 μL of the bacterial solution was added (final inoculum 5×10^5 CFU/mL). The final volume employed for the assay was 200 μL .

The absorbance at 600 nm was measured with the Multiskan GO plate reader (Thermo Fisher Scientific, Finland) or Varioskan LUX plate reader (Thermo Fisher Scientific, Finland) at time 0 and different time points and used for quantifying bacterial growth. Plates were incubated in a plate shaker (500 rpm) at 37 °C between the measurements.

Selected studies were conducted in the presence of an efflux pump substrate (25 μ M of Phe-Arg β -naphthylamide dihydrochloride (#P4157, Sigma-Aldrich, USA) with *Escherichia coli* BW25-113 (wild type) and isogenic *E. coli* JW0427-1 (Δ *clpP*),⁶⁷ (derived from *E. coli* BW25-113).

Nitric oxide stress was induced by adding 2 mM of DPTA NONOate ((Z)-1-[N-(3-aminopropyl)-N-(3-ammoniopropyl)amino]diazene-1,1,2-diolate, Cayman Chemical), dissolved in NaOH (0.14 mM stock solution) in a *E. coli* BW25-113 culture at OD₆₀₀ = 0.1 in M9 media supplemented with 10 mM glucose. The bacterial culture was previously grown in MHB overnight, shaking at 250 rpm at 37 °C. On the experimental day, fresh M9 medium supplemented with 10 mM glucose was inoculated at 1:100 ratio with the overnight culture and grown until approximately OD₆₀₀ = 0.3. The assay was conducted at 37 °C while shaking (500 rpm), and the bacterial growth was monitored measuring the OD₆₀₀ every hour for 15 or 24 h. The isogenic mutant *E. coli* JW0427-1 was used as control, an internal control or as strain study in a control experiment. The mutant was treated in the same way of the WT.

Molecular Docking. Molecular docking was performed using GOLD version 5.4.1 (Cambridge Crystallographic Data Centre, Cambridge, UK). Selection of the optimal scoring function was carried out by redocking the only co-crystallized non-covalently bound ClpP inhibitor AV145 (PDB ID 5DL1) into *S. aureus* ClpP. For newly identified inhibitors, the *E. coli* ClpP X-ray crystal structure 2FZS was selected. All protein structures were prepared with the molecular modeling software suite Molecular Operating Environment (MOE, Chemical Computing Group Inc., Montreal, Canada) version 2016.0802 and energy minimized using an Amber10:EHT force field with implicit solvation model (R-Field). Three-dimensional coordinates of ligands to be docked were generated within MOE. Redocking of co-crystallized AV145 into the ClpP structure 5DL1 (all 14 chains) revealed the scoring function GoldScore to be best suited for docking non-covalent compounds into ClpP, resulting in top-ranked docking poses with heavy atom root-mean-square deviation (RMSD) values 0.68 Å for all 14 monomers. The search space for compounds to be docked into *E. coli* ClpP was defined by a sphere of 15 Å radius centered on atom C10 of the ligand. For each compound, 50 docking runs were conducted. The early termination option was switched off. The Asn150 amide side was allowed to flip by 180 degrees.

2.5.2 Chemistry

All the chemistry procedure are reported in section **Appendix, Part 1. Chemistry procedures – Chapter 2**

2.6 Abbreviation list

3,4-DCI	3,4-dichloroisocoumarin
AAA+	ATPases Associated with diverse cellular Activities
ADEPs	Acyldepsipeptides
ClpA	Caseinolytic protease subunit A
ClpP	Caseinolytic protease proteolytic subunit
ClpX	Caseinolytic protease subunit X
Comp.	Compound
DPTA NONOate	(Z)-1-[N-(3-Aminopropyl)-N-(3-ammoniopropyl)amino]diazen-1-ium-1,2-diolate
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
HTS	High-throughput screening
IC ₅₀	Concentration of a drug that is required for 50% inhibition <i>in vitro</i>
LB	Luria-Bertani
LCMS	Liquid chromatography–mass spectrometry
MCF-7	Michigan Cancer Foundation-7 human cell line
MHA	Müller-Hinton agar
MHB	Müller-Hinton broth
MOE	Molecular Operating Environment
Ni-NTA	Nickel-charged affinity
NO•	Nitric oxide
PaβN	Phenylalanine-arginine beta-naphthylamide
PDB	Protein data bank
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SDS-PAGE	Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
SPR	Surface Plasmon Resonance Spectroscopy
Suc-LY-AMC	4-(((S)-1-(((S)-2-(4-Hydroxyphenyl)-1-(4-methyl-2-oxo-2H-chromen-7-yl)ethyl)amino)-4-methyl-1-oxopentan-2-yl)amino)-4-oxobutanoic acid
WT	Wild type
Z'	Z prime
Z-LY-CMK	Benzyl ((S)-1-(((S)-4-chloro-1-(4-hydroxyphenyl)-3-oxobutan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate
$\Delta clpP$	Mutant defective in <i>clpP</i>
$\Delta lpxC$	Mutant impaired in lipid A synthesis

2.7 References

1. Bartlett, J. G., Gilbert, D. N. & Spellberg, B. Seven ways to preserve the miracle of antibiotics. *Clin. Infect. Dis.* 56 (10), 1445-1450 (2013).
2. Chellat, M. F., Raguz, L. & Riedl, R. Targeting antibiotic resistance. *Angew. Chem. Int. Ed. Engl.* 55 (23), 6600-6626 (2016).
3. Kupferschmidt, K. Resistance fighters. *Science.* 352 (6287), 758-761 (2016).
4. WHO *Antimicrobial Resistance: Global Report on Surveillance 2014.* (2014).
5. Rossolini, G. M., Arena, F., Pecile, P. & Pollini, S. Update on the antibiotic resistance crisis. *Curr. Opin. Pharmacol.* 18, 56-60 (2014).
6. Butler, M. S., Blaskovich, M. A. & Cooper, M. A. Antibiotics in the clinical pipeline at the end of 2015. *J. Antibiot.* 70 (1), 3-24 (2017).
7. Goodreid, J. D., Janetzko, J., Santa Maria, J. P. Jr., Wong, K. S., Leung, E., Eger, B. T., *et al.* Development and characterization of potent cyclic acyldepsipeptide analogues with increased antimicrobial activity. *J. Med. Chem.* 59 (2), 624-646 (2016).
8. Brotz-Oesterhelt, H. & Sass, P. Bacterial caseinolytic proteases as novel targets for antibacterial treatment. *Int. J. Med. Microbiol.* 304 (1), 23-30 (2014).
9. Frees, D., Brondsted, L. & Ingmer, H. Bacterial proteases and virulence. *Subcell. Biochem.* 66, 161-192 (2013).
10. Arribas, J. & Castaño, J. G. A comparative study of the chymotrypsin-like activity of the rat liver multicatalytic proteinase and the ClpP from *Escherichia coli*. *J. Biol. Chem.* 268 (28), 21165-21171 (1993).
11. Frees, D., Sorensen, K. & Ingmer, H. Global virulence regulation in *Staphylococcus aureus*: pinpointing the roles of ClpP and ClpX in the sar/agr regulatory network. *Infect. Immun.* 73 (12), 8100-8108 (2005).

12. Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. & Berche, P. The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol. Microbiol.* 35 (6), 1286-1294 (2000).
13. Gaillot, O., Bregenholt, S., Jaubert, F., Di Santo, J. P. & Berche, P. Stress-induced ClpP serine protease of *Listeria monocytogenes* is essential for induction of listeriolysin O-dependent protective immunity. *Infect Immun.* 69 (8), 4938-4943 (2001).
14. Kwon, H. Y., Ogunniyi, A. D., Choi, M. H., Pyo, S. N., Rhee, D. K. & Paton, J. C. The ClpP protease of *Streptococcus pneumoniae* modulates virulence gene expression and protects against fatal pneumococcal challenge. *Infect. Immun.* 72 (10), 5646-5653 (2004).
15. Park, C. Y., Kim, E. H., Choi, S. Y., Tran, T. D., Kim, I. H., Kim, S. N., *et al.* Virulence attenuation of *Streptococcus pneumoniae* clpP mutant by sensitivity to oxidative stress in macrophages via an NO-mediated pathway. *J. Microbiol.* 48 (2), 229-235 (2010).
16. Robinson, J. L. & Brynildsen, M. P. An ensemble-guided approach identifies ClpP as a major regulator of transcript levels in nitric oxide-stressed *Escherichia coli*. *Metab. Eng.* 31, 22-34 (2015).
17. Flynn, J. M., Neher, S. B., Kim, Y.-I., Sauer, R. T. & Baker, T. A. Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell.* 11, 671-683 (2003).
18. Zhao, B. B., Li, X. H., Zeng, Y. L. & Lu, Y. J. ClpP-deletion impairs the virulence of *Legionella pneumophila* and the optimal translocation of effector proteins. *BMC Microbiol.* 16 (1), 174 (2016).
19. Qiu, D., Eisinger, V. M., Head, N. E., Pier, G. B. & Yu, H. D. ClpXP proteases positively regulate alginate overexpression and mucoid conversion in *Pseudomonas aeruginosa*. *Microbiology.* 154, 2119-2130 (2008).
20. Alexopoulos, J. A., Guarne, A. & Ortega, J. ClpP: a structurally dynamic protease regulated by AAA+ proteins. *J. Struct. Biol.* 179 (2), 202-210 (2012).
21. Ma, W., Tang, C. & Lai, L. Specificity of trypsin and chymotrypsin: Loop-motion-controlled dynamic correlation as a determinant. *Biophys. J.* 89 (2), 1183-1193 (2005).
22. Schelin, J., Lindmark, F. & Clarke, A. K. The ClpP multigene family for the ATP-dependent Clp protease in the cyanobacterium *Synechococcus*. *Microbiology.* 148, 2255-2265 (2002).

-
23. Gur, E., Ottofueling, R. & Dougan, D. A. *Regulated Proteolysis in Microorganisms*. 1 ed. Springer Netherlands. (2013).
 24. Olivares, A. O., Baker, T. A. & Sauer, R. T. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat. Rev. Microbiol.* *14* (1), 33-44 (2016).
 25. Brotz-Oesterhelt, H., Beyer, D., Kroll, H. P., Endermann, R., Ladel, C., Schroeder, W., *et al.* Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat. Med.* *11* (10), 1082-1087 (2005).
 26. Malik, I. T. & Brotz-Oesterhelt, H. Conformational control of the bacterial Clp protease by natural product antibiotics. *Nat. Prod. Rep.* *34* (7), 815-831 (2017).
 27. Bottcher, T. & Sieber, S. A. beta-Lactones as privileged structures for the active-site labeling of versatile bacterial. *Angew. Chem., Int. Ed.* *47* (24), 4600-4603 (2008).
 28. Bottcher T. & Sieber, S. A. beta-Lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of *Staphylococcus aureus*. *J. Am. Chem. Soc.* *130* (44), 14400-14401 (2008).
 29. Bottcher, T. & Sieber, S. A. beta-Lactones decrease the Intracellular virulence of *Listeria monocytogenes* in macrophages. *Chemmedchem.* *4* (8), 1260-1263 (2009).
 30. Bottcher, T. & Sieber, S. A. Structurally refined beta-lactones as potent inhibitors of devastating bacterial virulence factors. *Chembiochem.* *10* (4), 663-666 (2009).
 31. Hackl, M. W., Lakemeyer, M., Dahmen, M., Glaser, M., Pahl, A., Lorenz-Baath, K., *et al.* Phenyl esters are potent inhibitors of caseinolytic protease P and reveal a stereogenic switch for deoligomerization. *J. Am. Chem. Soc.* *137* (26), 8475-8483 (2015).
 32. Pahl, A., Lakemeyer, M., Vielberg, M. T., Hackl, M. W., Vomacka, J., Korotkov, V. S., *et al.* Reversible Inhibitors Arrest ClpP in a defined conformational state that can be revoked by ClpX association. *Angew. Chem. Int. Ed. Engl.* *54* (52), 15892-15899 (2015).
 33. Moreira, W., Ngan, G. J., Low, J. L., Poulsen, A., Chia, B. C., Ang, M. J., *et al.* Target mechanism-based whole-cell screening identifies bortezomib as an inhibitor of caseinolytic protease in *mycobacteria*. *mBio.* *6* (3), e00253-15 (2015).

-
34. Akopian, T., Kandrор, O., Tsu, C., Lai, J. H., Wu, W. G., Liu, Y. X., *et al.* Cleavage specificity of *Mycobacterium tuberculosis* ClpP1P2 protease and identification of novel peptide substrates and boronate inhibitors with anti-bacterial activity. *J. Biol. Chem.* 290 (17), 11008-11020 (2015).
 35. Mundra, S., Thakur, V., Bello, A. M., Rathore, S., Asad, M., Wei, L., Y *et al.* A novel class of *Plasmodial* ClpP protease inhibitors as potential antimalarial agents. *Bioorg. Med. Chem.* 25 (20), 5662-5677 (2017).
 36. Szyk, A. & Maurizi, M. R. Crystal structure at 1.9Å of *E. coli* ClpP with a peptide covalently bound at the active site. *J. Struct. Biol.* 156 (1), 165-174 (2006).
 37. Lamden, L. A. B. P. A. Aminoalkylphosphonofluoridate derivatives: rapid and potentially selective inactivators of serine peptidases. *Biochem. Biophys. Res. Commun.* 112 (3), 1085-1090 (1983).
 38. Joossens, J., Van der Veken, P., Surpateanu, G., Lambeir, A. M., El-Sayed, I., Ali, O. M., *et al.* Diphenyl phosphonate inhibitors for the urokinase-type plasminogen activator: Optimization of the P4 position. *J. Med. Chem.* 49 (19), 5785-5793 (2006).
 39. Joossens, J., Ali, O. M., El-Sayed, I., Surpateanu, G., Van der Veken, P., Lambeir, A. M., *et al.* Small, potent, and selective diaryl phosphonate inhibitors for urokinase-type plasminogen activator with in vivo antimetastatic properties. *J. Med. Chem.* 50 (26), 6638-6646 (2007).
 40. Van der Veken, P., Soroka, A., Brandt, I., Chen, Y. S., Maes, M. B., Lambeir, A. M., *et al.* Irreversible inhibition of dipeptidyl peptidase 8 by dipeptide-derived diaryl phosphonates. *J. Med. Chem.* 50 (23), 5568-5570 (2007).
 41. Winiarski, L., Oleksyszyn, J. & Sienczyk, M. Human neutrophil elastase phosphonic inhibitors with improved potency of action. *J. Med. Chem.* 55 (14), 6541-6553 (2012).
 42. Pietrusewicz, E., Sienczyk, M. & Oleksyszyn, J. Novel diphenyl esters of peptidyl alpha-aminoalkylphosphonates as inhibitors of chymotrypsin and subtilisin. *J. Enzyme Inhib. Med. Chem.* 24 (6), 1229-1236 (2009).
 43. Burchacka, E., Skorenski, M., Sienczyk, M. & Oleksyszyn, J. Phosphonic analogues of glutamic acid as irreversible inhibitors of *Staphylococcus aureus* endoproteinase GluC: An efficient synthesis and inhibition of the human IgG degradation. *Bioorg. Med. Chem. Lett.* 23 (5), 1412-1415 (2013).

-
44. Burchacka, E., Zdzalik, M., Niemczyk, J. S., Pustelny, K., Popowicz, G., Wladyka, B., *et al.* Development and binding characteristics of phosphonate inhibitors of SplA protease from *Staphylococcus aureus*. *Protein Sci.* 2014, 23 (2), 179-189. (2014).
 45. Dess, D. B. & Martin, J. C. Readily accessible 12-I-5 oxidant for the conversion of primary and secondary alcohols to aldehydes and ketones. *J. Org. Chem.* 48 (22), 4155-4156. (1983).
 46. Van der Veken, P., Sayed, I., Joossens, J., Stevens, C., Augustyns, K. & Haemers, A. Lewis acid catalyzed synthesis of *N*-protected diphenyl 1-aminoalkylphosphonates. *Synthesis.* 36, 634-638 (2005).
 47. Okano, K., Okuyama, K. I., Fukuyama, T. & Tokuyama, H. Mild debenzoylation of aryl benzyl ether with BCl₃ in the presence of pentamethylbenzene as a non-Lewis-basic cation scavenger. *Synlett.* 13, 1977-1980 (2008).
 48. Van Soom, J., Crucitti, G. C., Gladysz, R., Van der Veken, P., Di Santo, R., Stuyver, I., *et al.* The first potent diphenyl phosphonate KLK4 inhibitors with unexpected binding kinetics. *MedChemComm.* 6 (11), 1954-1958 (1979).
 49. Oleksyszyn, J. S. L. & Mastalerz, P. Diphenyl 1-aminoalkanephosphonates. *Synthesis.* 12, 985-986 (1979).
 50. Powers, J. C., Boduszek, B. & Oleksyszyn, J. Basic α -Aminoalkylphosphonate Derivatives. US 5686419 (1997).
 51. Mazur, R. H. *N*-Adamantane-Substituted Tetrapeptide Amides. US4273704A. (1981).
 52. Grundl, M., Oost, T., Pautsch, A., Peters, S., Riether, D. & Wienen, W. Substituted *N*-[1-Cyano-2-(phenyl)ethyl]-2-azabicyclo[2.2.1]heptane-3-carboxamide Inhibitors of Cathepsin C. WO 2013041497 (2013).
 53. Augustyns, K., Van der Veken, P., Lambeir, A. M. V. R., Scharpe, S. & Haemers, A. Novel Urokinase Inhibitors. WO2007045496 (2007).
 54. Augustyns, K., Lambeir, A. M., Messaggio, J. & Van der Veken, P. Activity-Based Probes for the Urokinase Plasminogen Activator. WO2012152807 (2012).
 55. Joossens, J., Augustyns, K., Lambeir, A. M., Van der Veken, P., Van Soom, J. & Magdolen, V. Novel KLK4 inhibitors. WO2015144933 A1 (2015).

-
56. Burchacka, E., Sienczyk, M., Frick, I. M., Wysocka, M., Lesner, A. & Oleksyszyn, J. Substrate profiling of *Fingoldia magna* SufA protease, inhibitor screening and application to prevent human fibrinogen degradation and bacteria growth *in vitro*. *Biochimie*. 103, 137-143 (2014).
 57. Boduszek, B. Synthesis of novel phosphonopeptides derived from pyridylmethylphosphonate diphenyl esters. *Phosphorus, Sulfur Silicon Relat. Elem.* 176 (1), 119-124 (2001).
 58. Sienczyk, M. & Oleksyszyn, J. A convenient synthesis of new α -aminoalkylphosphonates, aromatic analogues of arginine as inhibitors of trypsin-like enzymes. *Tetrahedron Lett.* 45 (39), 7251-7254 (2004).
 59. Lejczak, B., Kafarski, P., Soroka, M. & Mastalerz, P. Synthesis of the phosphonic acid analog of serine. *Synthesis*. 7, 577-580. (1984).
 60. Ali, O. M. Design and synthesis of small and potent inhibitors of urokinase as antitumor agents. *World J. Chem.* 7 (1), 01-06 (2012).
 61. Sienczyk, M., Lesner, A., Wysocka, M., Legowska, A., Pietruszewicz, E., Rolka, K., *et al.* New potent cathepsin G phosphonate inhibitors. *Bioorg. Med. Chem.* 16 (19), 8863-8867 (2008).
 62. Oleksyszyn, J., Marcinkowska, A., Sienczyk, M., Drąg-Zalesińska, M. & Wysocka, T. Application of Aromatic Amidines and Guanidines, Derivatives of Diphenyl Esters of 1-Aminoalkanephosphonic Acids for Induction of Apoptosis of Cancer Cells. PL 213133 (2013).
 63. Yang, D., Fan, L., Su, X., Wang, C., Li, H., Wang, L., *et al.* Preparation of L-(m-Aminophenyl)glycine and its Derivatives. CN 101633626 A (2010).
 64. Andrew, R. G., Barker, A. J., Boyle, F. T. & Wardleworth, J. M. Anti-Tumor Compounds. US5280027 (1994).
 65. Sass, P. & Bierbaum, G. Lytic activity of recombinant bacteriophage phi11 and phi12 endolysins on whole cells and biofilms of *Staphylococcus aureus*. *Appl Environ Microbiol.* 73 (1), 347-352 (2007).
 66. Zhang, J. H. & Chung, T. D. Y. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screening.* 4 (2), 67-73 (1999).
 67. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 2006.0008 (2006).

2.8 Supplementary data

2.8.1 Biological evaluation

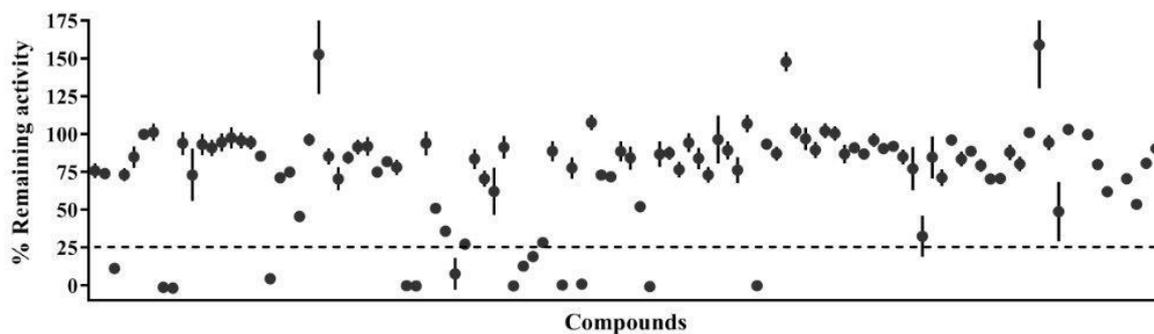


Figure S1. Overview of primary screening results. ClpP inhibition was investigated at 200 μ M compound concentration. Compounds resulting in $\leq 25\%$ remaining ClpP activity were selected for hit confirmation (dotted line). Each value represents the mean of three independent experiment \pm standard deviation.

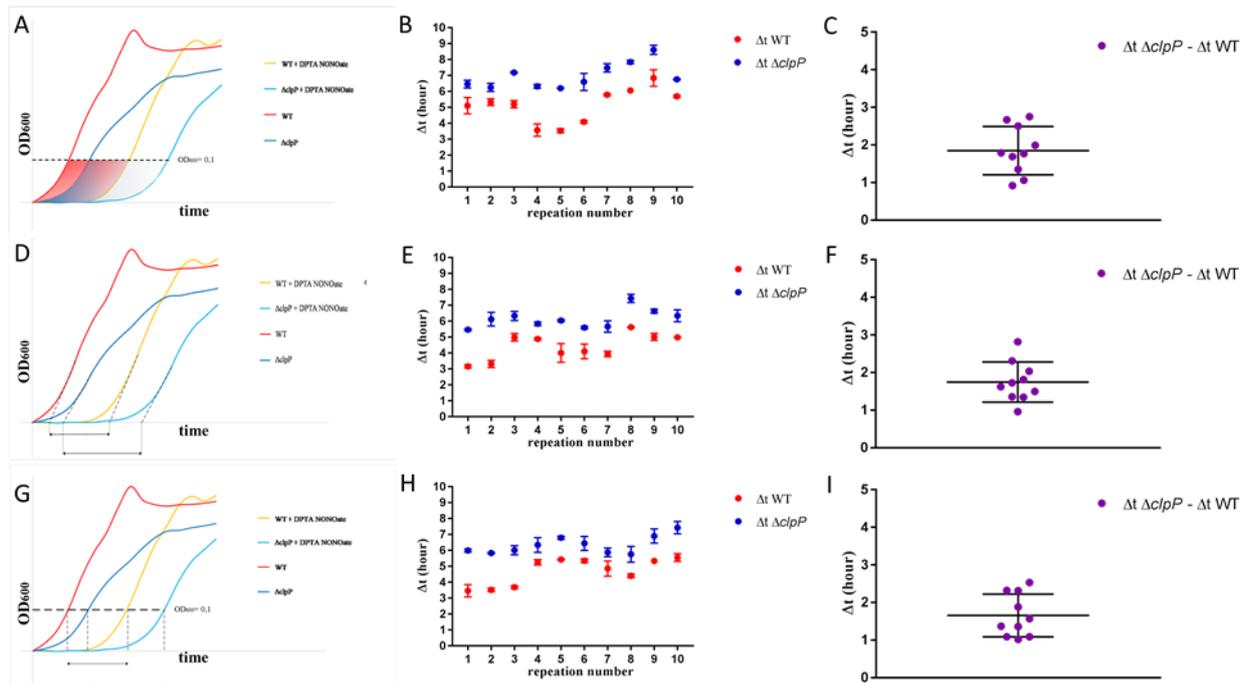


Figure S2. Analysis of the *E. coli* wild type (WT) and *E. coli* *clpP* defective strain ($\Delta clpP$) growth curve, in presence and absence of nitric oxide stress (WT, WT + DPTA NONOate, $\Delta clpP$, $\Delta clpP$ + DPTA NONOate), to evaluate the relative delay of the growth restart of the mutant compared to the wild type under stress condition in ten experiments with three replicates each. In this regard, three criteria have been defined. For each analysis, a bilateral t-test was carried out calculated (allowing different variances for the populations). Figure S2 A, D and G, graphically describe each criteria; Figure S2 B, E, H are the representation of a single experiments (in red the time for the WT, Δt WT, and in blue for the mutant, Δt $\Delta clpP$); Figure S2 C, F, I, show the re-growth time difference, Δt , between WT and $\Delta clpP$ (Δt $\Delta clpP$ - Δt WT).

Analysis 1 (A-C): an equivalent time of delay for each type of population (i.e. between wild type and mutant) was defined. The equivalent delay time is calculated from the definition of the area delimited by the 2 curves, stressed and not stressed, of each population (A: the red area between the red and the yellow curves describes the wild type area, the blue area between the dark blue and light blue curves, the mutant) and an arbitrary threshold, set at OD₆₀₀ = 0.1, representative of the early stage growth. The equivalent delay time of the strains was calculated by dividing the area values for the threshold.

Analysis 2 (D-F): the tangent (dotted line in Fig. S2 A) to each curve has been calculated and the point (in the time) of intersection between the tangent and the X-axis was used for the analysis. The tangent was defined within the interval where the curve has exponential growth with constant speed. The selection of the interval range has been chosen by applying a threshold selection algorithm to the derivate.

Analysis 3 (G-I); a bacterial growth threshold has been set at $OD_{600} = 0,1$ and, for this, the time at which each bacterium reaches this value of absorbance was evaluated. Piece-wise linear interpolation has been used.

For all the different analysis methods (Fig. S2 B, E, H) the difference between the time of growth of WT stressed and not stressed and the mutant stressed and not stressed are statistically different, for analysis B p-value $< 0,00005$, for analysis E p-value $< 0,00005$, for analysis H p-value $< 0,0005$). The delay in time for the stressed mutant growth, shown in Fig. S2 C, F, I, is consistently higher than 1 hour by all methods.

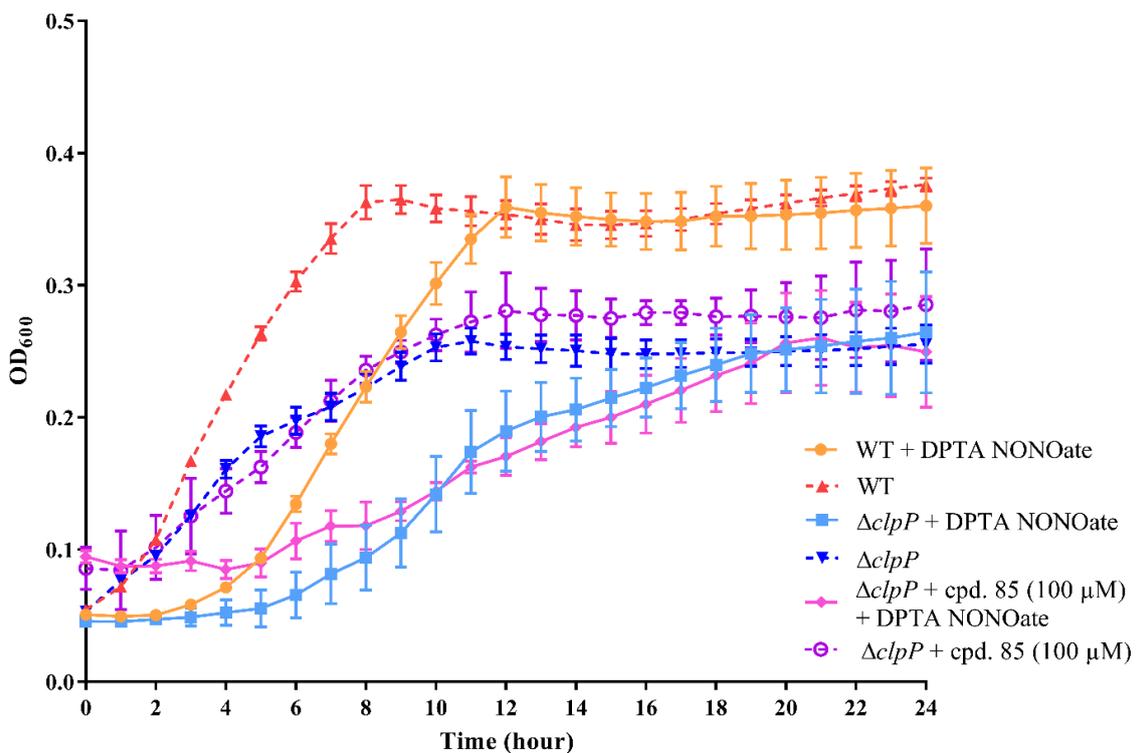


Figure S3. Control experiment to check the *clpP*-related activity of compound 85 suggested by nitric oxide stress test in *E. coli* wild type (WT) reported in Figure 5. Compound 85 was tested in *E. coli clpP*-defective strain ($\Delta clpP$) in presence ($\Delta clpP + cpd. 85 (100 \mu M) + DPTA NONOate$) or absence of nitric oxide stress ($\Delta clpP + cpd. 85 (100 \mu M)$), in order to unravel a potential off-target effect. The comparison between the growth curves of the mutant in presence of compound 85 and the controls curves (absence of compound 85, WT, WT + DPTA NONOate, $\Delta clpP$, $\Delta clpP + DPTA NONOate$) shown that the compound does not influence significantly the growth of *E. coli clpP* defective strain ($\Delta clpP$). The *clpP*-related effect of compound 85 has been confirmed. Each value represents the mean of three independent experiment \pm standard deviation.

Table S1. Summary of antibacterial activity (growth inhibition) of compounds. The fourteen hits compounds have been tested in all the summarized assays in this table. The fourteen hits compounds have been tested in all the summarized assays in this table at concentration of 100 μ M. A single bacteria colony, from a Mueller Hinton Agar (MHA) plate of the bacteria strain, was used to inoculate 5 mL sterile saline solution. Fresh and sterile Mueller Hilton Broth (MHB) media was used to adjust the CFU/mL of the bacteria suspension to 1×10^6 . The compounds were transferred in 96-well, flat bottom, sterile plates (screening format) in triplicate, together with the controls (DMSO and ciprofloxacin). As final step, the bacterial solution was added (final inoculum 5×10^5 CFU/mL). The final volume employed for the assay was 200 μ L. The plates were incubated at 37 °C in a plate shaker for 24 hours and the absorbance (600 nm) was measured at time 0, after 24 hours and different times, accordingly with each experiment set-up. For more information about the experimental procedure, see **Antibacterial assays** in the **EXPERIMENTAL SECTION**. Compounds have been considered active if showing a growth inhibition \geq of the 5 % of the the bacteria growth control. However, the minimum of inhibitions shown by the compounds tested in this work was 69.9 %.

Bacterial strain	Active compounds	Inhibition of growth (%)
<i>Acinetobacter baumannii</i> (ATCC 19606)	-	
<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> (ATCC 700603)	-	
<i>Pseudomonas aeruginosa</i> (ATCC 27853)	-	
<i>Staphylococcus aureus</i> (ATCC 29213)	115	99.9
<i>Escherichia coli</i> (BW25113)	-	
<i>Escherichia coli</i> Δ <i>tolC</i> (JW5503)	90, 107, 115	98.9, 69.9, 99.0
<i>Escherichia coli</i> Δ <i>lpxC</i> (JD17464)	-	
<i>Escherichia coli</i> Δ <i>clpP</i> (JW0427-1)	-	

<i>Escherichia coli</i> (BW25113) + PAβN	90	98.5
<i>Escherichia coli</i> Δ <i>clpP</i> (JW0427-1) + PAβN	90	99.4
<i>Escherichia coli</i> (BW25113) + DPTA NONOate	85	N.D.

2.8.3 Synthetic procedure for UAMC library compounds

All the chemistry procedure are reported in section **Appendix, Part 2. Synthetic procedure for UAMC library compounds – Chapter 2**

2.8.4 References

1. Joossens, J., Augustyns, K., Lambeir, A. M., Van der Veken, P., Van Soom, J. & Magdolen, V. Novel KLK4 inhibitors. WO2015144933 A1, (2015).
2. Augustyns, K., Van Der Veken, P., Messagie, J., Joossens, J. & Lambeir, A. M. Activity-based probes for the urokinase plasminogen activator. US 20140079632, (2014).
3. Jiang, J., Crabtree, R. H. & Brudvig, G. W. One-Step Trimethylstannylation of Benzyl and Alkyl Halides. *J Org Chem*, 81 (19), 9483-9488, (2016).
4. Sreenath, K., Allen, J. R., Davidson, M. W. & Zhu, L. A FRET-based indicator for imaging mitochondrial zinc ions. *Chem. Commun. (Cambridge, U. K.)*, 47 (42), 11730-11732, (2011).
5. Shen, Y.-M., Song, L.-L., Qian, X.-H. & Yang, Y. J. A scalable synthesis of 1-amino-5-cyanonaphthalene, a precursor for a nitric oxide probe (NO550) designed via the “dye assembly” principle. *Chin. Chem. Lett.*, 24 (1), 7-8, (2013).

Chapter 3

Identification and characterization of approved drugs and drug-like compounds as covalent *Escherichia coli* ClpP inhibitors

The content of this chapter is partially based on the manuscript “**Identification and characterization of approved drugs and drug-like compounds as covalent *Escherichia coli* ClpP inhibitors**”

Elisa Sassetti,^{1,3} Cristina Durante Cruz,² Päivi Tammela,² Mathias Winterhalter,³ Koen Augustyns⁴, Philip Gribbon,¹ Björn Windshügel,¹

In revision, *Frontiers in Microbiology*.

¹ *Fraunhofer Institute for Molecular Biology and Applied Ecology, ScreeningPort, Schnackenburgallee 114, 22525, Hamburg, Germany.*

² *Drug Research Program, Division of Pharmaceutical Biosciences, University of Helsinki, Viikinkaari 5E, FI-00014 Helsinki, Finland.*

³ *Department of Life Sciences and Chemistry, Jacobs University Bremen gGmbH, Campus Ring 1, 28759 Bremen, Germany.*

⁴ *Laboratory of Medicinal Chemistry, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium*

Contribution to the work. I have designed, performed (except one experiment and two repetition of bacterial growth assays, conducted by Dr. Cristina Durante Cruz) and analyzed personally all the biological experiments and the computational tests included in this part of the work. I also gave a substantial contribution in writing and reviewing the manuscript. The rest of the work has been performed by the co-author of the manuscript.

3.1 Abstract

The serine protease Caseinolytic protease subunit P (ClpP) is important for protein turnover and homeostasis in bacteria and has been shown to be involved in controlling virulence and biofilm formation. Therefore, ClpP represents an interesting potential target for therapeutic intervention in infections involving both Gram-positive and Gram-negative bacteria. In this study, we screened small molecule libraries of approved and investigational drugs for *Escherichia coli* ClpP inhibition, using a biochemical assay. Concentration-response analyses confirmed the approved drugs bortezomib, cefmetazole, cisplatin, the investigational drug cDPCP, as well as the protease inhibitor 3,4-dichloroisocoumarin to be ClpP inhibitors with IC₅₀ values ranging between 0.04 and 31 μ M. Of the five novel inhibitors, cefmetazole and cisplatin did not show any inhibition of the exemplary serine protease bovine α -chymotrypsin. Surface plasmon resonance studies revealed all compounds bind covalently to the ClpP enzyme. Cytotoxicity studies in selected human cell lines showed no toxicity of cefmetazole. Molecular docking calculations suggest a potential mode of binding for cefmetazole involving interactions with serine 97 and threonine 168 of ClpP. The antibacterial potential of the compounds was assessed using a panel of *E. coli* strains in the presence and absence of nitric-oxide induced stress.

3.2 Introduction

Caseinolytic protease subunit P (ClpP) is a well-conserved serine protease found in bacteria and many eukaryotes, although it is absent in archaea and mollicutes^{1,2}. Studies suggest that 80 % of the cellular proteolysis in bacteria is dependent upon the two proteases ClpP and Lon (ATP-dependent serine protease), thus suggesting a major role of ClpP in bacterial homeostasis and pathogenesis^{3,4}.

ClpP is a compartmentalized protease that is composed of two heptameric rings, assembled back-to-back, which form an interior proteolytic chamber with a narrow pore entrance⁵. A typical Ser-His-Asp catalytic triad is present in each subunit, located inside the inner chamber and thus protected from the cytoplasmic environment⁶. The peptidase activity of ClpP generates peptides with lengths of 7-8 residues⁷. However, in the presence of ATPases (AAA⁺) which are specific protein binding partners of ClpP and which are themselves associated with multiple cellular activities, the degree of ClpP protease activity is further enhanced^{8,9}. In the case of ClpP in *E. coli*, the relevant AAA⁺ proteins are ClpX and ClpA. These both form hexameric structures able to bind to the surface of tetradecameric ClpP¹⁰. These both form hexameric structures able to bind to the surface of tetradecameric ClpP¹⁰. The ClpP related role of these AAA⁺ proteins is to unfold and translocate proteins into the ClpP chamber, where they undergo proteolysis in an ATP-dependent manner⁹.

To date, about 70 proteins have been identified as substrates for *E. coli* ClpP-mediated proteolytic cleavage¹¹. In *E. coli* ClpP is involved in metabolism, damage repair, cell division, transcriptional regulation and stress response¹². More broadly, it is involved in oxidative stress response in *Streptococcus pneumoniae*¹³, virulence in *Staphylococcus aureus*^{14,15}, biofilm formation and in motility in *Pseudomonas aeruginosa*¹⁶ and has shown to be essential for macrophage infection in *Listeria monocytens*¹⁷. Due to its wide range of essential functions,

ClpP has been considered a highly promising antibiotic target and several studies have focused on identification of molecules which modulate its various activities³.

Several molecules have been described which inhibit ClpP activity by binding to its active site, including phenyl esters, boronates, AV145 and the covalently binding benzyloxycarbonyl-leucyltyrosine chloromethyl ketone (Z-LY-CMK)^{18, 19, 21}. The β -lactone compounds also showed activity against several pathogenic strains³. More recently, we report diarylphosphonates showing potent inhibition of the ClpP enzymatic activity in *E. coli* (see chapter 2)³⁸. In contrast, additional compounds activating ClpPs have been described. These Acyldepsipeptides (ADEPs) bind to the interface of the ClpP-ATPase protein-protein interaction site and lead to uncontrolled proteolysis²⁰.

In this study, we aimed to identify small molecule compounds or new warheads that inhibit *E. coli* ClpP in either a covalent or non-covalent manner. For this purpose, we used a high-throughput biochemical assay to screen three libraries of small molecules, containing approved and investigational drugs. Emerging hit molecules were then further characterised using biophysical and cell-based methods.

3.3 Results

The small molecule libraries SCREEN-WELL[®] FDA approved drug library version 2 (774 compounds.), LOPAC^{®1280} (1280 compounds) and a subset of the SPECS protease-focussed collection (329 compounds) were screened for *E. coli* ClpP inhibitors using Z-LY-CMK as positive and DMSO as negative controls. Plate level statistics were used to calculate the screen quality parameter value, Z' ²¹. The screen gave an overall $Z' > 0.6$ (Figure S1), with no evidence for spatial, intra- or inter-plate effects and a uniform spread of active compounds across the plates, indicating a valid screen. In the primary screen, a cut-off for the compound selection was set at 90 % relative inhibition of ClpP. Six compounds from the primary screen as hits, four from LOPAC^{®1280} and two from the SCREEN-WELL[®] FDA approved drug library V2 were selected and purchased as dry solids from commercial vendors (purity by LCMS was $> 90\%$) for further analysis.

Dose-response analysis of the identified ClpP inhibitors confirmed five of the selected compounds as ClpP inhibitors with IC_{50} values ranging from 0.04 μ M to 31.0 μ M (Figure 1, Table 1). Bortezomib proved the most potent enzyme inhibitor in this study, with a more than 100-fold lower IC_{50} value compared to cefmetazole and 3,4-dichloroisocoumarin (Table 1). Cisplatin and cDPCP showed lower potency and were less effective inhibitors than the positive control (Z-LY-CMK). Notably, none of the other 45 chemically related carbapenems and penems class antibiotics included in the LOPAC¹²⁸⁰, SCREENWELL[®] and matrix metalloproteinase (MMP) inhibitors subset, showed a comparable inhibition to cefmetazole (Table S2).

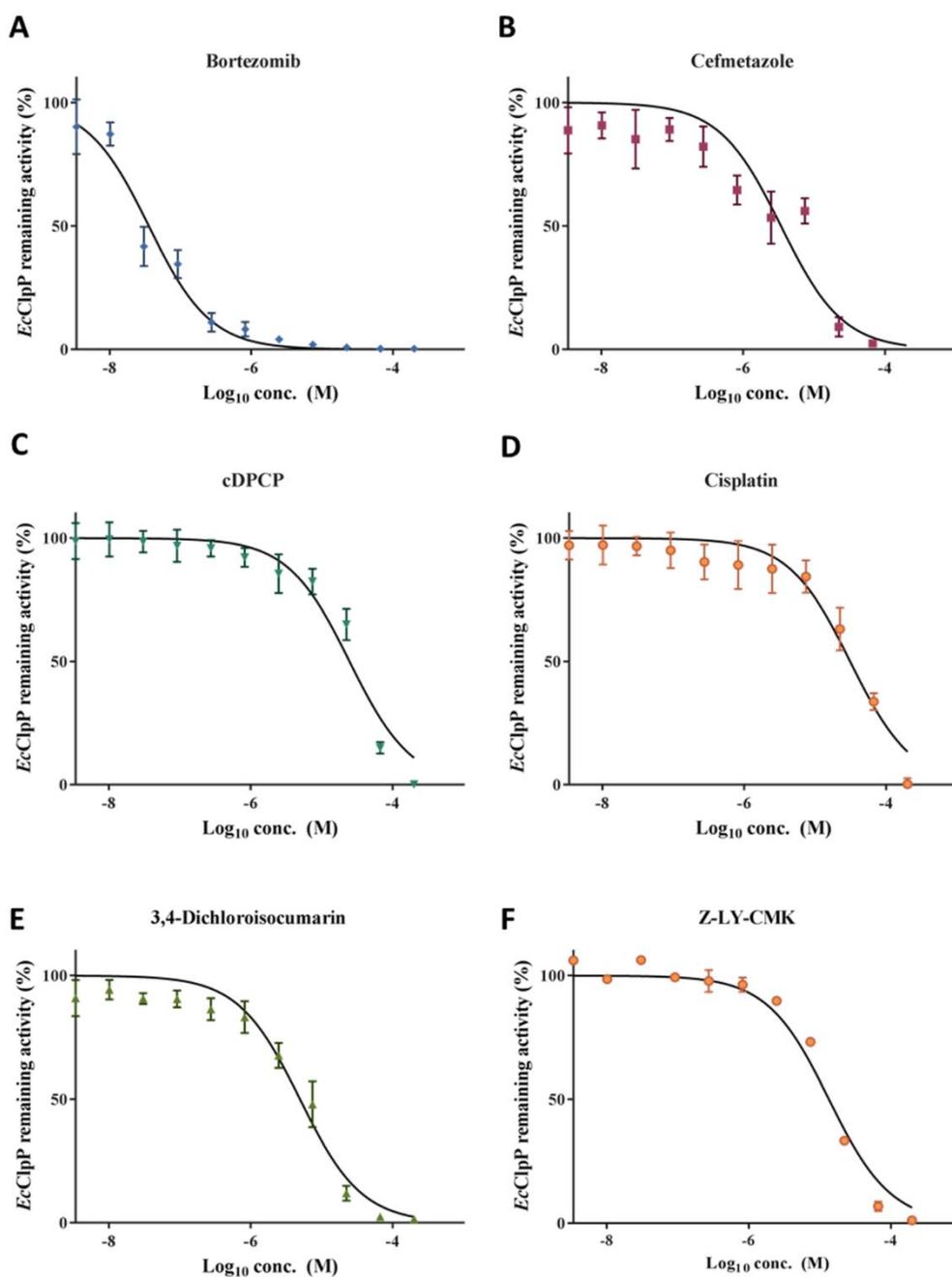


Figure 1. Concentration response curves for high-throughput screening hit compounds (A-E) and the known inhibitor Z-LY-CMK (F). All measurements were conducted in eleven points, 1:3 dilution starting from 200 μ M as top concentration. Error bars represent biological triplicates with three internal replicates each.

Table 1. Overview of IC₅₀ values of identified ClpP inhibitors and positive control (Z-LY-CMK). Comparison with inhibition of bovine α -chymotrypsin at 200 μ M compound concentration. The signal was normalized for the substrate background and compare to the maximum fluorescent level (signal generated from substrate and protein in absence of compounds).

	<i>E. coli</i> ClpP IC ₅₀ (μ M)	<i>E. coli</i> ClpP IC ₅₀ (μ M) confidence interval	% inhibition <i>E. coli</i> ClpP at 200 μ M	% inhibition chymotrypsin at 200 μ M
Bortezomib	0.04	0.02 to 0.05	100.1	97.8
Cefmetazole	3.4	1.5 to 7.6	94.5	0
cDPCP	24.6	17 to 36	91.2	84.7
Cisplatin	31.0	21.4 to 44.8	92.3	0
3,4-Dichloroisocoumarin	4.9	3.3 to 7.2	98.9	101.4
Z-LY-CMK	14.0	11.0 to 16.6	98.6	99.8

In order to investigate whether any of the five ClpP inhibitors antagonize other serine proteases, the inhibition of bovine α -chymotrypsin was investigated (Table 1). At 200 μ M compound concentration, bortezomib, 3,4-dichloroisocoumarin as well as the positive control Z-LY-CMK completely inhibited chymotrypsin and also cDPCP showed potent enzyme inhibition. In contrast, neither cefmetazole nor cisplatin inhibited chymotrypsin.

The cytotoxic potential of the ClpP inhibitors and Z-LY-CMK was tested in three human cell lines: HTC-116 (colon); MCF-7 (breast); and A549 (lung) in concentration response (top concentration 200 μ M, 1:3 dilution series). Bortezomib was highly toxic in all tested cell lines, even at the lowest concentrations (Figure 2). Relative to its ClpP inhibition (IC₅₀ = 25 μ M), cDPCP was toxic towards human cell (< 5 fold window). Cefmetazole did not show significant cytotoxicity against the human cell lines (>50 fold window). It is interesting to note that the cellular cytotoxicity of Z-LY-CMK, was second only to that of Bortezomib and was also significantly more than Z-LY-CMK's IC₅₀ against ClpP (Table 1 and Table 2), providing evidence that substantial off-target effects were occurring.

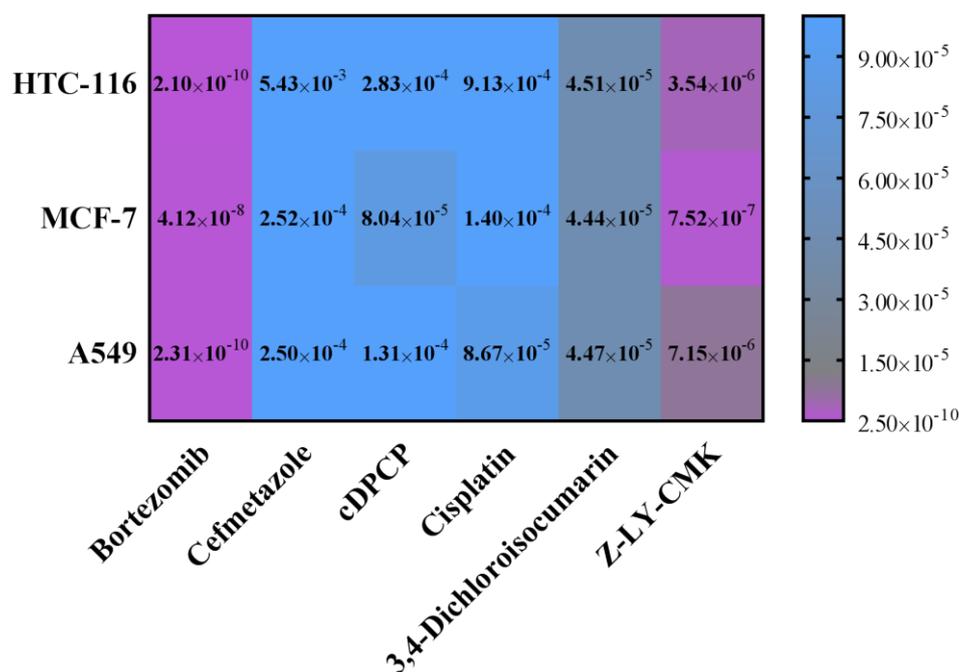


Figure 2. Overview of IC_{50} (M) values in cell viability tests using human cell lines HTC-116, MCF-7 and A549.

The binding properties of the compounds for the protein were evaluated by surface plasmon resonance (SPR). Upon immobilization on the sensor surface, test compounds were injected into the flow cell at 80 μ M concentration. The resulting sensorgrams were evaluated and compared with that of Z-LY-CMK, which binds *E. coli* ClpP covalently¹⁸. All tested compounds showed essentially irreversible binding mode, as SPR signals did not return to baseline after 30 minutes of dissociation time, showing similar behavior to Z-LY-CMK (Figure3). Caffeine was used as a negative non-binding control (FigureS2 supplementary information). The high RU values obtained for cisplatin and cDPCP indicates unspecific binding of these compounds to the protein (Figure 3B&C). Also, a low level of absolute RU change is seen in SPR upon bortezomib binding, even though this compound exhibits nanomolar IC_{50} 's in the enzymatic assay. This effect could be due to partial dissociation effects or induction of intramolecular conformational changes upon binding, (Figure 3A).

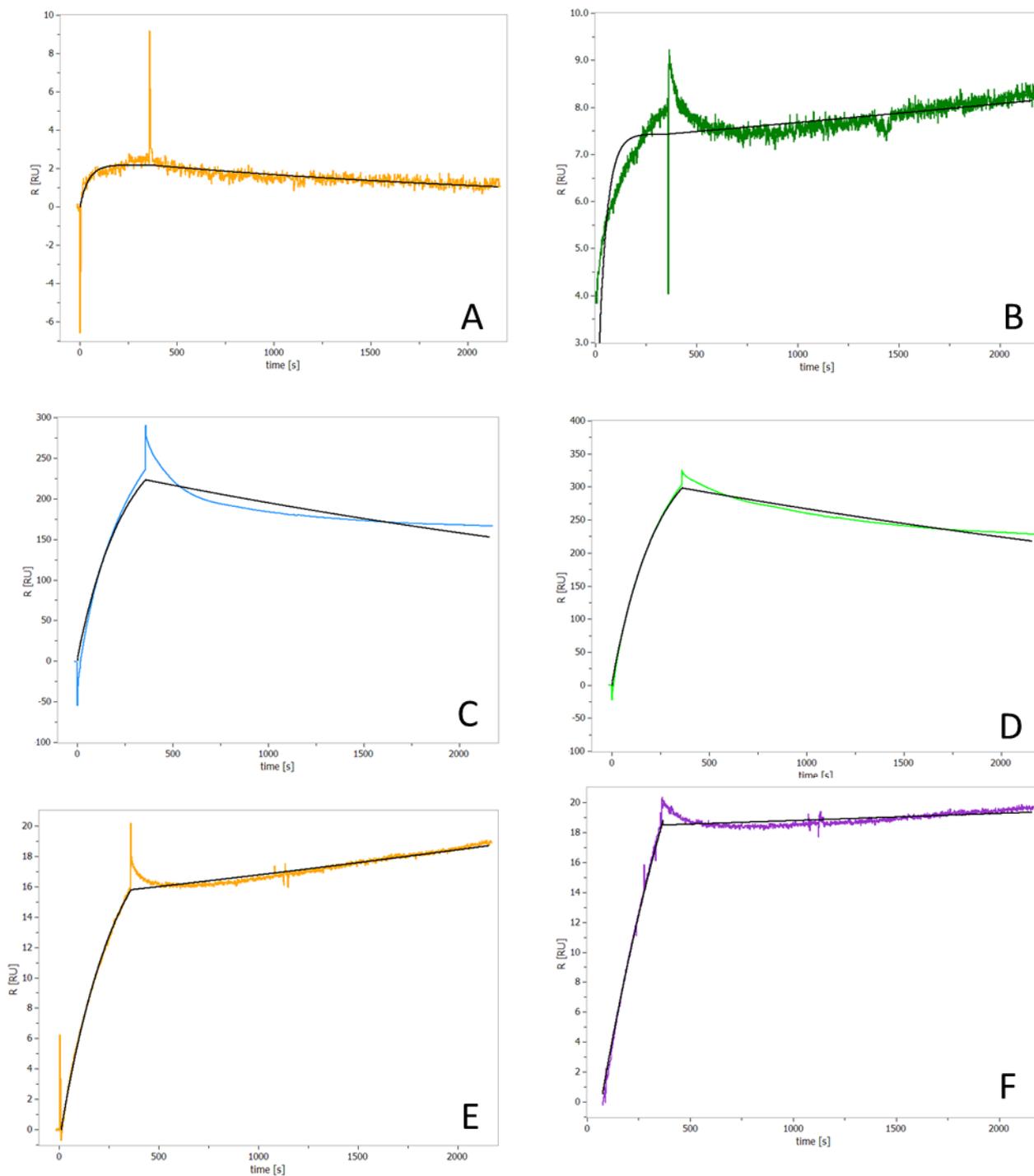


Figure 3. SPR sensorgrams of the ClpP inhibitors at a concentration of 80 μM . Each compound was injected over 7 seconds, followed by 1,800 seconds of dissociation time. A. Bortezomib, B. Cefmetazole, C. cDPCP, D. Cisplatin, E. 3,4-Dichloroisocoumarin, F. Z-LY-CMK (covalent binder, control).

In order to investigate the potential binding mode of cefmetazole in ClpP, the compounds were docked into the active site of the protein structure. Due to the covalent binding mode, indicated by the SPR measurements, cefmetazole was docked covalently with the β -lactam ring in the open form to serine 97 (Figure 4) which has been shown to irreversibly interact with the known inhibitor Z-LY-CMK¹⁸.

In order to evaluate the potential binding mode of cefmetazole in the active site and to investigate its affinity for ClpP relative to other members of this class of antibiotics, cefmetazole was docked into the X-ray crystal structure of *E. coli* ClpP (PDB ID 2FZS). Based on the results from surface plasmon resonance studies, the compound was docked covalently, in its post-reactive form, to serine 97, (which also shows covalent attachment to the co-crystallised Z-LY-CMK¹⁸).

The hypothetical binding modes into X-ray crystal structure of *E. coli* ClpP of cefmetazole were predicted using GOLD. The compound was docked covalently to serine 97, with the most frequent penem antibiotic class binding mode.^{33, 22} In our model (see Figure 4), the serine 97 interacts with the carboxyl group created upon the β -lactam ring opening. Due to the position of the nitrile group deep in the pocket, we speculate that, after the activation of the group by the asparagine150, a covalent attach with threonine 168 can occur (unless the nitrile group is already degraded). The reactivity of nitrile groups with protein residues that are known to interact covalently has been reported by Magnin et al²³.

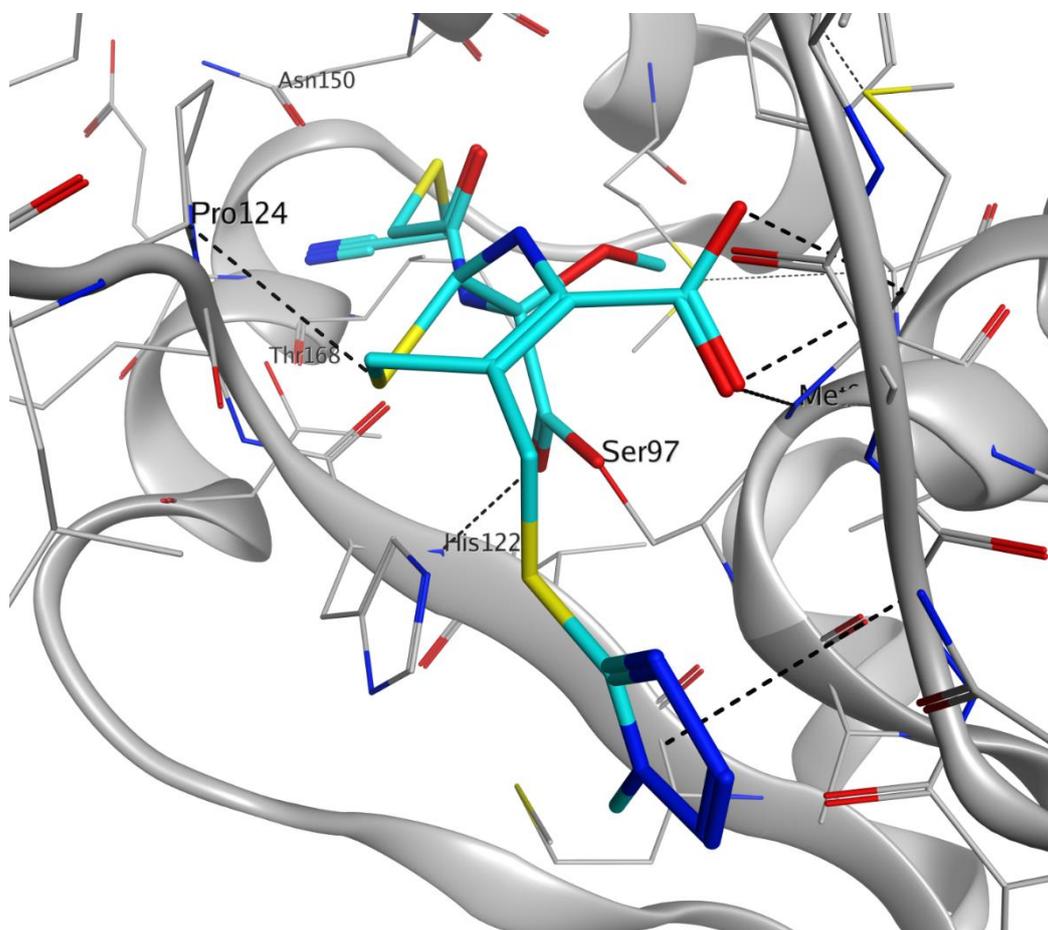


Figure 4. Binding mode of top-ranked docking pose of covalently docked cefmetazole (carbon atoms in cyan) in PDB ID 2FZS using GOLD with the compound attached to atom OG of serine 97. Hydrogen bonds are indicated as black dashed lines.

Compounds were tested for their intrinsic antimicrobial activity, firstly against *Escherichia coli* BW25-113, wild type strain, at a concentration of 100 μM in rich medium (MHB). In this assay, the only cefmetazole showed inhibition of bacteria growth at this concentration, while all remaining compounds were inactive. That is not surprising since cefmetazole is a marketed antibacterial.

The remaining compounds were then tested in *Escherichia coli* JW5503 (ΔtolC), efflux pump deficient strain, and none of them showed improvements in their activities against bacteria. Further activity runs were carried out in *Escherichia coli* BW25-113 and the isogenic ClpP-defective strain *Escherichia coli* JW0427 (ΔclpP) in the presence of the known efflux pump substrate phenylalanine-arginine beta-naphthylamide (Pa β N, 25 μM). With the efflux pump inactivated Bortezomib gave 55% reduction on growth in the WT strain and, interestingly, almost zero growth (100% of reduction) in the *clpP* deficient mutant.

Recently, it has been shown that ClpP-deficient *E. coli* shows an impaired nitric oxide detoxification capacity compared to wild type *E. coli* after nitric oxide stress induction using DTPA NONOate²⁴. To investigate compound effects on the capacity of bacteria to undergo NO \bullet detoxification, compounds were tested (at 50 and 100 μM) in *Escherichia coli* BW25-113 (WT strain) in minimal media M9, in the presence and absence of nitric oxide, using the isogenic mutant *Escherichia coli* JW0427 (ΔclpP) as the positive control. Compounds cDPCP, cisplatin and 3,4-dichloroisocoumarin activity are compared to positive (ΔclpP *E. coli* under nitric oxide stress) and negative controls (*E. coli* wild type under nitric oxide stress and ΔclpP and wild type *E. coli* without NO \bullet stress) seen Figure 5.

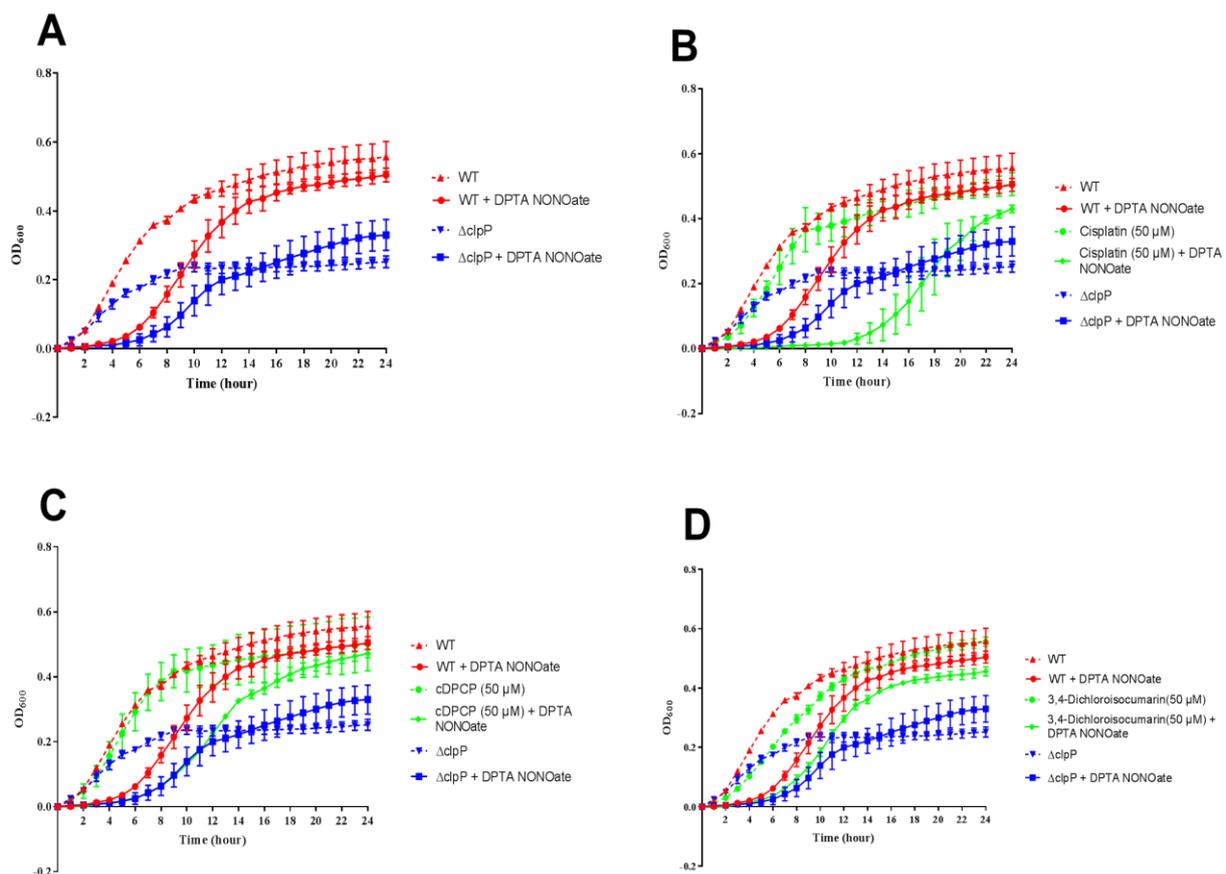


Figure 5. Bacterial growth assays for WT and $\Delta clpP$ *E. coli* strains in minimal media, in the presence (solid lines) and absence (dotted lines) of DTPA NONOate ($\text{NO}\cdot$) induced stress for 3 selected compounds. A) WT and $\Delta clpP$ alone. B) Cisplatin, BC) cDPCP and D) 3,4-dichloroisocoumarin. Compounds at 50 μM .

The growth recovery time of the $\Delta clpP$ *E. coli* strain is delayed by one hour compared to the isogenic WT, see Figure 5A, blue solid curve (mutant growth curve) versus red solid line (WT growth curve). This observation has been reported previously from Robinson *et al.*²⁴, but with a longer delay of approximately 4-5 hours was seen, (for more details see S2 Chapter 2 of this thesis). Another effect that must be noted here is that the mutant, either in presence or absence of nitric oxide stress, at 24 hours never reaches the same optical density level of the WT.

The evaluation of the effect of cisplatin in WT under stress condition is shown in Figure 5B. At 50 μM cisplatin concentration, growth in absence of NO stress (dotted green line) is comparable with the WT type alone (dotted red line). However, in presence of stress, the growth of WT plus cisplatin (solid green line) is retarded compared to both stressed WT (solid red line) and stressed $\Delta clpP$ mutant (solid blue line), with no compound. At 100 μM cisplatin bacterial growth (see Figure 3 B, 3.8 Supplementary data) is affected by the compound concentration already in absence of nitric oxide stress (dotted pink line), probably due to compound toxicity. Interestingly, the growth of WT in presence of 50 μM cisplatin in media (MHB) at 24 hours is not affected

by the compound (data reported in this section, see Table 2): this effect is apparently correlated with low availability of resources due to minimal media culture conditions.

For cDPDC (see figure 5 C and S3 C) at 50 μM , the compound does not influence bacterial growth curves. In the presence of stress cDPDC treated WT cells recover later than the stressed WT cells without compound, resembling the stressed mutant cells lacking ClpP (Figure 5C). In the experiments at 100 μM concentration (see Figure 3 B, 3.8 Supplementary data) WT cell growth is clearly influenced by the presence of the compound itself (dotted pink line). Interestingly, cDPDC does not seem to interfere under richer media conditions (MHB).

Similar behavior has been seen in 3,4-Dichloroisocoumarin (see Figure 5 D and S3 D) at 50 μM and 100 μM . A reduced effect can be seen in the nonstressed growth curves of WT in presence of compound (dotted green lines) and a delay is shown in the curves with 3,4-Dichloroisocoumarin and $\text{NO}\bullet$ (solid green and pink lines) compare to the stressed WT (solid red line). The curve of 100 μM (solid pink lines, Figure S3 D) has a greater effect and is closer to the mutant under stress conditions (solid green lines, Figure 5 D).

At 24h all the three above mentioned compounds (except cisplatin at 100 μM) under $\text{NO}\bullet$ stress reach the OD_{600} level of the wt, exceeding the growth levels of the mutant.

3.4 Discussion

Based on a biochemical high-throughput screen we identified and profiled five structurally distinct small molecules as inhibitors of *E. coli* ClpP which bind in an irreversible manner. Among them, 3 anticancer drugs, one proteasome inhibitor and one antibiotic, known to be active against Gram-positive and Gram-negative bacteria via a different mechanism.

Two of the three anti-cancer drugs, cisplatin and cDPCP are metal-containing compounds used in anticancer therapy³⁷. Cisplatin treats many types of cancers, including carcinomas, germ cell tumors, lymphomas, and sarcomas²⁵ and it is also known to interact with serine residues and to induce their phosphorylation in p53, for instance²⁶. The difference major difference between the two chemotherapeutic agents is in their mode of action: while cisplatin cross-links the DNA, cDPCP makes a single mono-functional adduct with DNA due to singular chlorine that can act as leaving group³⁷. The observed cytotoxicity is not surprising since the main activity of the drugs but must be underlined that both do not have activity against a classical serine protease, like chymotrypsin, at the relatively high-test concentration of 200 μM used in this study. Both cisplatin and cDPCP showed antibacterial activity, in a possibly *clpP*-mediated manner, in the presence of nitric oxide stress (Figure 5 B&C and S3 B&C).

The third anticancer drug, bortezomib, is used in mantle cell lymphoma therapy. In mammalian cells it acts a reversible inhibitor of the proteolytic activity of the proteasome complex²⁷ and has been previously described as ClpP inhibitor in a “Mechanism-Based Whole-Cell” screen directed against *Mycobacterium tuberculosis*²⁸.

Bortezomib is a dipeptidyl boronic acid which is a modified derivative of leucine and phenylalanine²⁹, similar to benzyloxycarbonyl-leucyltyrosine chloromethyl ketone, another ClpP known inhibitor¹⁸. In this study, we showed that bortezomib potently inhibits ClpP with an *in vitro* activity in the nanomolar range, although it is highly cytotoxic. Interestingly, in the experiments in *E. coli* with the addition of the efflux pump substrate PaβN, bortezomib behaves differently between the wt strain and in $\Delta clpP$: the growth of the $\Delta clpP$ mutant is inhibited at 100% but the WT cell line growth is just reduced to only 55 %.

3,4-dichloroisocoumarin (3,4-DCI), is a general serine protease inhibitor acting against elastase-like, chymotrypsin-like and trypsin-like proteases. It inactivates these proteases through acylation of catalytic serines³⁰. Not surprisingly, this compound was identified as a hit in the screen and was active but not selective at high concentration. Also, for this compound, we observe no evidence for nitric oxide-*clpP* mediated effects (Figure 5 D and S3 D).

The last identified drug is cefmetazole, a semisynthetic cephamycin antibiotics that interfere of transpeptidation and cross-linking during cell wall synthesis^{22,31}. In this study, we identified a new intracellular target for cefmetazole. The enzyme inhibition assay revealed cefmetazole has a relatively low IC₅₀ *in vitro* on *E. coli* ClpP and it is not active against chymotrypsin. The screened libraries in this study contained other 45 cepheems and penem compounds (Table S2 supplementary information). However, none of these compounds showed activity in the primary screen above >70 % indicating some degree of intra-compound-class specificity of this particular compound for ClpP. Cefmetazole, carries a nitrile moiety, not present in all other tested compounds. Nitrile moieties are known to be reactive species in protease inhibitors field (see Magnin et al.²³) and they are present in commercialized antidiabetic drugs. In Cefmetazole two moieties potentially reactive towards serine 97 are present: the penem ring and the nitrile group. Covalent binding docking experiments revealed the possibility of a double binding of Cefmetazole (at serine 97 with penem and at threonine 68) to ClpP: this possibility could explain the high affinity of the compound for *E. coli* ClpP compared to those of the other penems and carbapenems screened. This hypothesis, if confirmed by structural studies, could open new routes for structure-based design of novel “broad spectrum ClpP inhibitors” as antibiotics.

3.5 Materials and methods

Protein production. Plasmid pETDclpPec (ORF ECK0431)³² was cloned into the *E. coli* SG1146a strain. After overnight incubation, bacteria were diluted in fresh LB medium containing 100 µg/ml ampicillin and were grown until an OD₆₀₀ of 0.6 was reached. Isopropyl β-D-1-thiogalactopyranoside (1 mM) was added for induction and flasks were placed in an orbital shaker for 5 hours at 30 °C and 180 rpm. Cells were lysed on ice with a homogenizer (Precellys Evolution, Bertin Technologies, France). Protein was purified by IMAC bench method with a Ni-NTA His-tag resin column (Sigma Aldrich, USA) using buffers with increasing concentration of imidazole (Buffer A: pH 7.6, 50 mM Tris-HCl buffer, 150 mM NaCl, 10 mM imidazole.

Buffer B: pH 7.6, 50 mM Tris-HCl buffer, 150 mM NaCl, 20 mM imidazole. Buffer C: pH 7.6, 50 mM Tris-HCl buffer, 150 mM NaCl, 500 mM imidazole).

Assay Development and Screening. The proteolytic activity of ClpP in presence of compounds was assessed with a fluorescence-based assay in 384 plate format (well black, flat-bottom microtiter plates, #3820, Corning Inc., Corning, USA). The assay was carried out at a concentration of 75 μ M of fluorogenic substrate Suc-LY-AMC (Enzo Life Sciences, Germany) and 25 ng/ μ l of *E. coli* ClpP in the assay buffer (100 mM NaCl, 100 mM Hepes pH 7.5, 0.05 % Brij[®] 35). Protein and compounds were incubated for 10 minutes at 30 °C before starting the reaction with the addition of the substrate. Changes in fluorescence (excitation 350 nm, emission 435 nm) were monitored at 1 minute intervals over one hour using a microplate reader (Infinite M1000, Tecan, Switzerland) at 30 °C. The assay was carried out using the Fluent[™] 780 laboratory automation system (Tecan) and an acoustic liquid handling (Echo 550, Labcyte, USA) for compound transfer. Compound stocks were dissolved at 10 mM concentration in 99.8 % DMSO (ROTIPURAN[®] CAS No.67-68-5, Carl Roth GmbH, Germany) and stored at -20 °C until usage. The primary screen was carried out at a compound concentration of 200 μ M. The concentration-response analysis was carried out using dry-stock compounds dissolved in DMSO. The maximum fluorescence level was defined in wells where the compounds were absent and in presence of DMSO. The substrate back-ground signal was calculated in as well in absence of compounds and of protein. Z-LY-CMK was used as a positive inhibition control.

Compound selectivity was assessed by using bovine α -chymotrypsin (#C4129, Sigma-Aldrich, USA) at 200 μ M concentration. The assay was carried out as described before but with 1 ng/ μ l of α -chymotrypsin and 100 μ M of Suc-LY-AMC substrate dissolved in an assay buffer containing 150 mM NaCl, 10 mM CaCl₂, 50 mM Tris-HCl, and 0.05 % Brij[®] 35.

The SCREEN-WELL[®] FDA approved drug library V2 was purchased from Enzo Life Sciences (Lörrach, Germany), and the LOPAC^{®1280} library was obtained from Sigma-Aldrich (Taufkirchen, Germany). Moreover, from the Specs library (Specs, Netherlands), a subset of matrix metalloproteinase (MMP) inhibitors was selected using training data from the ChEMBL database³³. The Specs library was filtered on commonly applied properties and the Specs screening guidelines to give a selection of 329 compounds.

The initial concentration for all the compounds was 10 mM dissolved in DMSO 99.9%.

Hits emerging from the primary screen were re-purchased as dry-stocks. 3,4-dichloroisocoumarin (#D7910), cDPCP (#C0996) and Cisplatin (#P4394) were obtained from Sigma Aldrich (USA), Cefmetazole sodium (#KS-1415) and Guanabenz acetate (#KS-1362) from Key Organics (UK), Bortezomib (#MolPort-006-393-007) from MolPort (Latvia).

Concentration response experiments were done at least 3 times independently, each in technical triplicates.

Surface Plasmon Resonance Spectroscopy. *E. coli* ClpP (80 µg/ml) protein was immobilized, using amino-based coupling (SPR-AS-HCA-03, Bruker Daltonics Inc., USA) according to the manufacturer's recommendations, onto the surface of the sensor chips using a SIERRA SPR-16 instrument (Bruker Daltonics Inc., USA). Measurements were conducted in acetate pH 4 with a buffer containing 150 mM NaCl, 10 mM Hepes, 3 mM EDTA and 0.05 % (v/v), Tween-20 and a flow rate for protein immobilization of 10 µL/min. Binding assays were carried out using the same buffer with the addition of 0.8 % (v/v) of DMSO and 80 µM test substance (flow rate of 20 µL/min for 7 min and up to 1,800 seconds of dissociation time). As a positive control for irreversible binding, Z-LY-CMK was used.

The SPR data were analysed with the software Analyzer 3 (Bruker Daltonics, USA), which included algorithms for subtracting reference measurement and accounting for non-specific interactions of the solvent with the sensor chip surface.

Cell viability assays. Potential cell toxicity of the tested compounds was evaluated by ATP quantification using the CellTiter-Glo[®] viability assay kit (Promega). A549, HTC-116, and MCF-7 cells were seeded at 2,000 cells/well (20 µl) in white, flat bottom, sterile 96-well plates (# 781073, Greiner Bio-One, Germany) and incubated for 24 h. 200 nL ClpP inhibitors plus positive (valinomycin, 10 µM) and negative (DMSO, 1 %) controls were transferred in the plate, using an Echo[®] 550 liquid handler (Labcyte, USA), in dose-response at 1:3 dilutions starting from 100 µM and incubated for an additional 48 h (95 % air, 5 % CO₂ at 37° C). After incubation, 20 µl of CellTiter-Glo reagent was added into the culture and incubated at room temperature in the dark, for a further 30 minutes. Luminescence was quantified by an EnVision plate reader (PerkinElmer, Germany)

Antibacterial assays. The capacity of hit compounds to interfere with the bacteria growth was evaluated in *E. coli* BW25-113 (wild type) and *tolC* deficient strain *E. coli* JW5503. Single colonies of the bacteria strains from Mueller Hinton Agar (MHA) culture, were used to inoculate 5 ml of saline solution. The suspension was transferred into fresh sterile Mueller Hinton Broth (MHB) media to give a final concentration of 10⁶ CFU/ml. Growth assays used 96-well, flat bottom, sterile clear plates (#167008, Nunc, VWR, USA) and compounds were tested at 50 and 100 µM concentration (2 % DMSO). Absorbance measurements (600 nm) were made with the Multiskan GO or Varioskan LUX plate reader (Thermo Fisher Scientific, USA) from 0 to 24h at hourly intervals. Between readings, plates were stored in shaking incubator (500 rpm) at 37 °C.

The effect of the compounds on wild type strain *E.coli* BW25-113 and the isogenic *clpP*-defective *E.coli* JW0427-1³⁴ were measured in presence of an efflux pump substrate Phe-Arg β-naphthylamide dihydrochloride (#P4157, Sigma-Aldrich, USA) at 25 µM.

Combination tests in presence of nitric oxide were carried out as described previously (Chapter 2). Briefly, nitric oxide stress was induced by adding 2 mM of DTPA NONOate ((Z)-1-[N-(3-aminopropyl)-N-(3-ammoniopropyl)amino]diazene-1,1,2-diolate, Cayman Chemical, Ann Arbor, MI, USA), dissolved in

NaOH (final concentration 0.14 mM) in a *E. coli* BW25-113 culture at OD₆₀₀=0.1 in M9 media supplemented with 10 mM glucose. *E. coli* JW0427-1 culture was used as growth control to observe no ClpP mediated effects of the compounds. The bacterial cultures were grown in MHB overnight, shaking at 250 rpm at 37 °C. On the day of the experiment, fresh M9 media plus 10 mM glucose was inoculated 1:100 with the overnight culture and grown until approximately OD₆₀₀=0.3, and then used for the assay.

Data Analysis. Calculation of Z prime (Z') for HTS assay performance was performed according to ³⁴ and plates were considered valid for further analyses if Z' > 0.5 ³⁵. IC₅₀ values were determined by a non-linear fit of dose response using the equation for sigmoidal dose response with variable slope (GraphPad Prism 7.03, GraphPad Software, La Jolla, CA, USA)

Molecular Docking. The ClpP X-ray crystal structure was obtained from the Protein Data Bank (PDB) ID 2ZFS. Of the 14 chains, chain B was selected based on a stereochemical investigation of all chains using PROCHECK^{36, 37} and ProSa³⁵. Protein structures were prepared using Molecular Operating Environment (MOE, Chemical Computing Group Inc., Montreal, Canada) version 2016.0802. Three-dimensional coordinates of cefmetazole were generated within MOE. Molecular docking and re-docking were carried out using GOLD version 5.6.1 (Cambridge Crystallographic Data Centre, Cambridge, UK) and the GoldScore scoring function. The scoring function was selected based on the re-docking results of the co-crystallised covalently bound ligand Z-LY-CMK into 2FZS using all four implemented scoring functions. This resulted in lowest deviations of the top-ranked docking pose from the experimentally determined Z-LY-CMK coordinates for GOLD in combination with GoldScore ³⁶. The search space was defined as a 15 Å diameter sphere centered on atom OG of serine 97. Altogether, 50 docking runs were conducted. The early termination option was disabled and the asparagine 150 side chain was defined as flexible.

3.6 Abbreviation list

3,4-DCI	3,4-dichloroisocoumarin
AAA+	ATPases Associated with diverse cellular Activities
ADEPs	Acyldepsipeptides
ClpA	Caseinolytic protease subunit A
ClpP	Caseinolytic protease proteolytic subunit
ClpX	Caseinolytic protease subunit X
DMSO	dimethyl sulfoxide
DPTA NONOate	(Z)-1-[N-(3-Aminopropyl)-N-(3-ammoniopropyl)amino]diazene-1,1,2-diolate
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
FDA	Food and Drug Administration
His-Ser-Asp	histidine-serine-aspartic acid

HTS	High-throughput screening
IC ₅₀	Concentration of a drug that is required for 50% inhibition <i>in vitro</i>
IMAC	Immobilized metal ion affinity chromatography
LB	Luria-Bertani
LCMS	Liquid chromatography–mass spectrometry
MCF-7	Michigan Cancer Foundation-7 human cell line
MHA	Müller-Hinton agar
MHB	Müller-Hinton broth
MOE	Molecular Operating Environment
MW	Molecular weight
Ni-NTA	Nickel-charged affinity
NO•	Nitric oxide
PaβN	Phenylalanine-arginine beta-naphthylamide
PDB	Protein data bank
SMILES	Simplified Molecular Input Line Entry System
SPR	Surface Plasmon Resonance Spectroscopy
Suc-LY-AMC	4-(((S)-1-(((S)-2-(4-Hydroxyphenyl)-1-(4-methyl-2-oxo-2H-chromen-7-yl)ethyl)amino-4-methyl-1-oxopentan-2-yl)amino)-4-oxobutanoic acid
WT	Wild type
Z'	Z prime
Z-LY-CMK	Benzyl ((S)-1-(((S)-4-chloro-1-(4-hydroxyphenyl)-3-oxobutan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate
Δ <i>clpP</i>	Mutant defective in <i>clpP</i>
Δ <i>tolC</i>	Mutant defective in <i>tolC</i>

3.7 References

1. Wong, P. & Houry, W. A. Chaperone networks in bacteria: Analysis of protein homeostasis in minimal cells. in *Journal of Structural Biology*. 146 (1-2), 79-89 (2004).
2. Yu, A. Y. H. & Houry, W. A. ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Letters*. 581 (19), 3749-57 (2007).
3. Bhandari, V., Wong, K. S., Zhou, J. L., Mabanglo, M. F., Batey, R. A. & Houry, W. A. The Role of ClpP Protease in Bacterial Pathogenesis and Human Diseases. *ACS Chem. Biol.* 13 (6), 1413–1425 (2018).

4. Goldberg, A. L., Moerschell, R. P., Hachung, C. & Maurizi, M. R. ATP-dependent protease La (Lon) from *Escherichia coli*. *Methods Enzymol.* 244, 350–375 (1994).
5. Brotz-Oesterhelt, H. & Sass, P. Bacterial caseinolytic proteases as novel targets for antibacterial treatment. *Int. J. Med. Microbiol.* 304 (1), 23-30 (2014).
6. Alexopoulos, J. A., Guarne, A. & Ortega, J. ClpP: a structurally dynamic protease regulated by AAA+ proteins. *J. Struct. Biol.* 179 (2), 202-210 (2012).
7. Schelin, J., Lindmark, F. & Clarke, A. K. The ClpP multigene family for the ATP-dependent Clp protease in the cyanobacterium *Synechococcus*. *Microbiology.* 148, 2255–2265 (2002).
8. Olivares, A. O., Baker, T. A. & Sauer, R. T. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat. Rev. Microbiol.* 14 (1), 33-44 (2016).
9. Him, D., Chung, Y. S., Gloyd, M., Joseph, E., Ghirlando, R., Wright, G. D., *et al.* Acyldepsipeptide Antibiotics Induces the Formation of a Structured Axial Channel in ClpP: a Model for the ClpX/ClpA Bound State of ClpP. *Chem Biol.* 17 (9), 959–969 (2011).
10. Beuron, F., Maurizi, M. R., Belnap, D. M., Kocsis, E., Booy, F. P., Kessel, M., *et al.* At Sixes and Sevens: Characterization of the Symmetry Mismatch of the ClpAP Chaperone-Assisted Protease. *J. Struct. Biol.* 123 (3) , 248–259 (1998).
11. Flynn, J. M., Neher, S. B., Kim, Y.-I., Sauer, R. T. & Baker, T. A. Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell.* 11, 671–683 (2003).
12. Robinson, J. L. & Brynildsen, M. P. An ensemble-guided approach identifies ClpP as a major regulator of transcript levels in nitric oxide-stressed *Escherichia coli*. *Metab. Eng.* 31, 22-34 (2015).
13. Park, C. Y., Kim, E. H., Choi, S. Y., Tran, T. D., Kim, I. H., Kim, S. N., *et al.* Virulence attenuation of *Streptococcus pneumoniae* clpP mutant by sensitivity to oxidative stress in macrophages via an NO-mediated pathway. *J. Microbiol.* 48 (2), 229-235 (2010).
14. Böttcher, T. & Sieber, S. A. beta-Lactones as privileged structures for the active-site labeling of versatile bacterial. *Angew. Chem., Int. Ed.* 47 (24), 4600-4603 (2008).
15. Conlon, B. P., Nakayasu, E. S., Fleck, L. E., LaFleur, M. D., Isabella, V. M., Coleman, K., *et al.* Activated ClpP kills persisters and eradicates a chronic biofilm infection. *Nature.* 503 (7476), 365-70 (2013).

16. Fernández, L., Breidenstein, E. B. M., Song, D. & Hancock, R. E. W. Role of Intracellular Proteases in the Antibiotic Resistance, Motility, and Biofilm Formation of *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother.* 56(2), 1128–1132 (2012).
17. Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. & Berche, P. The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol. Microbiol.* 35 (6), 1286–1294 (2000).
18. Szyk, A. & Maurizi, M. R. Crystal structure at 1.9Å of *E. coli* ClpP with a peptide covalently bound at the active site. *J. Struct. Biol.* 156 (1), 165–174 (2006).
19. Ye, F., Li, J. & Yang, C. G. The development of small-molecule modulators for ClpP protease activity. *Mol. Biosyst.* 13 (1), 23–31 (2017).
20. Brotz-Oesterhelt, H. & Sass, P. Bacterial caseinolytic proteases as novel targets for antibacterial treatment. *Int. J. Med. Microbiol.* 304 (1), 23–30 (2014).
21. Zhang, J. H. & Chung, T. D. Y. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screening.* 4 (2), 67–73 (1999).
22. Utsui, Y., Ohya, S., Takenouchi, Y., Masazo, T. & Sugawara, S. Release of lipoteichoic acid from *Staphylococcus aureus* by treatment with Cefmetazole and other B-lactam antibiotics. *J. Antibiot.* 36, 1380–1386 (1983).
23. Magnin, D. R., Robl, J. A., Sulsky, R. B., Augeri, D. J., Huang, Y., Simpkins, L. M., *et al.* Synthesis of Novel Potent Dipeptidyl Peptidase IV Inhibitors with Enhanced Chemical Stability: Interplay between the N-Terminal Amino Acid Alkyl Side Chain and the Cyclopropyl Group of α -Aminoacyl-L-cis-4,5-methanoprolinenitrile-Based Inhibitors. *J. Med. Chem.* 47 (10), 2587–2598 (2004).
24. Robinson, J. L. & Brynildsen, M. P. An ensemble-guided approach identifies ClpP as a major regulator of transcript levels in nitric oxide-stressed *Escherichia coli*. *Metab. Eng.* 31, 22–34 (2015).
25. Dasari, D. & Tchounwou P. B. Cisplatin in cancer therapy: molecular mechanisms of action. *Eur J Pharmacol* 5, 364–378 (2015).
26. Damia, G., Filiberti, L., Vikhanskaya, F., Carrassa, L., Taya, Y., D'Incalci, M., *et al.* Cisplatin and Taxol Induce Different Patterns of p53 Phosphorylation. *Neoplasia.* 3 (1), 10–16 (2001).
27. Kane, R. C., Bross, P. F., Farrell, A. T. & Pazdur, R. Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy. *Oncologist.* 8 (6), 508–513 (2003).

-
28. Moreira, W., Ngan, G. J., Low, J. L., Poulsen, A., Chia, B. C., Ang, M. J., *et al.* Target mechanism-based whole-cell screening identifies bortezomib as an inhibitor of caseinolytic protease in mycobacteria. *mBio*. 6 (3), e00253-15 (2015).
 29. Kane, R. C., Dagher, R., Farrell, A., Ko, C. W., Sridhara, R., Justice, R., *et al.* Bortezomib for the treatment of mantle cell lymphoma. *Clin. Cancer Res.* 13 (18), 5291–5294 (2007).
 30. Rusbridge, N. M. & Beynon, R. J. 3,4-Dichloroisocoumarin, a serine protease inhibitor, inactivates glycogen phosphorylase b. *FEBS Lett.* 268 (1), 133–136 (1990).
 31. Luckner, P. & Brandsch, M. Interaction of 31 β -lactam antibiotics with the H⁺/peptide symporter PEPT2: analysis of affinity constants and comparison with PEPT1. *Eur. J. Pharm. Biopharm.* 59 (1), 17–24 (2005).
 32. Sass, P. & Bierbaum, G. Lytic activity of recombinant bacteriophage phi11 and phi12 endolysins on whole cells and biofilms of *Staphylococcus aureus*. *Appl Environ Microbiol.* 73 (1), 347-352 (2007).
 33. Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., *et al.* ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* 40, 1100–1107 (2012).
 34. Zhang, J. H. & Chung, T. D. Y. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screening.* 4 (2), 67-73 (1999).
 35. Iversen, P. W., Eastwood, B. J., Sittampalam, G. S. & Cox, K. L. A comparison of assay performance measures in screening assays: Signal window, Z' factor, and assay variability ratio. *J. Biomol. Screen.* 11 (3), 247–252 (2006).
 36. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 267 (3), 727–748 (1997).
 37. Bonetti A., Leone R., Muggia F. M. & Howell S. B. Platinum and Other Heavy Metal Compounds in Cancer Chemotherapy. *Molecular Mechanisms and Clinical Applications*, Springer. (2009).
 38. Moreno-Cinos, C., Sasseti, E., Salado, I. G., Witt, G., Benramdane, S., Reinhardt, L., *et al.* α -Amino Diphenyl Phosphonates as Novel Inhibitors of Escherichia coli ClpP Protease. *Journal of Medicinal Chemistry Article ASAP.* 62 (2), 774–797 (2019).

3.8 Supplementary data

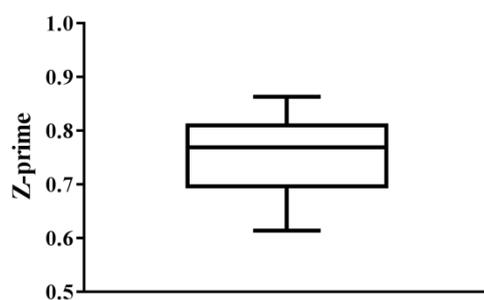


Figure S1. Overall Z' (Z-prime) for the libraries screening in this study.

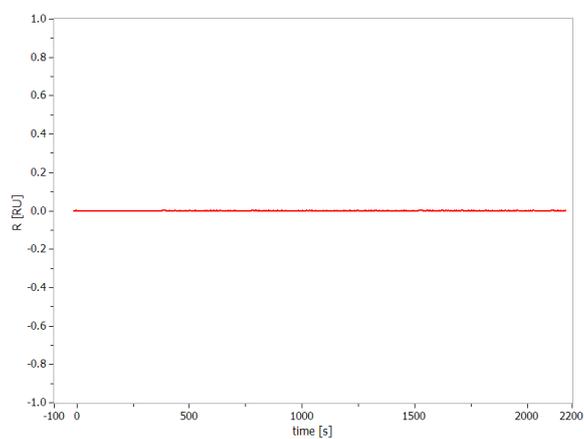


Figure S2. SPR sensorgram of the negative control caffeine at 80 μM compound concentration.

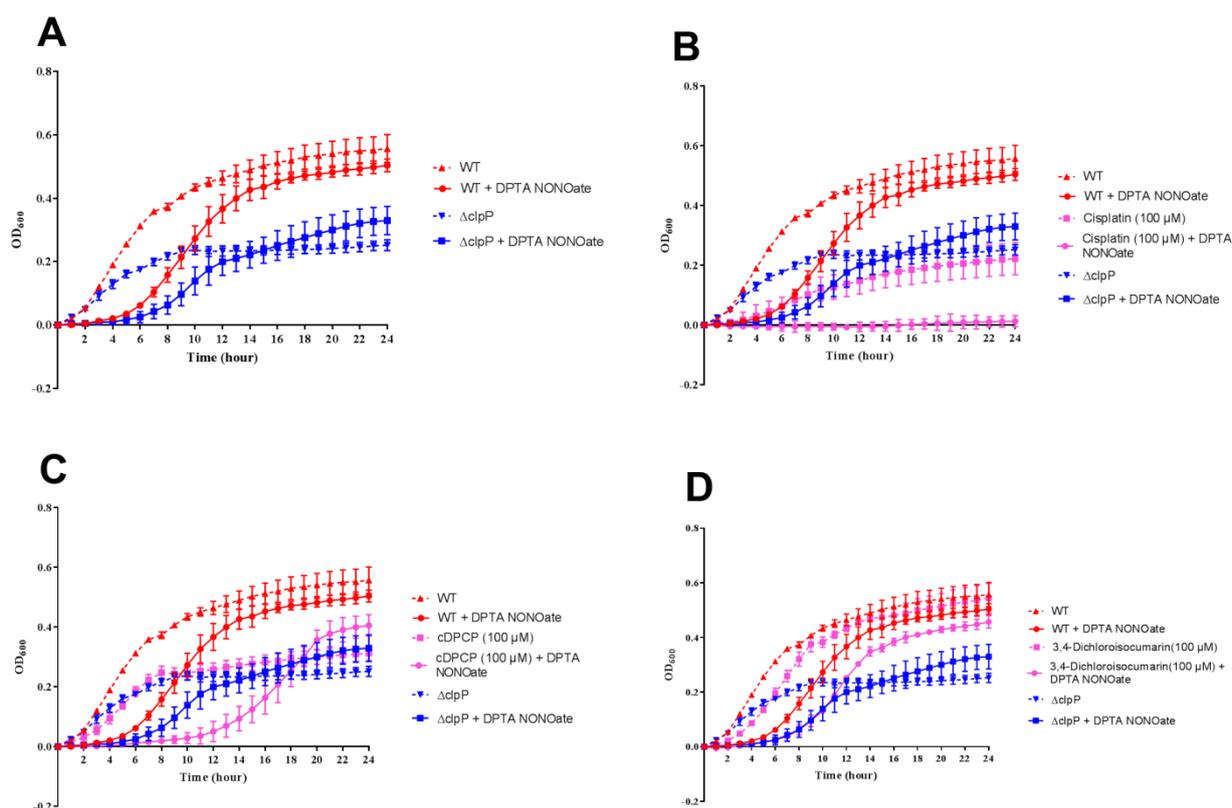


Figure S3. As, shown in Figure 5 but compounds at Bacterial growth assays for wt and $\Delta clpP$ *E. coli* strains in minimal media, in the presence (solid lines) and absence (dotted lines) of DTPA NONOate ($\text{NO}\cdot$) induced stress for 3 selected compounds. A) wt and $\Delta clpP$ alone. B) Cisplatin, BC) cDPCP and D) 3,4-dichloroisocoumarin. As shown in Figure 5 but compounds at 100 μM .

Table S1. List of compounds showing less than 70 % of the remaining activity of *E. coli* ClpP (200 μM) in the primary screen. Guanabenz acetate results to be active in the initial screening but not in the in the dose-response validation after the compound was newly purchased from a supplier, probably due to the different composition of isomers. The compound was excluded from further analysis.

Compound name	Library	CAS	% of inhibition in <i>E. coli</i> ClpP	MW	SMILES
Bortezomib	SCREEN- WELL®	17932 4-69-7	100,00	384,2	CC(C)C[C@@H](NC(=O) [C@H](CC1=CC=CC=C1) NC(=O)C1=NC=CN=C1) B(O)O

Atazanavir	SCREEN- WELL®	19890 4-31-3	85,03	704,9	<chem>COC(=O)N[C@H](C(=O)N[C@@H](CC1=CC=CC=C1)[C@@H](O)CN(CC1=CC=C(C=C1)C1=NC=C(C=C1)NC(=O)[C@@H](NC(=O)OC)C(C)(C)C)C)C</chem>
Sulfasalazine	SCREEN- WELL®	599- 79-1	77,33	398,4	<chem>OC(=O)C1=CC(=CC=C1O)\N=N\C1=CC=C(C=C1)S(=O)(=O)NC1=CC=CC=N1</chem>
Balsalazide	SCREEN- WELL®	80573- 04-2	76,46	357,3	<chem>O=C(O)CCNC(=O)c1ccc(/N=N/c2cc(C(=O)O)c(O)cc2)cc1</chem>
Rifapentine	SCREEN- WELL®	61379- 65-5	71,54	877	<chem>O=C(O[C@H]1[C@H](C)[C@H](O)[C@H](C)[C@@H](O)[C@@H](C)/C=C/C=C(/C)\C(=O)Nc2c(/C=N/N3CCN(C4CCCC4)CC3)c(O)c3c(c(O)c(C)c4O[C@@](C)(C(=O)c34)O/C=C/[C@H](OC)[C@H]1C)c2O)C</chem>
Silver Sulfadiazine	SCREEN- WELL®	22199- 08-2	88,38	357,1	<chem>S(=O)([O-])(=Nc1ncccn1)c1ccc(N)cc1.[Ag+]</chem>
Sunitinib Malate	SCREEN- WELL®	34103 1-54-7	80,67	532,6	<chem>Fc1cc2/C=C/c3c(C)c(C(=O)NCCN(CC)CC)c(C)[nH]3)/C(=O)Nc2cc1.O=C(O)[C@@H](O)CC(=O)O</chem>
Nitazoxanide	SCREEN- WELL®	55981- 09-4	79,96	307,3	<chem>O=[N+](O-)]c1sc(NC(=O)c2c(OC(=O)C)cccc2)nc1</chem>
Cisplatin (Cis-Diamineplatinum(II) Dichloride)	SCREEN- WELL®,	15663- 27-1	96,54	300,1	<chem>[Pt+2](Cl)(Cl)([N+H3])[N+H3]</chem>

	LOPAC®1				
	280				
Ezatiostat	LOPAC®1	16868	88,80	529,7	<chem>S(C[C@H](NC(=O)CC[C@H](N)C(=O)OCC)C(=O)N[C@@H](C(=O)OCC)c1cccc1)Cc1cccc1</chem>
	280	2-53-9			
Cefmetazole sodium	LOPAC®1	56796-	94,51	493,5	<chem>S(CC#N)CC(=O)N[C@]1(OC)C(=O)N2C(C(=O)[O-])=C(CSc3n(C)nnn3)CS[C@H]12.[Na+]</chem>
	280	39-5			
10058-F4	LOPAC®1	40381	72,82	249,4	<chem>S=C1S/C(=C\c2ccc(CC)cc2)/C(=O)N1</chem>
	280	1-55-2			
Guanabenz acetate	LOPAC®1	23256-	100,00	291,1	<chem>Clc1c(/C=N/NC(=N)N)c(C1)ccc1.O=C(O)C</chem>
	280	50-0			
Bisdemethoxycurcumin	LOPAC®1	33171-	76,54	308,3	<chem>O=C(/C=C/c1ccc(O)cc1)C(C(=O)/C=C/c1ccc(O)cc1</chem>
	280	05-0			
SR 27897 hydrate	LOPAC®1	13638	84,71	411,9	<chem>Clc1c(-c2nc(NC(=O)c3n(CC(=O)O)c4c(c3)cccc4)sc2)cccc1</chem>
	280	1-85-6			
S 24795	LOPAC®1	30467	76,27	418,1	<chem>BrC1ccc(C(=O)Cc2[n+](C)cccc2)cc1.[I-]</chem>
	280	9-75-2			
3, 4-Dichloroisocoumarin	LOPAC®1	51050-	98,92	215,0	<chem>ClC1=C(Cl)c2c(C(=O)O1)cccc2</chem>
	280	59-0			
Retinoic acid p-hydroxyanilide	LOPAC®1	65646-	74,12	391,6	<chem>O=C(Nc1ccc(O)cc1)/C=C(\C=C\C=C(/C=C/C=1C(C)(C)CCCC=1C)\C)/C</chem>
	280	68-6			
Myricetin	LOPAC®1	529-	66,04	318,2	<chem>O=C1C(O)=C(c2cc(O)c(O)c(O)c2)Oc2c1c(O)cc(O)c2</chem>
	280	44-2			
Tyrphostin 51	LOPAC®1	12643	72,80	268,2	<chem>Oc1c(O)cc(/C(/C#N)=C(/N)\C=C(/C#N)\C#N)cc1O</chem>
	280	3-07-6			
Tyrphostin 23	LOPAC®1	11840	79,28	186,2	<chem>Oc1c(O)ccc(/C=C(\C#N)/C#N)c1</chem>
	280	9-57-7			
cDPCP	LOPAC®1	10634	91,16	377,1	<chem>[Pt](Cl)(N)(N)[n+]1cccc1.[Cl-]</chem>
	280	3-54-8			

2-({4- [(cyclohexylamino) sulfonyl]anilino}ca rbonyl)cyclohexane carboxylic acid	MMP- Spec	n.a.	72,93	408,5	(=O)(=O)(NC1CCCCC1)c 1ccc(NC(=O)[C@H]2[C@ @H](C(=O)O)CCCC2)cc1
3-cyclopentyl-N- [2-(4- morpholinyl)ethyl] propanamide	MMP- Spec	n.a.	84,44	254,4	O=C(NCCN1CCOCC1)C CC1CCCC1

Table S2. Complete list of cepheids and penems compounds present in the screened compound libraries.

Compound name	Library	CAS	SMILES
Cefepime hydrochloride Hydrate	SCREEN- WELL®	123171- 59-5	[H+].[O-].[Cl-].[Cl-].CO\N=C(/C(=O)N[C@H]1[C@H]2SCC(C[N+]3(C)CCC C3)=C(N2C1=O)C(O)=O)C1=CSC(N)=N1
Dicloxacillin sodium Salt Monohydrate	SCREEN- WELL®	13412- 64-1	O.[Na+].[H][C@]12SC(C)(C)[C@@H](N1C(=O)[C@H]2 NC(=O)C1=C(C)ON=C1C1=C(Cl)C=CC=C1Cl)C([O-])=O
Doripenem	SCREEN- WELL®	148016- 81-3	[H][C@]12[C@@H](C)C(S[C@@H]3CN[C@H](CNS(N) (=O)=O)C3)=C(N1C(=O)[C@]2([H])[C@@H](C)O)C(O) =O
Imipenem	SCREEN- WELL®	64221- 86-9	[H][C@](C)(O)[C@@]1([H])C(=O)N2C(C(O)=O)=C(C[C@ @]12[H])SCCNC=N
Orlistat (Tetrahydrolipstatin)	SCREEN- WELL®	96829- 58-2	CCCCCCCCCCC[C@@H](C[C@@H]1OC(=O)[C@H]1 CCCCC)OC(=O)[C@H](CC(C)C)NC=O
Meropenem	SCREEN- WELL®	96036- 03-2	[H][C@]1([C@@H](C)O)C(=O)N2C(C(O)=O)=C(S[C@ @H]3CN[C@@H](C3)C(=O)N(C)C)[C@H](C)[C@]12[H]
Ampicillin Trihydrate	SCREEN- WELL®	7177- 48-2	O.O.O.CC1(C)S[C@@H]2[C@H](NC(=O)[C@H](N)C3= CC=CC=C3)C(=O)N2[C@H]1C(O)=O
Aztreonam	SCREEN- WELL®	78110- 38-0	C[C@H]1[C@H](NC(=O)C(=N/OC(C)(C)C(O)=O)\C2=C SC(N)=N2)C(=O)N1S(O)(=O)=O
Ceftazidime	SCREEN- WELL®	72558- 82-8	CC(C)(O)N=C(/C(=O)N[C@H]1[C@H]2SCC(C[N+]3=C C=CC=C3)=C(N2C1=O)C([O-])=O)C1=CSC(N)=N1)C(O)=O

Oxacillin sodium salt monohydrate	SCREEN- WELL®	7240- 38-2	<chem>CC1=C(C(=NO1)C2=CC=CC=C2)C(=O)NC3C4N(C3=O)C(C(S4)(C)C)C(=O)[O-].O.[Na+]</chem>
Penicillin Potassium	V SCREEN- WELL®	132-98- 9	<chem>[K+].CC1(C)S[C@@H]2[C@H](NC(=O)COC3=CC=CC=C3)C(=O)N2[C@H]1C([O-])=O</chem>
Piperacillin	SCREEN- WELL®	61477- 96-1	<chem>CCN1CCN(C(=O)N[C@@H](C(=O)N[C@H]2[C@H]3S C(C)(C)[C@@H](N3C2=O)C(O)=O)C2=CC=CC=C2)C(=O)C1=O</chem>
Amoxicillin	SCREEN- WELL®	26787- 78-0	<chem>CC1(C)S[C@@H]2[C@H](NC(=O)[C@H](N)C3=CC=C(O)C=C3)C(=O)N2[C@H]1C(O)=O</chem>
Cefadroxil	SCREEN- WELL®	66592- 87-8	<chem>O.[H][C@]12SCC(C)=C(N1C(=O)[C@H]2NC(=O)[C@H](N)C1=CC=C(O)C=C1)C(O)=O</chem>
Cefdinir	SCREEN- WELL®	91832- 40-5	<chem>[H][C@]12SCC(C=C)=C(N1C(=O)[C@H]2NC(=O)C(=N/O)\C1=CSC(N)=N1)C(O)=O</chem>
Cefditoren Pivoxil	SCREEN- WELL®	117467- 28-4	<chem>[H][C@]12SCC(\C=C/C3=C(C)N=CS3)=C(N1C(=O)[C@H]2NC(=O)C(=N/OC)\C1=CSC(N)=N1)C(=O)OCOC(=O)C(C)(C)C</chem>
Cefixime	SCREEN- WELL®	79350- 37-1	<chem>[H][C@]12SCC(C=C)=C(N1C(=O)[C@H]2NC(=O)C(=N/OCC(O)=O)\C1=CSC(N)=N1)C(O)=O</chem>
Cefotetan Disodium	SCREEN- WELL®	74356- 00-6	<chem>[Na+].[Na+].[H][C@]12SCC(CSC3=NN=NN3C)=C(N1C(=O)[C@]2(NC(=O)C1SC(S1)=C(C(N)=O)C([O-])=O)OC)C([O-])=O</chem>
Cefotaxime Acid	SCREEN- WELL®	63527- 52-6	<chem>[H][C@]12SCC(COC(C)=O)=C(N1C(=O)[C@H]2NC(=O)C(=N/OC)\C1=CSC(N)=N1)C(O)=O</chem>
Cefpodoxime Proxetil	SCREEN- WELL®	87239- 81-4	<chem>[H][C@]12SCC(COC)=C(N1C(=O)[C@H]2NC(=O)C(=N/OC)\C1=CSC(N)=N1)C(=O)OC(C)OC(=O)OC(C)C</chem>
Cefprozil	SCREEN- WELL®	92665- 29-7	<chem>[H][C@]12SCC(\C=C\C)=C(N1C(=O)[C@H]2NC(=O)[C@H](N)C1=CC=C(O)C=C1)C(O)=O</chem>
Ceftibuten	SCREEN- WELL®	97519- 39-6	<chem>[H][C@]12SCC=C(N1C(=O)[C@H]2NC(=O)C(=C/CC(O)=O)\C1=CSC(N)=N1)C(O)=O</chem>
Ceftizoxim sodium	SCREEN- WELL®	68401- 82-1	<chem>[Na+].[H][C@]12SCC=C(N1C(=O)[C@H]2NC(=O)C(=N/OC)\C1=CSC(N)=N1)C([O-])=O</chem>
Cefuroxime Axetil	SCREEN- WELL®	64544- 07-6	<chem>[H][C@]12SCC(COC(N)=O)=C(N1C(=O)[C@H]2NC(=O)C(=N/OC)\C1=CC=CO1)C(=O)OC(C)OC(C)=O</chem>
Cefuroxime sodium	SCREEN- WELL®	56238- 63-2	<chem>[Na+].[H][C@]12SCC(COC(N)=O)=C(N1C(=O)[C@H]2NC(=O)C(=N/OC)\C1=CC=CO1)C([O-])=O</chem>

Cephalexin	SCREEN-	23325-	O.CC1=C(N2[C@H](SC1)[C@H](NC(=O)[C@H](N)C1=
Monohydrate	WELL®	78-2	CC=CC=C1)C2=O)C(O)=O
Clavulanate	SCREEN-	61177-	[K+].[H][C@@]12CC(=O)N1[C@@H](C([O-
Potassium	WELL®	45-5])=O)\C(O2)=C\CO
Cloxacillin sodium	SCREEN-	7081-	O.[Na+].[H][C@]12SC(C)(C)[C@@H](N1C(=O)[C@H]2
	WELL®	44-9	NC(=O)C1=C(C)ON=C1C1=C(Cl)C=CC=C1)C([O-])=O
Ezetimibe	SCREEN-	163222-	O[C@@H](CC[C@@H]1[C@H](N(C1=O)C1=CC=C(F)
	WELL®	33-1	C=C1)C1=CC=C(O)C=C1)C1=CC=C(F)C=C1
Nafcillin sodium	SCREEN-	985-16-	[H][C@]12SC(C)(C)[C@@H](N1C(=O)[C@H]2NC(=O)
	WELL®	0	C1=C(OCC)C=CC2=C1C=CC=C2)C(=O)O[Na]
Penicillin	G SCREEN-	113-98-	CC1(C(N2C(S1)C(C2=O)NC(=O)CC3=CC=CC=C3)C(=O)
Potassium	WELL®	4)[O-])C.[K+]
(Benzylpenicillin)			
Cefaclor	SCREEN-	53994-	[H]C12SCC(Cl)=C(N1C(=O)C2NC(=O)C(N)C1=CC=CC
	WELL®,	73-3	=C1)C(O)=O
	LOPAC®		
		1280	
Cefazolin sodium	SCREEN-	27164-	[Na+].[H][C@]12SCC(CSC3=NN=C(C)S3)=C(N1C(=O)[
	WELL®,	46-1	C@H]2NC(=O)CN1C=NN=N1)C([O-])=O
	LOPAC®		
		1280	
Ceftriaxone sodium	SCREEN-	104376-	CN1C(=NC(=O)C(=O)N1)SCC2=C(N3[C@@H]([C@@H
	WELL®,	79-6](C3=O)NC(=O)/C(=N\OC)/C4=CSC(=N4)N)SC2)C(=O)[
	LOPAC®		O-].[Na+]
		1280	
Cephalexin hydrate	LOPAC®	15686-	O=C(N[C@@H]1C(=O)N2C(C(=O)O)=C(C)CS[C@H]12
		1280)C(N)c1ccccc1
		71-2	
Cefsulodin sodium	LOPAC®	52152-	S(=O)(=O)([O-
salt hydrate	1280	93-9])[C@@H](C(=O)N[C@@H]1C(=O)N2C(C(=O)[O-
])=C(C[n+]3ccc(C(=O)N)cc3)CS[C@H]12)c1ccccc1.[Na+
]
Cefmetazole sodium	LOPAC®	56796-	S(CC#N)CC(=O)N[C@]1(OC)C(=O)N2C(C(=O)[O-
	1280	39-5])=C(CSc3n(C)nn3)CS[C@H]12.[Na+]
Imipenem	LOPAC®	74431-	S(CCNC=N)C1=C(C(=O)O)N2C(=O)[C@H]([C@H](O)C
monohydrate	1280	23-5)[C@H]2C1

Cephalosporin C zinc salt	LOPAC® 1280	59143- 60-1	<chem>O=C([O-])[C@H](N)CCCC(=O)N[C@@H]1C(=O)N2C(C(=O)[O-])=C(COC(=O)C)CS[C@H]12.[Zn+2]</chem>
Cephalothin sodium	LOPAC® 1280	58-71-9	<chem>O=C(N[C@@H]1C(=O)N2C(C(=O)[O-])=C(COC(=O)C)CS[C@H]12)Cc1sccc1.[Na+]</chem>
Cephradine	LOPAC® 1280	38821- 53-3	<chem>O=C(N[C@@H]1C(=O)N2C(C(=O)O)=C(C)CS[C@H]12)C(N)C1=CCC=CC1</chem>
Cefotaxime sodium	LOPAC® 1280	64485- 93-4	<chem>O=C(N[C@@H]1C(=O)N2C(C(=O)[O-])=C(COC(=O)C)CS[C@H]12)/C(=N/OC)/c1nc(N)sc1.[Na+]</chem>
Pivmecillinam	LOPAC® 1280	32886- 97-8	<chem>O=C(OCOC(=O)[C@H]1C(C)(C)S[C@@H]2[C@H](/N=C/N3CCCCC3)C(=O)N12)C(C)(C)C</chem>
N-cyclohexyl-1-(2-oxo-1-azetidiny)cyclohexanecarboxamide	MMP-Spec	n.a.	<chem>(=O)(=O)(NC1CCCCC1)c1ccc(NC(=O)[C@H]2[C@@H](C(=O)O)CCCC2)cc1</chem>
N-cyclohexyl-1-(2-oxo-1-azetidiny)cyclopentanecarboxamide	MMP-Spec	n.a.	<chem>O=C(NCCN1CCOCC1)CCC1CCCC1</chem>

Chapter 4

Crystallography

Contribution to the work. I have performed the optimization of the protein purification protocol, purified the protein itself and participated in the optimization of crystallization and co-crystallization conditions with help from experts in the field (in particular Anne-Sophie Humm and Dr. Irina Cornaciu). Also, I have analyzed the resultant crystal structure and compared with other structures available. The rest of the work was conducted by Anne-Sophie Humm under the supervision of Dr. Irina Cornaciu and Dr. José Marquéz.

4.1 Introduction

4.1.1 X-ray crystallography

Since the first crystal structure for myoglobin was solved at a resolution of 6 Å, in 1957, structural biology has played a critical role in advancing the understanding of biological phenomena. Initially structural studies were focused on the structures of proteins, then moved into the fields of virus ¹ and DNA ², lately, it has been also applied to complexes as protein-protein, nucleic-acid-protein and antigen-antibody complexes³.

After the first structures were published, there was the need of a comprehensive database to collect results solved and serve as a repository for work on upcoming proteins, and in 1971, the Protein Data Bank (PDB), was created ². Accordingly with the information reported in PDB website (<https://www.rcsb.org/stats/summary>), today the number of deposited structures in PDB stands at over 144210. Structural biology utilises several techniques besides X-ray crystal structure determination: cryo-electron microscopy and NMR spectroscopy, neutron diffraction and X-ray solution scattering. Nevertheless, X-ray crystallography remains at this point the most widely applied methodology ⁴. The fundamentals of X-ray crystallography are based on the X-ray diffraction by ordered crystal structures, followed by detection and deconvolution of the atom distribution ⁴. In protein X-ray crystallization, high quality and sufficient concentrated protein and soluble material are required together with crystallization conditions that promotes crystal growth instead of buffer salt precipitation ³. The size of the crystal is also important since high intensity synchrotron X-ray beams can damage the crystal during measurements, however, this problem is partially reduced now, allowing the study of smaller crystals, with the use of micro-focused beam lines and cryogenic temperatures during data collection ⁵.

Lately, high throughput concepts have been applied to crystallography and it has been employed routinely in drug discovery programs. Automatization is now a standard procedure in crystallization and handling processes on synchrotron beam lines. Miniaturization to 96 well format with lower volume is also reducing material requirements and increasing throughput ⁶. Several drug-discovery based on crystallography have led to clinical trials drugs ^{6,7}, and commercial companies are using library of fragments with crystal soaking and a mixture of 2 to 8 compounds per condition to perform initial hit finding using X-ray crystallography ⁸. The European consortium INSTRUMENT (research infrastructure in structural biology) and the related consortium iNEXT (infrastructures for NMR, EM and X-rays for translational research) are performing X-ray crystallography fragment based drug discovery in several targets. Eventually, the target of this PhD work (*Escherichia coli* ClpP) will be employed in this iNEXT project as follow-up study.

4.1.2 ClpP crystal structures

An analysis of the crystal structures of caseinolytic protease subunit P reveals the presence of 54 structures from 14 different organisms in the PDB. Among these are 12 bacteria and 2 eukaryote (human and *Plasmodium falciparum*) forms (see Table 1). The Homo sapiens structures correspond to the mitochondria associated form since in man this is the only location of ClpP. In general, the available co-crystallized structures (13) involve ADEP class compounds. Interestingly published human ClpP structures also include a complex with a compound from this class, indicating potential selectivity issues. Overall just 5 complexes are in presence of a non-ADEPs. Only 3 of the structures (Table 1) from bacteria and involve inhibiting compounds and target the pocket of the protein where the catalytic triad is located. Those structures are two from Gram-positive bacteria *Staphylococcus aureus* (PDB ID 5DL1, ligand name AV145) ⁹ and *Bacillus subtilis* (PDB ID 3TT7, ligand name DFP) ¹⁰, only one from the Gram-negative *Escherichia coli* (PDB ID 2FZS, ligand name Z-LY-CMK) ¹¹. Lastly, a *Staphylococcus aureus* structure (PDB ID 5DL1) ⁹ represent the only co-crystallized ClpP structure with a non covalent inhibitor.

Table 1. Summary of reported ClpP crystal structures in PDB (PROTEIN DATA BANK). *unrealized structures; bold: ligand effecting as inhibitor.

Organism	Number of structures	number of co-crystal structures	PDB ID co-crystal structure and ligand
<i>Staphylococcus aureus</i>	13	3	5WI8 (urea depsipeptide); 5VZ2 (ADEP); 5DL1 (AV145)
<i>Bacillus subtilis</i>	7	4	3TT7 (DFP) ; 3KTI (ADEP 1); 3KTJ (ADEP 2); 3KTK (ADEP 2)

Chapter 4

<i>Escherichia coli</i>	6	2	3MT6 (ADEP 1); 2FZS (Z-LY-CMK)
<i>Mycobacterium tuberculosis</i>	6	1	4U0G (ADEP)
<i>Helicobacter pylori</i>	4		
<i>Francisella tularensis</i>	4		
<i>Lysteria monocytogenes</i>	4		
<i>Homo sapiens</i>	3	2	6H23 (activating molecule); 6BBA (ADEP 28)
<i>Pseudomonas aeruginosa</i>	2*		
<i>Nisseria meningidis</i>	1	1	5DKP (ADEP A54556)
<i>Streptococcus pneumoniae</i>	1		
<i>Enterococcus faecium</i>	1		
<i>Coxiella brunetii</i>	1		
<i>Plasmodium falciparum</i>	1		

4.2 Overview

This ongoing aspect of the thesis work is being performed in conjunction with the of the European Horizon 2020 iNEXT consortium, specifically the High Throughput Crystallization Facility (Xtal lab) at EMBL Grenoble which is equipped to perform fully automated crystallization screening process. The application of crystallographic techniques in this project has the primary aim of structural elucidation of *Escherichia coli* ClpP protein in complex with priority compounds in order to better understand the mechanism of action of the inhibitors. Specifically, we intend to crystallize the most active and interesting inhibitors reported in Chapter 2 and Chapter 3. The information obtained will complement the activity-related and biophysical data-sets in place for the compounds and inform further rational optimization of inhibitor properties.

A second part of the project will rely on our access to the iNEXT hit finding platform for large scale crystallographic screening of libraries of compound fragments against ClpP. This will enable identification of fragments-based ligands suitable for further expansion into leads with inhibitory properties.

4.3 Status

Since the project is still running, this section reports upon the current status of the project: final results and conclusions are not yet in place. To obtain the high purity *Escherichia coli* ClpP protein suitable for protein crystallography, the purification protocol described in the sections 2.5.1 and 3.5 was optimized further

Purification protocol. 10 ml of overnight culture were used to inoculate 1 liter of fresh LB + 100 µg/ml Ampicillin and left to grow at 37° C, shaking at 130 rpm until the culture reach $OD_{600} = 0,6$. The induction was performed with 1mM IPTG and growth for 5h at 30° C, shaking at 130 rpm. The cultures were centrifuged at 4000 g. for 20 minutes at 4° C and the supernatant discarded. The lysis of the resuspended resultant bacterial pellet (in buffer containing 50 mM Tris, 150 mM NaCl, 150 mM NaCl, pH 7.6) underwent microfluidization and DNase was added. The output was centrifuged at 18 000 rpm for 45 min at 4° C and the collected supernatant filtered through 45µm filter. Purification used immobilized metal affinity chromatography (IMAC), with ÄKTAprime plus (GE Heath care life sciences, USA) and a 1mL prepack column HisTrap HP (GE Heath care life sciences, USA) with 2 component gradient (buffer A 50 mM Tris, 150 mM NaCl, 150 mM NaCl, pH 7.6, buffer B 50 mM Tris, 150 mM NaCl, 150 mM NaCl, 500 mM imidazole pH 7.6). SDS-gel analysis was performed at the end of purification steps (see Figure 1). Buffer exchange, to avoid protein precipitation, was performed in a PD-10- column (GE Heath care life sciences, USA). After protein concentration, the protein was at 20 mg/ml.

The most suitable screening conditions for crystallization for ClpP were identified using several commercially available 96 well plate-based screening platforms. In particular ProPlex, JCSG and Morpheus (Molecular dimension, UK), Classic-Suite, ProComplex, MPD, PEGs and Nucleix (Quiagen, Germany) and Index

(Hampton Research, USA). The drops for the crystallization process (protein concentration 20 mg/ml) were disposed by the dispensing workstation Cartesian PixSys (Cartesian Technologies, USA) and the liquid handling robot nanolitre pipetting Mosquito HTS (TTP Labtech, UK) using the sitting drop approach. The drops were continually monitored by an automated image system, Rock Imager R1000 (Formulatrix, USA), the harvesting of the formed crystal operated by CrystalDirect, crystal harvester of EMBL and the data management system was performed by the in-house developed “Crystallization Information Management System (CRIMS)”.

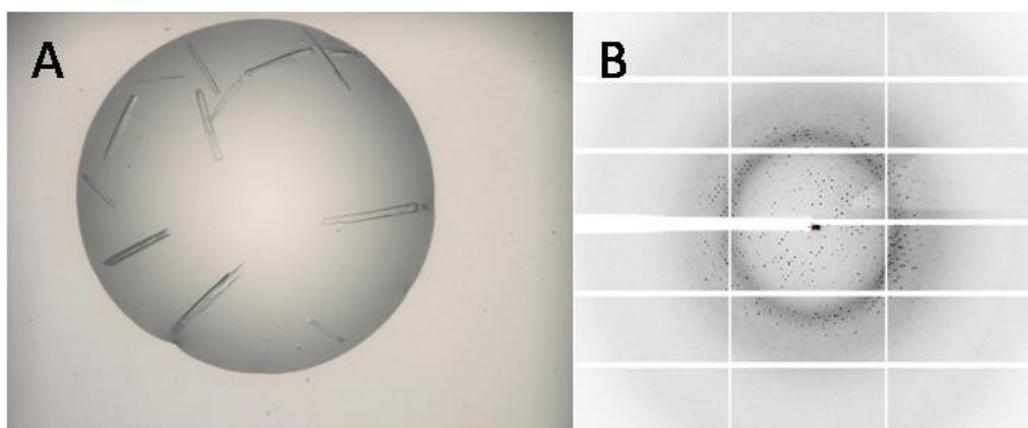


Figure 1. A) Drop containing crystals, B) scattering plot of the harvested crystal of A.

To date, a crystal structure at 2.02 Å of *Escherichia coli* ClpP is available as an intermediate output of the project. The structure does not present any binding compounds, but the apo-structure. The specific crystal harvested (Figure 1 A) to perform this diffraction were from Index (Hampton Research, USA), screening plate. The conditions were the following: 0.2 M ammonium acetate, 0.1 M bis-tris pH 5.5, 45 % (v/v) 2-methyl-2,4-pentanediol. The data were collected at the synchrotron of Grenoble (France) at ID30-A1 (massif beam-line) (see Figure 1 B). The data processing was done with Autoproc (Global Phasing suite) and refinement performed with Buster (Global Phasing suite).

4.4 Comments on the crystal structure

Our structure, solved in space group P1 21 1 containing a large unit cell with 2 times the 14 subunits (the entire structure is tetradecameric), shows as an “impurity” the presence of the solvent molecule 2-methyl-2,4-pentanediol (MPD) located in the active site. There is no evidence of direct binding to the protein, rather it occupies the active site pocket in a “space-filling” manner (Figure 2). The presence of MPD in the active site should does not influence the co-crystallization since the compounds have a higher affinity for the protein compare to the solvent, on the other hand, this might influence a possible soaking procedure. In fact, soaking involved an already crystallized protein (with the presence of MPD co-crystallized in the active site) and just after the addition of the compound.

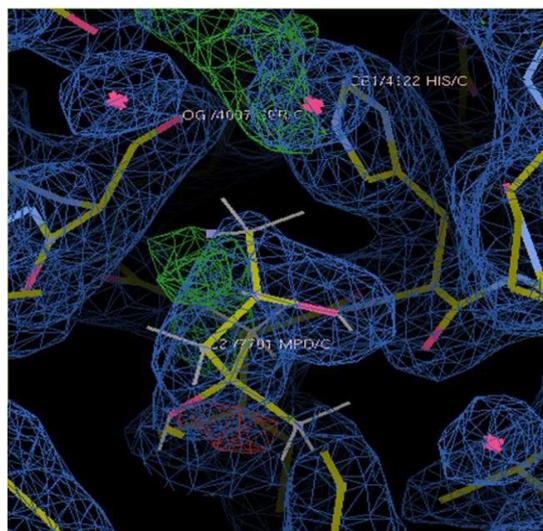


Figure 2. Electron density in the active site, with the presence of 2-methyl-2,4-pentanediol (MPD). Picture provided by Anne-Sophie Humm.

Among the published *Escherichia coli* ClpP structures, besides the co-crystallized structures with modulators (reported in Table 1 both with a resolution of 1.9 Å), are present other two with mutations in the amino acid sequences (A153C, PDB ID 3HLN, resolution 3.2 Å and V6A, PDB ID 1YG8, resolution 2.6 Å) and two apo-structures, the more recent of the two, with a resolution of 1.9 Å (PDB ID 1YG6) and the first deposited *Escherichia coli* ClpP with 2.3 Å (PDB ID 1TYF).

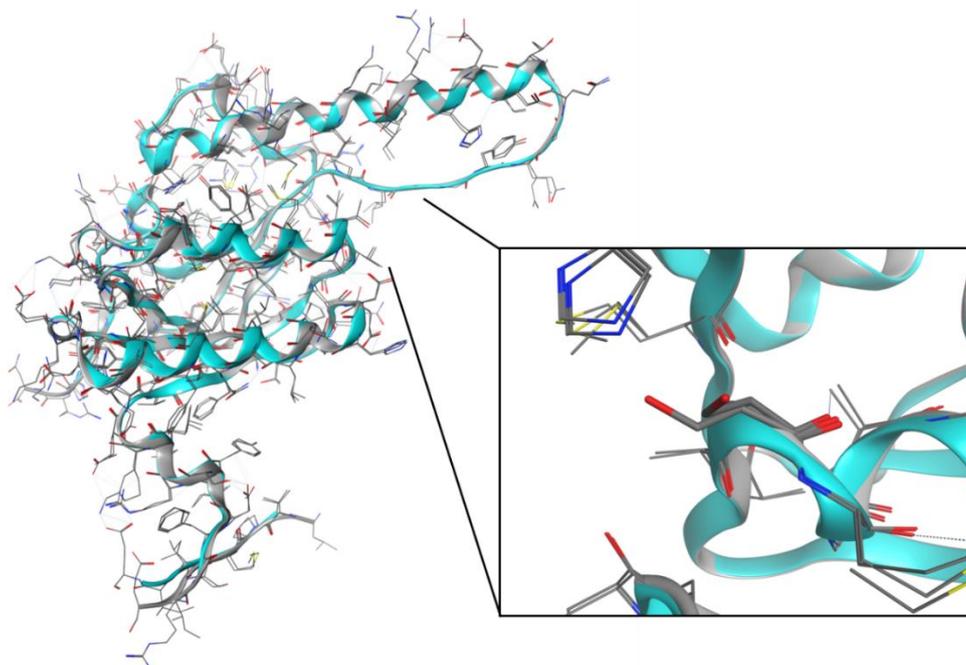


Figure 3. Superimposition of chain A of our structure (in grey) and chain A of 2FZS (in light blue). In detail the active site.

In complex, the structure defined by here results are similar in orientation to that which was used for docking studies, (2FZS)¹¹. Both structures have a comparable resolution (2.02 Å versus 1.9 Å). An analysis of the structures-comparison has been performed, first aligning and after superimposing the chains A of the two structures using the molecular modeling software suite Molecular Operating Environment (MOE, Chemical Computing Group Inc., Montreal, Canada) version 2016.0802, and the result is shown in Figure 3. Here the comparison of our structure, colored in grey and the *E. coli* ClpP published structure in light blue (2FZS) reveals a pretty strong similarity among the two, except a discrepancy that for the orientation of the catalytic serine, shown here in the zoomed section of the picture. The different orientation of the catalytic serine 97, principal discrepancy observed, is most likely due to the presence of the covalent inhibitor (Z-LY-CMK) bounded with the residue in 2FZS.

RMSD = 0.298 Å

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
1: EcClpP.A	0.00	0.23	0.29	0.25	0.32	0.39	0.20	0.23	0.22	0.22	0.18	0.19	0.23	0.18	0.23	0.24	0.32	0.58	0.30	0.26	0.37	0.40	0.29	0.27	0.30	0.26	0.34	0.30	0.30
2: EcClpP.B	0.21	0.00	0.30	0.29	0.31	0.23	0.23	0.23	0.22	0.20	0.22	0.23	0.20	0.23	0.25	0.29	0.50	0.32	0.26	0.36	0.44	0.27	0.26	0.28	0.28	0.34	0.29	0.29	0.29
3: EcClpP.C	0.29	0.30	0.00	0.27	0.37	0.28	0.27	0.23	0.24	0.30	0.26	0.27	0.27	0.30	0.26	0.35	0.50	0.28	0.30	0.38	0.42	0.32	0.31	0.30	0.28	0.33	0.33	0.33	
4: EcClpP.D	0.25	0.28	0.27	0.00	0.38	0.23	0.28	0.25	0.24	0.25	0.27	0.23	0.24	0.29	0.25	0.31	0.52	0.27	0.27	0.32	0.44	0.32	0.27	0.27	0.25	0.31	0.27	0.30	
5: EcClpP.E	0.37	0.30	0.37	0.38	0.00	0.30	0.30	0.30	0.28	0.30	0.36	0.34	0.33	0.34	0.33	0.35	0.53	0.36	0.35	0.43	0.40	0.32	0.33	0.29	0.35	0.35	0.36	0.37	
6: EcClpP.F	0.19	0.23	0.28	0.22	0.32	0.00	0.23	0.19	0.19	0.20	0.19	0.20	0.16	0.24	0.22	0.29	0.52	0.27	0.23	0.33	0.44	0.26	0.22	0.25	0.23	0.33	0.27	0.29	
7: EcClpP.G	0.20	0.21	0.27	0.28	0.33	0.21	0.00	0.21	0.22	0.18	0.19	0.21	0.21	0.20	0.25	0.34	0.51	0.32	0.27	0.40	0.47	0.31	0.29	0.31	0.28	0.35	0.33	0.32	
8: EcClpP.H	0.21	0.23	0.23	0.25	0.30	0.38	0.21	0.00	0.20	0.20	0.22	0.21	0.18	0.24	0.23	0.30	0.50	0.28	0.26	0.37	0.44	0.28	0.25	0.28	0.25	0.31	0.29	0.28	
9: EcClpP.I	0.22	0.22	0.24	0.24	0.28	0.38	0.23	0.20	0.00	0.22	0.22	0.23	0.20	0.25	0.22	0.27	0.49	0.28	0.24	0.33	0.40	0.23	0.23	0.22	0.25	0.28	0.26	0.28	
10: EcClpP.J	0.18	0.20	0.30	0.25	0.33	0.20	0.18	0.20	0.22	0.00	0.23	0.23	0.18	0.24	0.25	0.31	0.51	0.32	0.26	0.38	0.40	0.30	0.27	0.29	0.28	0.34	0.31	0.31	
11: EcClpP.K	0.19	0.22	0.26	0.27	0.36	0.38	0.39	0.22	0.22	0.23	0.00	0.23	0.23	0.25	0.25	0.33	0.54	0.33	0.26	0.38	0.40	0.30	0.28	0.32	0.27	0.34	0.31	0.33	
12: EcClpP.H	0.21	0.23	0.27	0.23	0.34	0.28	0.21	0.21	0.21	0.21	0.00	0.23	0.24	0.23	0.27	0.46	0.27	0.23	0.32	0.40	0.28	0.26	0.26	0.23	0.29	0.26	0.26		
13: EcClpP.M	0.18	0.20	0.27	0.24	0.33	0.38	0.23	0.18	0.20	0.23	0.23	0.23	0.00	0.23	0.23	0.29	0.53	0.29	0.25	0.36	0.43	0.27	0.25	0.28	0.24	0.31	0.29	0.29	
14: EcClpP.N	0.23	0.23	0.30	0.29	0.34	0.24	0.20	0.24	0.25	0.24	0.23	0.24	0.23	0.00	0.26	0.31	0.51	0.33	0.26	0.38	0.45	0.31	0.29	0.30	0.28	0.35	0.31	0.29	
15: 2FZS.A	0.24	0.25	0.26	0.25	0.33	0.22	0.25	0.23	0.22	0.25	0.25	0.23	0.23	0.26	0.00	0.21	0.44	0.17	0.15	0.28	0.36	0.19	0.17	0.19	0.14	0.22	0.18	0.19	
16: 2FZS.B	0.32	0.29	0.35	0.31	0.35	0.29	0.34	0.30	0.27	0.35	0.33	0.27	0.29	0.31	0.23	0.00	0.45	0.21	0.19	0.22	0.37	0.18	0.20	0.19	0.19	0.25	0.16	0.20	
17: 2FZS.C	0.34	0.30	0.30	0.32	0.33	0.32	0.33	0.30	0.40	0.33	0.34	0.46	0.33	0.33	0.44	0.41	0.00	0.42	0.45	0.48	0.24	0.46	0.46	0.46	0.43	0.40	0.43	0.38	
18: 2FZS.D	0.30	0.30	0.28	0.27	0.36	0.27	0.32	0.28	0.26	0.32	0.31	0.27	0.29	0.33	0.37	0.25	0.42	0.00	0.22	0.28	0.36	0.24	0.21	0.21	0.19	0.22	0.19	0.21	
19: 2FZS.E	0.28	0.28	0.30	0.27	0.35	0.23	0.27	0.26	0.24	0.26	0.26	0.23	0.23	0.26	0.15	0.19	0.45	0.22	0.00	0.28	0.38	0.19	0.17	0.20	0.15	0.24	0.17	0.18	
20: 2FZS.F	0.37	0.36	0.38	0.32	0.43	0.33	0.40	0.37	0.53	0.38	0.38	0.32	0.36	0.38	0.28	0.22	0.48	0.28	0.28	0.00	0.43	0.25	0.25	0.25	0.26	0.29	0.19	0.22	
21: 2FZS.G	0.45	0.45	0.42	0.44	0.45	0.44	0.47	0.44	0.40	0.40	0.40	0.40	0.43	0.43	0.36	0.37	0.24	0.36	0.38	0.43	0.00	0.40	0.40	0.36	0.36	0.33	0.37	0.35	
22: 2FZS.H	0.29	0.27	0.32	0.32	0.32	0.28	0.31	0.28	0.25	0.30	0.30	0.28	0.27	0.31	0.19	0.18	0.46	0.24	0.19	0.25	0.40	0.00	0.17	0.20	0.19	0.24	0.19	0.22	
23: 2FZS.I	0.27	0.28	0.31	0.27	0.33	0.23	0.29	0.25	0.23	0.27	0.28	0.26	0.25	0.28	0.17	0.20	0.46	0.21	0.17	0.25	0.40	0.17	0.00	0.17	0.14	0.17	0.20	0.20	
24: 2FZS.J	0.30	0.30	0.30	0.27	0.29	0.25	0.31	0.26	0.22	0.29	0.32	0.28	0.28	0.30	0.19	0.19	0.46	0.21	0.20	0.25	0.38	0.20	0.17	0.00	0.20	0.22	0.17	0.23	
25: 2FZS.K	0.26	0.26	0.28	0.25	0.35	0.23	0.28	0.25	0.25	0.26	0.27	0.23	0.24	0.26	0.14	0.19	0.43	0.19	0.15	0.26	0.36	0.19	0.17	0.20	0.00	0.23	0.17	0.18	
26: 2FZS.L	0.34	0.34	0.31	0.31	0.35	0.31	0.31	0.31	0.28	0.34	0.34	0.29	0.31	0.31	0.22	0.25	0.40	0.22	0.24	0.29	0.33	0.24	0.24	0.22	0.23	0.00	0.22	0.24	
27: 2FZS.M	0.30	0.30	0.33	0.27	0.36	0.27	0.31	0.29	0.26	0.35	0.31	0.26	0.29	0.31	0.18	0.16	0.43	0.19	0.17	0.19	0.37	0.19	0.17	0.17	0.17	0.22	0.00	0.17	
28: 2FZS.N	0.30	0.29	0.30	0.30	0.37	0.29	0.32	0.29	0.28	0.31	0.31	0.26	0.29	0.29	0.19	0.20	0.38	0.23	0.18	0.27	0.35	0.22	0.20	0.23	0.19	0.24	0.17	0.00	

Figure 4. Root-mean-square deviation (RMSD) between the C_{alpha} of the single chain of our structure (*EcClpP*) and the one used as reference structure (2FZS).

An overall evaluation of the distance between the atoms of the two structures in analysis has been performed with the measurement of the Root-mean-square deviation of the atomic position (RMSD), using MOE software. The calculation has been done for all the chains of our structure, versus all the chains of 2FZS, previously superimposed. The resultant distance, or the difference between the structures, is calculated in Ångström. Here the RMSD calculated between the carbon alpha (C_{alpha}) was equal to 0.289 Å, as reported in Figure.4 (RMSD of all the structures, not divided by chains, was resulting as 0,841, RMSD MAIN = 0,320).

In conclusion, the final results in quite nice proximity of the structures atomic coordinates, showing a strong similarity among those.

4.5 References

1. Harrison, S. C. & Jack, A. Structure of tomato bushy stunt virus: III. Three-dimensional X-ray diffraction analysis at 16 Å resolution. *J. Mol. Biol.* 97 (2), 173–191 (1975).
2. Shi, Y. A glimpse of structural biology through X-ray crystallography. *Cell.* 159 (5), 995–1014 (2014).
3. Smyth, M. S. & Martin, J. H. J. x Ray crystallography. *J. Clin. Pathol. - Mol. Pathol.* 53 (1), 8–14 (2000).
4. Wang, H. W. & Wang, J. W. How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Sci.* 26 (1), 32–39 (2017).
5. Clabbers, M. T. B. & Abrahams, J. P. Electron diffraction and three-dimensional crystallography for structural biology. *Crystallogr. Rev.* 24 (3), 176–204 (2018).
6. Zheng, H., Handing, K. B., Zimmerman, M. D., Shabalin, I. G., Almo, S. C. & Minor, W. X-ray crystallography over the past decade for novel drug discovery – where are we heading next? *Expert Opin Drug Discov.* 10 (9), 975–989 (2016).
7. Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I. & Ben Jelloul, M. Towards rationalization of crystallization screening for small- To medium-sized academic laboratories: The PACT/JCSG+ strategy. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 61 (10), 1426–1431 (2005).
8. Hartshorn, M. J., Murray, C. W., Cleasby, A., Frederickson, M., Tickle, I. J. & Jhoti, H. Fragment-Based Lead Discovery Using X-ray Crystallography. *J. Med. Chem.* 48 (2), 403–413 (2005).
9. Pahl, A., Lakemeyer, M., Vielberg, M. T., Hackl, M. W., Vomacka, J., Korotkov, V. S., *et al.* Reversible Inhibitors Arrest ClpP in a Defined Conformational State that Can Be Revoked by ClpX Association. *Angew. Chemie Int. Ed.* 54 (52), 15892–15896 (2015).
10. Lee, B. G., Kim, M. K. & Song, H. K. Structural insights into the conformational diversity of ClpP from *Bacillus subtilis*. *Mol. Cells.* 32 (6), 589–595 (2011).
11. Szyk, A. & Maurizi, M. R. Crystal structure at 1.9Å of *E. coli* ClpP with a peptide covalently bound at the active site. *J. Struct. Biol.* 156 (1), 165–174 (2006).

Chapter 5

Conclusions and outlooks

Overview

Due to the low number of new drugs entering the clinic and the increasing number of multi drug resistance bacteria, alternative strategies are urgently required to identify new antibiotic treatments, especially for Gram-negative organisms. In the work presented in this thesis, covering ClpP and using *Escherichia coli* as a model organism, the aim was to extend the small but still important panorama of antibacterial drug research focussed on non-classical Gram-negative targets. The strategy adopted was multidisciplinary, using competencies in biochemistry, biophysics, microbiology, computational chemistry which was greatly supported by medicinal chemistry input from across the INTEGRATE-ITN consortium. Following the development of a bioactivity assay for monitoring ClpP enzymatic function, new compounds were identified and profiled for their potency, selectivity and, toxicity versus mammalian cell. Applying further biochemical and biophysical methods, computational chemistry and microbiology techniques, selected compounds were profiled and predictions made of the binding interactions with ClpP. In a still ongoing part of the project, the production of the crystal structure of the target protein with active compounds is studied. This will help to clarify the binding mode of the protein-ligand interaction and assist in the process of further optimising the hits in future work.

A key finding presented in this thesis is the discovery of a new class of *Escherichia coli* modulating compounds, alpha-amino diaryl phosphonates, that inhibit ClpP proteolysis activity (Table 5, 2.3 Results). In addition, new ClpP modulating activities are reported for several already-approved drugs and commercially available bioactive compounds (Figure 1 and Table 1, 3.3 Results). Taken together, this work has increased the number of active molecules and warheads inhibiting the activity of a Gram-negative caseinolytic protease P, and represents a wider range of starting points for future lead optimisation studies. In addition to the reported compounds, an important finding of this work is the confirmation of reports by 2015 by Robinson *et al.* of phenotypic *clpP*-mediated effects in *Escherichia coli* growth, under specific stress conditions. These effects were confirmed in an improved medium throughput format 96 well screening assay which can be further adapted to the study of other compound synergism-related effects under stress conditions (Figure 5 of the paragraph 2.3 Results, Figure 5 of the paragraph 3.3 Results and Figure S2, 2.8 Supplementary data). Also, a new and different effect on bacterial growth has been shown, for mutant strains with ClpP deleted, when cultured under minimal resources conditions (Figure 5 of the paragraph 2.3 Results).

Identification of a new class of inhibitors of ClpP

In a first part of the research, the starting point was the design and synthesis, by the Medicinal Chemistry group of the University of Antwerp, of a focussed compound library. The set contained some 200 compounds, mainly diarylphosphonates, plus diverse warheads (arylphosphonates and nitriles) designed specifically to inhibit serine proteases. In addition, a set of newly synthesised and commercially sourced compounds representing previously described ClpP inhibitors, in particular, Z-LY-CMK and boronates (Lys-boroMet) were used for assay validation and benchmarking. In the course of multiple rounds of compound evaluation in the ClpP activity assay, further optimised compounds were designed and synthesized based on hydrophilic exploration and analysis of the most active compound moieties. Subsequently, 14 active compounds were highlighted, 4 with sub-micromolar IC₅₀ values (78, 85, 92 and 107) and 10 with IC₅₀ 's within the low micromolar range (Table 5, 2.3 Results). Surface Plasmon Resonance studies show reversible binding between the protein and tested ligands (Figure 4, 2.3 Results). The selectivity of the compounds was tested against chymotrypsin, a serine protease, and cytotoxicity experiments were performed in 3 human cell lines, followed by testing for antibacterial effects against several bacteria strains (*Escherichia coli* wild type and *clpP* deleted mutant strain) in the presence and absence of nitric oxide stress. Taking into account all the on and off-target profile results, 2 compounds were highlighted (92 and 85), as the more interesting for further study. The first (92) for its good in-vitro activity whilst showing relatively clear effects on human cells (Table 5, 2.3 Results) and the second (85) for the effect in the nitric oxide stress experiment (Figure 5, 2.3 Results). For those two compounds, a computational analysis of the most favorable predicted poses was performed. This evaluation indicated that the main moiety which characterized the respective compounds (aniline for the compound 85 and benzamide for the compound 92) were deeply embedded in the proposed binding pocket (Figure 6, 2.3 Results), suggesting a plausible mechanism for the underlying mechanisms of inhibition.

Identification of marketed drugs as inhibitors of ClpP activity

In an analogy to a drug repurposing approach, around 2380 approved drugs and compounds under investigation with defined bioactivities were screened in the biochemical assay under automated conditions. for activity against *Escherichia coli* ClpP. A group of 5 molecules was identified with IC₅₀ values between 0.04 and 31 μM (Table 1, 3.3 Results), all were subsequently shown to bind covalently ClpP in surface plasmon resonance (Figure 3, 3.3 Results). Among those are listed 3 anticancer (cDPCP, cisplatinium, and bortezomib), one proteasome inhibitor and serine protease inhibitor (3,4-dichloroisocoumarin) and an approved antibiotic (cefmetazole). Not surprisingly, 3,4-dichloroisocoumarin was found here since its known activity in serine proteases, as well bortezomib, reported previously to inhibit *Micobacterium tuberculosis* ClpP, was the most active compound in a biochemical assay with nanomolar IC₅₀. Interestingly, cefmetazole is usually reported to interfere with the stages of cell wall trans-peptidation and cross-linking, was found here to have also an intracellular target. The data obtained in the repurposing study would suggest a significant specificity for *Escherichia coli* ClpP with cefmetazole since none of the other 45 compounds screened belonging to the same

or related compound class (cephems and penem) was inhibitory to the same degree. (Table S2, 3.8 Supplementary data). In the nitric stress induced experiments, 3 compounds which might have *clpP*-related effect in bacteria growth were cDPCP, cisplatinum and 3,4-dichloroisocoumarin (Figure 5, 3.3 Results).

Confirmation of the role of ClpP in stress response and its potential as a targeting

In this work, we have validated the phenotypic growth properties of *clpP* defective mutant *Escherichia coli*, and compared it to the growth of an isogenic wild type strain, in minimal media (representing a resource shortage condition) and in presence and absence of nitric oxide donor induced stress (Robinson *et al.*). The phenotypic assay was further optimised and adapted to a 96 well plate suitable for compound screening campaigns. For reasons of availability and access, different strains (WT *Escherichia coli* BW25-113 and its isogenic mutant for ClpP, *Escherichia coli* JW0427-1) and a different minimal media (M9) were used compared to the previous report. Similarly, the protocol was modified to an optimized screening format. However, similar retardation effects were seen in WT versus ClpP mutant growth under stress conditions, albeit at a lower magnitude. Under the conditions of this study, in nitric oxide stress conditions, the mutant strain starts to re-grow with a delay time between 1 and 2.5 hours. This time is shorter compared to the one reported by Robinson *et al.* (around 4 to 5 hours) but was reproducible and statistically significant as shown in Figure S2 of Chapter 2 (2.8 Supplementary data). Moreover, in the same set of experiments, ClpP achieved a significantly lower level of OD₆₀₀ during the stationary phase compare to the wild type which was independent of the presence of nitric oxide donor induced stress (data shown in Figure 5 of the paragraph 2.3 Results), demonstrating the existence of additional other *clpP*-related phenotypic effects.

The research activities are ongoing with respect to the crystallography studies are being carried out in collaboration with EMBL Grenoble, Marquez group, which the goal of producing and solving crystal structures of *Escherichia coli* ClpP with selected compounds from the study in order to could obtain information about the interactions and actual binding pose for further development of the molecules. At present, a crystal structure without the presence of compounds have been solved and an optimised co-crystallization method is under development. Furthermore, after the conclusion of this study, a fragment screening campaign with co-crystallization strategy will commence in order to generate additional starting points for hit finding and lead optimisation.

Appendix

Part 1. Chemistry procedures – Chapter 2

Reagents were obtained from commercial sources and were used without further purification. Characterization of all compounds was done with ^1H and ^{13}C NMR and mass spectrometry. ^1H and ^{13}C NMR spectra were recorded on a 400 MHz Bruker Avance III Nanobay spectrometer with Ultrashield working at 400 MHz and 100 MHz respectively; and analyzed by use of MestReNova analytical chemistry software. Chemical shifts are in ppm, and coupling constants are in hertz (Hz). The UPLC (Ultra Performance liquid chromatography), used to quantify the purity of the products was an ACQUITY UPLC H-Class System with a TUV detector Waters coupled to a MS detector Waters QDa. An Acquity UPLC BEH C18 1.7 μm (2.1 x 50 mm) column was used and as eluent a mixture of 0.1% FA in H_2O , 0.1% FA in MeCN, H_2O and MeCN. The wavelengths for UV detection were 254 nm and 214 nm. Key target compounds for the activity were analysed by High Resolution Mass: 10 μL of each sample (conc. = 10^{-5} M) was injected using the CapLC system (Waters, Manchester, UK) and electrosprayed using a standard electrospray source. Samples were injected with an interval of 5 min. Positive ion mode accurate mass spectra were acquired using a Q-TOF II instrument (Waters, Manchester, UK). The MS was calibrated prior to use with a 0.2% H_3PO_4 solution. The spectra were lock mass corrected using the known mass of the nearest H_3PO_4 cluster. Where necessary, flash column chromatography was performed on a Biotage ISOLERA One flash system equipped with an internal variable dual wavelength diode array detector (200–400 nm). For normal phase purifications SNAP cartridges (4 – 100 g, flow rate of 10 – 100 mL/min) were used, and reverse phase purifications were done making use of KP-C18 cartridges (4 - 30 g, flow rate of 10 – 50 mL/min). Dry sample loading was done by self-packing samplet cartridges using Celite 545. Gradients used varied for each purification.

The following sections comprise the synthetic procedures and analytical data for all compounds reported in this manuscript. Every reaction was performed under N_2 atmosphere if not stated otherwise. Several synthetic procedures that were used in the preparation of intermediates and final products are summarized here as “General Procedures”. Target compounds were obtained with a purity >95% and as amorphous solids, unless stated otherwise.

General Procedure A. K_2CO_3 (3 eq) was added to a solution of the selected **aromatic alcohol** (1 eq) in anhydrous DMF (1.5 M) and the reaction mixture was stirred at rt for 30 min. Benzyl bromide (1.05 eq) was added dropwise to the reaction mixture, that was left stirring for 4 h at rt. The reaction mixture was quenched with H_2O and extracted with EtOAc. The combined EtOAc were washed with H_2O , brine, dried over MgSO_4 , filtered and the solvent was evaporated in vacuo to yield the corresponding **protected alcohol**.

General Procedure B. Dess-Martin periodinane (1.2 eq) was added portionwise to a stirred solution of the selected **primary alcohol** (1 eq) in anhydrous DCM (0.2 M) at 0 °C. The mixture was stirred at rt for 4 h and then the solvent was evaporated in vacuo. The crude was purified by flash column chromatography (SiO₂, EtOAc in heptane, 0/100 to 100/0). The desired fractions were collected and concentrated to yield the corresponding **aldehyde**.

General Procedure C. Selected **aldehyde** (1 eq), benzyl carbamate (if not stated otherwise) (1 eq) and triphenyl phosphite (if not stated otherwise) (1.1 eq) were dissolved in anhydrous DCM (0.3 M). Then, copper(II) triflate (0.1 eq) was added and the mixture was stirred at rt for 16 h. Then, solvent was evaporated and the residue dissolved in the minimum amount of MeOH. The solution was kept at -20 °C for 48 h and then filtrated. When precipitation did not succeed, the crude was purified by flash column chromatography (SiO₂, EtOAc in heptane, 0/100 to 100/0) and if still not pure, by reverse phase column chromatography (C18, MeOH in H₂O, 0/100 to 100/0). The desired fractions were collected and concentrated to yield the corresponding **-amino diarylphosphonate**□ as a racemic mixture.

General Procedure D. To a stirred solution of the selected **protected alcohol** (1 eq) and pentamethylbenzene (3 eq) in anhydrous DCM (0.3 M) was added boron trichloride (1 M in hexanes) (2 eq) dropwise at -78 °C. After 15 min, the reaction was quenched with CHCl₃:MeOH (10:1, 1 mL) at -78 °C, and the resulting mixture was allowed to reach rt. The organic solvents were evaporated in vacuo. The residue was purified by flash column chromatography (SiO₂, EtOAc in heptane 0/100 to 100/0) and then by reverse column chromatography (C18, MeOH in H₂O 0/100 to 100/0). The desired fractions were then collected and evaporated to yield the corresponding **deprotected alcohol**.

General Procedure E. To a solution of the selected **acid** (1 eq) in anhydrous DCM (0.3 M) at 0 °C was added 4-methylmorpholine (1.2 eq). This was followed by dropwise addition of isobutyl chloroformate (1.2 eq) over 20 min. After 30 min of stirring at 0 °C, NH₃ (25%, aq. sol.) (6 eq) was added portionwise over 5 min. The reaction was stirred for 16 h at rt and then the DCM was evaporated in vacuo. The remaining solution was extracted with EtOAc, washed with citric acid citric acid (5% aq. sol.), NaHCO₃ (sat. sol.) and brine, dried over Na₂SO₄, filtered and the solvents were evaporated in vacuo to yield the corresponding **amide**.

General Procedure F. A solution of Burgess reagent (2.1 eq) in anhydrous DCM (0.3 M) was added over a suspension of the corresponding **amide** (1 eq) in anhydrous DCM (0.3 M) and the reaction mixture was stirred for 24 h. The reaction mixture was washed with AcOH (1% aq. sol.), brine, dried over Na₂SO₄, filtered and the solvents were evaporated in vacuo. The residue was purified by flash column chromatography (SiO₂, EtOAc in heptane 0/100 to 100/0). The desired fractions were collected and concentrated to yield the corresponding **nitrile**.

General Procedure G. Hydrochloric acid (4 M in dioxane) (20 eq) was added dropwise to a solution of the selected **protected amine** (1 eq) in anhydrous MeOH (0.1 M) at 0 °C. The reaction mixture was stirred at rt for 16 h. The mixture was concentrated. The solid was then dissolved in a mixture of Na₂CO₃ (10% aq. sol.). The free salt was extracted with EtOAc and the combined organic layers were then acidified with HCl (2 M) until pH = 1 to get the hydrochloric salt again. The organic layer was further extracted with HCl and the combined aqueous layers evaporated. The excess of HCl was removed by coevaporation with toluene. In case of final compounds, the crude was purified by reverse column chromatography (18C, MeOH in H₂O, 0/100 to 100/0). The desired fractions were then collected and evaporated to yield the corresponding **deprotected amine** as a hydrochloride salt.

General Procedure H. Selected **Boc-protected compound** (1 eq) was dissolved in anhydrous DCM (0.02 M) and TFA (100 eq) was added and the solution was stirred for 1 h at rt. The solvents were evaporated in vacuo and the mixture was co-evaporated with heptane to yield corresponding **deprotected amine** compound as a TFA salt.

General Procedure I. To a solution of the selected **amine** (1 eq) in anhydrous DCM (0.04 M) was added Et₃N (3 eq) followed by *N,N'*-bis-Boc-1-guanylpyrazole (2 eq). The reaction was stirred at rt for 48 h. After this time, the solvent was evaporated in vacuo and the crude was purified by flash column chromatography (SiO₂, EtOAc in heptane 0/100 to 100/0) to yield the corresponding protected guanidine.

General Procedure J. Selected **acid chloride** (1.2 eq) was added dropwise to a solution of the selected **aniline** (1 eq) and DIPEA (1.5 eq) in anhydrous DCM (0.02 M) and the reaction mixture was stirred for 2 h at rt. Then, the reaction was quenched with HCl (1 M). This mixture was extracted with DCM, combined organic layers were washed with brine, dried over Na₂SO₄, filtered and the solvent was evaporated in vacuo. The crude was then purified by flash column chromatography (SiO₂, EtOAc in heptane: 20/80 to 80/20). The desired fractions were collected and concentrated to yield the corresponding **carbamate**.

General Procedure K. Zinc was first purified by stirring commercial Zn dust with HCl (2% aq. sol.) for 1 min. The acid was removed by filtration, and the Zn was washed with HCl (2% aq. sol.), distilled H₂O, EtOH, and finally with Et₂O. Then, selected **nitrobenzyl compound** (1 eq) was dissolved in mixture of THF (0.03 M) and NH₄Cl (sat. aq. sol.) (0.03 M) and cooled to 0 °C. The mixture was treated with the pre-treated Zn (5 eq) at vigorous stirring. The reaction mixture was stirred at rt for 1 h. The reaction mixture was filtered through celite while rinsing with THF. The mixture was extracted with EtOAc, washed with brine, dried over Na₂SO₄ and concentrated. When conversion was not complete, the crude was purified by reverse phase column chromatography (C18, MeOH in H₂O, 0/100 to 100/0). The desired fractions were collected and concentrated to afford the corresponding **aniline**.

General Procedure L. A mixture of the selected **cyanophenyl compound** (1 eq), hydroxylammonium chloride (2 eq) and DIPEA (2 eq) in EtOH (0.05 M) was heated to 80 °C for 48 h. The crude was filtrated, the filtrate was evaporated and the crude was dissolved in MeCN (0.1 M). Acetyl ether (3 eq) was added and the reaction was stirred at rt for 1 h. Then, the crude was concentrated, dissolved in MeOH and kept at - 20 °C for 16 h. The solid was filtered and rinsed with cold MeOH, the filtrate was concentrated to yield the corresponding ***N*-acetoxy carbamimidoyl phenyl compound**.

General Procedure M. The selected ***N*-acetoxy carbamimidoyl phenyl compound** (1 eq) was dissolved in AcOH (0.03 M) and wet Pd(II)/C 10 wt. % (0.1 eq) was added. The reaction mixture was stirred at rt under H₂ atmosphere (1 atm) for 24 h. Then, the palladium was filtrated off through a pad of celite from the mixture and the solvent was evaporated in vacuo. The crude was dissolved in MeOH and kept at - 20 °C for 16 h. The solid was filtered and washed with cold MeOH. Then, the solid was purified by reverse phase column chromatography (C18, MeOH in H₂O, 0/100 to 100/0). The desired fractions were collected and concentrated to yield the corresponding **aromatic amidine**.

2-(4-(Benzyloxy)phenyl)ethan-1-ol (1). General procedure **A** with 2-(4-hydroxyphenyl) ethanol (2.00 g, 14.5 mmol) to yield 2-(4-(benzyloxy)phenyl)ethanol (2.75 g, 12.03 mmol, 83% yield). ¹H NMR (400 MHz, CDCl₃) δ: 7.46 - 7.28 (m, 5H), 7.16 - 7.11 (m, 2H), 6.96 - 6.89 (m, 2H), 5.04 (s, 2H), 3.81 (t, *J* = 6.5 Hz, 2H), 2.80 (t, *J* = 6.5 Hz, 2H). MS (ESI) *m/z* 211.0 [M-OH]⁺.

2-(3-(Benzyloxy)phenyl)ethan-1-ol (2). General procedure **A** with 2-(3-hydroxyphenyl)-ethanol (800 mg, 5.79 mmol) to yield 2-(3-(benzyloxy)phenyl)ethanol (1.25 g, 5.47 mmol, 94% yield) as a white solid. MS (ESI) *m/z* 211.0 [M-OH]⁺.

2-(*p*-Tolyl)acetaldehyde (3). General procedure **B** with 2-(4-methylphenyl) ethanol (800 mg, 5.87 mmol) to yield 2-(*p*-tolyl)acetaldehyde (580 mg, 4.32 mmol, 74% yield) as a colourless oil. No ionization found. ¹H NMR (400 MHz, CDCl₃) δ: 9.73 (t, *J* = 2.5 Hz, 1H), 7.18 (d, *J* = 8.0 Hz, 2H), 7.12 - 7.10 (m, 2H), 3.65 (d, *J* = 2.5 Hz, 2H), 2.35 (s, 3H).

2-(4-Methoxyphenyl)acetaldehyde (4). General procedure **B** with 4-methoxybenzeneethanol (600 mg, 3.94 mmol) to yield 2-(4-methoxyphenyl)acetaldehyde (292 mg, 1.94 mmol, 49% yield) as a colourless oil. ¹H NMR (400 MHz, CDCl₃) δ: 9.72 (t, *J* = 2.5 Hz, 1H), 7.17 - 7.10 (m, 2H), 6.94 - 6.88 (m, 2H), 3.81 (s, 3H), 3.63 (d, *J* = 2.5 Hz, 2H). No ionization found.

2-(Naphthalen-2-yl)acetaldehyde (5). General procedure **B** with 2-(naphthalen-1-yl)ethanol (100 mg, 0.58 mmol) to yield 2-(naphthalen-2-yl)acetaldehyde (56 mg, 0.33 mmol, 57% yield) as a colourless oil. ¹H NMR

(400 MHz, CDCl₃) δ : 9.80 (t, J = 2.5 Hz, 1H), 7.94 - 7.83 (m, 3H), 7.60 - 7.40 (m, 4H), 4.12 (d, J = 2.5 Hz, 2H). No ionization found.

2-(4-Fluorophenyl)acetaldehyde (6). General procedure **B** with 2-(4-fluorophenyl)-ethanol (0.89 mL, 7.31 mmol) to yield 2-(4-fluorophenyl)acetaldehyde (545 mg, 3.95 mmol, 55% yield) as a colourless oil. ¹H NMR (400 MHz, CDCl₃) δ : 9.75 (t, J = 2.0 Hz, 1H), 7.18 (m, 2H), 7.06 (m, 2H), 3.68 (d, J = 2.0 Hz, 2H). No ionization found.

2-(4-(Trifluoromethyl)phenyl)acetaldehyde (7). General procedure **B** with 2-(4-fluorophenyl)-ethanol (200 mg, 1.26 mmol) to yield 2-(4-(trifluoromethyl)phenyl)acetaldehyde (111 mg, 0.59 mmol, 56% yield) as a colourless oil. No ionization found. ¹H NMR (400 MHz, CDCl₃) δ : 9.79 (t, J = 2.0 Hz, 1H), 7.63 (d, J = 8.0 Hz, 2H), 7.34 (d, J = 8.0 Hz, 2H), 3.79 (d, J = 2.0 Hz, 2H).

2-(4-(Benzyloxy)phenyl)acetaldehyde (8). General procedure **B** with 2-(4-(benzyloxy)phenyl)ethanol (**1**) (2.75 g, 12.0 mmol) to yield 2-(4-(benzyloxy)phenyl)acetaldehyde (2.03 g, 8.99 mmol, 75% yield) as a colourless oil. ¹H NMR (400 MHz, CDCl₃) δ : 9.73 (t, J = 2.5 Hz, 1H), 7.51 - 7.34 (m, 5H), 7.22 - 7.10 (m, 2H), 7.07 - 6.98 (m, 2H), 5.09 (s, 2H), 3.63 (d, J = 2.5 Hz, 2H). No ionization found.

2-(3-(Benzyloxy)phenyl)acetaldehyde (9). General procedure **B** with 2-(3-(benzyloxy)phenyl)ethanol (**2**) (1.25 g, 5.78 mmol) to yield 2-(3-(benzyloxy)phenyl)acetaldehyde (789 mg, 3.49 mmol, 64% yield) as a colourless oil. ¹H NMR (400 MHz, CDCl₃) δ : 9.73 (t, J = 2.5 Hz, 1H), 7.41 (tdd, J = 7.5, 7.0, 1.5 Hz, 4H), 7.36 - 7.27 (m, 1H), 6.95 - 6.90 (m, 2H), 6.86 - 6.80 (m, 2H), 5.07 (s, 2H), 3.66 (d, J = 2.5 Hz, 2H). No ionization found.

Benzyl (1-(diphenoxyphosphoryl)-2-phenylethyl)carbamate (10). Procedure and characterization consistent with previously reported data.⁴⁶

Benzyl (1-(diphenoxyphosphoryl)-2-(*p*-tolyl)ethyl)carbamate (11). General procedure **C** with 2-(*p*-tolyl)acetaldehyde (**3**) (580 mg, 4.32 mmol), to give benzyl (1-(diphenoxyphosphoryl)-2-(*p*-tolyl)ethyl)carbamate (1.18 g, 2.36 mmol, 55% yield) as an off-white solid. ¹H NMR (400 MHz, CDCl₃) δ : 7.36 - 7.26 (m, 5H), 7.26 - 7.20 (m, 7H), 7.15 (dd, J = 16.0, 8.0 Hz, 4H), 7.08 (d, J = 8.0 Hz, 5H), 5.32 - 5.10 (m, 5H), 5.09 - 4.87 (m, 1H), 4.87 - 4.73 (m, 2H), 3.38 (ddd, J = 14.5, 10.0, 4.5 Hz, 1H), 3.09 - 2.86 (m, 1H), 2.32 (s, 3H). ¹³C NMR (100 MHz, CDCl₃) δ : 155.7, 150.3, 150.1, 136.7, 136.3, 132.8, 130.0, 129.8, 129.4, 129.3, 128.6, 128.2, 128.1, 125.6, 125.4, 120.8, 120.6, 7.2, 49.5 (d, C₁-P, J_{CP} = 158.0 Hz), 35.7, 21.25. MS (ESI) m/z 502.1 [M+H]⁺. MP = 114-116 °C

Benzyl (1-(diphenoxyphosphoryl)-2-(4-methoxyphenyl)ethyl)carbamate (12). General procedure **C** with 2-(4-methoxyphenyl)acetaldehyde (**4**) (269 mg, 1.19 mmol) to yield benzyl (1-(diphenoxyphosphoryl)-2-(4-

methoxyphenyl) ethyl)carbamate (354 mg, 0.68 mmol, 34% yield). ^1H NMR (400 MHz, CDCl_3) δ : 7.43 - 6.94 (m, 17H), 6.81 (d, $J = 8.5$ Hz, 2H), 5.18 (d, $J = 10.5$ Hz, 1H), 5.11 - 4.86 (m, 2H), 4.76 (dtd, $J = 15.0, 10.5, 4.5$ Hz, 1H), 3.78 (s, 3H), 3.35 (ddd, $J = 14.5, 10.0, 4.5$ Hz, 1H), 2.98 (dt, $J = 14.5, 10.0$ Hz, 1H). ^{13}C NMR (100 MHz, CDCl_3) δ : 158.7, 155.7, 150.3, 150.1, 136.2, 130.5, 130.0, 129.8, 128.6, 128.3, 128.1, 127.9, 127.8, 125.6, 125.4, 120.8, 120.8, 120.6, 120.5, 114.1, 67.3, 55.3, 49.5 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 157.5$ Hz), 35.3. MS (ESI) m/z 518.2 $[\text{M}+\text{H}]^+$.

Benzyl (benzofuran-5-yl(diphenoxyphosphoryl)methyl)carbamate (13). General procedure C with 1-benzofuran-5-carbaldehyde (500 mg, 3.42 mmol) to yield benzyl (benzofuran-5-yl(diphenoxyphosphoryl)methyl)carbamate (100 mg, 0.95 mmol, 6% yield). ^1H NMR (400 MHz, CDCl_3) δ : 7.30-6.73 (m, 20H), 5.88 (br s, 1H), 5.62 (m, 1H), 5.10 (m, 2H). MS (ESI) m/z 536.0 $[\text{M}+\text{Na}]^+$.

Benzyl (1-(diphenoxyphosphoryl)-2-(naphthalen-2-yl)ethyl)carbamate (14). General procedure C with 2-(naphthalen-1-yl)acetaldehyde (**5**) (56 mg, 0.33 mmol) to give benzyl (1-(diphenoxyphosphoryl)-2-(naphthalen-2-yl)ethyl)carbamate (51 mg, 0.10 mmol, 29% yield). ^1H NMR (400 MHz, CDCl_3) δ : 8.10 - 8.03 (m, 1H), 7.88 (dd, $J = 6.5, 3.0$ Hz, 1H), 7.78 (d, $J = 8.0$ Hz, 1H), 7.56 - 7.47 (m, 2H), 7.42 - 6.99 (m, 17H), 5.75 (d, $J = 10.5$ Hz, 1H), 5.08 - 4.87 (m, 3H), 3.95 (ddd, $J = 14.5, 8.0, 4.0$ Hz, 1H), 3.44 (dt, $J = 14.5, 10.5$ Hz, 1H). ^{13}C NMR (100 MHz, CDCl_3) δ : 155.8, 150.4, 150.1, 136.2, 134.0, 132.0, 130.0, 129.8, 129.1, 128.5, 128.1, 128.0, 127.9, 127.8, 126.6, 125.8, 125.6, 125.3, 123.2, 120.7, 120.5, 67.0, 49.1 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 158.5$ Hz), 33.12. MS (ESI) m/z 538.1 $[\text{M}+\text{H}]^+$.

Benzyl (1-(diphenoxyphosphoryl)-2-(4-fluorophenyl)ethyl)carbamate (15). General procedure C with 2-(4-fluorophenyl)acetaldehyde (**6**) (545 mg, 3.95 mmol), to give benzyl (1-(diphenoxyphosphoryl)-2-(4-fluorophenyl)ethyl)carbamate (1.39 g, 2.75 mmol, 70% yield) as an off-white solid. ^1H NMR (400 MHz, CDCl_3) δ : 7.51 - 7.01 (m, 17H), 6.87 (dt, $J = 17.0, 8.0$ Hz, 2H), 5.30 (d, $J = 10.5$ Hz, 1H), 5.17 - 4.85 (m, 2H), 4.75 (dtd, $J = 15.0, 10.5, 4.5$ Hz, 1H), 3.37 (ddd, $J = 14.0, 9.0, 4.5$ Hz, 1H), 3.00 (dt, $J = 14.5, 10.0$ Hz, 1H). ^{13}C NMR (100 MHz, CDCl_3) δ : 162.1, 155.8, 150.3, 150.0, 136.2, 131.0, 130.9, 130.0, 129.9, 128.6, 128.4, 128.1, 125.7, 125.5, 120.8, 120.7, 120.5, 120.5, 115.7, 115.5, 67.4, 49.4 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 158.5$ Hz), 35.4. MS (ESI) m/z 506.2 $[\text{M}+\text{H}]^+$. MP = 133-135 °C

Benzyl (1-(diphenoxyphosphoryl)-2-(4-(trifluoromethyl)phenyl)ethyl) carbamate (16). General procedure C with 2-(4-(trifluoromethyl)phenyl)acetaldehyde (**7**) (111 mg, 0.59 mmol) to give benzyl (1-(diphenoxyphosphoryl)-2-(4-(trifluoromethyl) phenyl)ethyl)carbamate (91 mg, 0.16 mmol, 28% yield) as an off-white solid. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ : 8.24 (d, $J = 9.5$ Hz, 1H), 7.66 (d, $J = 8.0$ Hz, 2H), 7.56 (d, $J = 8.0$ Hz, 2H), 7.45 - 7.34 (m, 4H), 7.30 - 7.09 (m, 11H), 4.95 (m, 2H), 4.66 - 4.51 (m, 1H), 3.43 - 3.36 (m, 1H), 3.10 (m, 1H). ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$) δ : 155.9, 150.1, 149.8, 142.1, 137.0, 130.1, 130.0, 129.9, 128.3, 127.8, 127.3, 127.12 (q, $\text{C}-\text{CF}_3$, $J_{\text{CF}} = 31.5$ Hz), 125.5, 125.3, 125.1 (q, $\text{C}-\text{C}-\text{CF}_3$, $J_{\text{CF}} = 3.5$ Hz), 124.5

(q, CF_3 , $J_{CF} = 272.0$ Hz), 120.7, 120.7, 120.5, 120.4, 65.6, 49.6 (d, C_1-P , $J_{CP} = 159.5$ Hz), 34.0. MS (ESI) m/z 556.0 $[M+Na]^+$, (95%). HRMS: Calc: 556.15 Found: 556.1481 $[M+H]^+$.

Benzyl (1-(diphenoxyphosphoryl)-3-(methylthio)propyl)carbamate (17). Procedure and characterization consistent with previously reported data.⁴⁷

Benzyl ((diphenoxyphosphoryl)(6-hydroxynaphthalen-2-yl)methyl)carbamate (18). General procedure C with 6-hydroxy-2-naphthaldehyde (289 mg, 1.68 mmol), to give benzyl ((diphenoxyphosphoryl)(6-hydroxynaphthalen-2-yl)methyl)carbamate (208 mg, 0.39 mmol, 23% yield) as a white solid. 1H NMR (400 MHz, $DMSO-d_6$) δ : 9.85 (s, 1H), 8.98 (d, $J = 10.0$ Hz, 1H), 8.00 (s, 1H), 7.78 - 7.63 (m, 3H), 7.42 - 7.26 (m, 9H), 7.21 - 7.05 (m, 6H), 6.98 (d, $J = 8.4$ Hz, 2H), 5.78 - 5.59 (m, 1H), 5.10 (dd, $J = 35.0, 12.5$ Hz, 2H). ^{13}C NMR (100 MHz, $DMSO-d_6$) δ : 156.9, 150.9, 137.6, 135.2, 130.8, 130.4, 129.3, 128.9, 128.5, 128.2, 127.4, 127.2, 126.2, 121.3, 120.1, 109.5, 67.1, 53.9 (d, C_1-P , $J_{CP} = 157.5$ Hz). MS (ESI) m/z 540.1 $[M+H]^+$. HRMS: Calc: 540.16 Found: 540.1584 $[M+H]^+$. MP = 166-168 °C.

Benzyl (2-(4-(benzyloxy)phenyl)-1-(diphenoxyphosphoryl)ethyl)carbamate (19) General procedure C with 2-(4-(benzyloxy)phenyl)acetaldehyde (**8**) (1.79 g, 7.89 mmol), to give benzyl (2-(4-(benzyloxy)phenyl)-1-(diphenoxyphosphoryl)ethyl) carbamate (3.55 g, 5.99 mmol, 76% yield). 1H NMR (400 MHz, $CDCl_3$) δ : 7.46 - 7.27 (m, 12H), 7.25 - 7.02 (m, 10H), 6.89 (d, $J = 8.5$ Hz, 2H), 5.22 (d, $J = 10.5$ Hz, 1H), 5.03 (s, 2H), 5.02 (s, 2H), 4.83 - 4.70 (m, 1H), 3.35 (ddd, $J = 14.5, 10.0, 4.5$ Hz, 1H), 2.99 (dt, $J = 14.5, 10.0$ Hz, 1H). ^{13}C NMR (100 MHz, $CDCl_3$) δ : 158.0, 155.8, 150.3, 150.1, 137.1, 136.2, 130.5, 130.0, 129.7, 128.7, 128.6, 128.3, 128.1, 127.6, 125.6, 125.4, 120.8, 120.5, 115.1, 70.1, 67.3, 49.6 (d, C_1-P , $J_{CP} = 157.5$ Hz), 35.3. MS (ESI) m/z 594.2 $[M+H]^+$.

Benzyl (2-(3-(benzyloxy)phenyl)-1-(diphenoxyphosphoryl)ethyl)carbamate (20). General procedure C with 2-(3-(benzyloxy)phenyl)acetaldehyde (**9**) (789 mg, 3.49 mmol), to give benzyl (2-(3-(benzyloxy)phenyl)-1-(diphenoxyphosphoryl)ethyl) carbamate (1.37 g, 2.31 mmol, 66% yield). 1H NMR (400 MHz, $CDCl_3$) δ : 7.48 - 7.06 (m, 21H), 6.96 - 6.85 (m, 3H), 5.33 (d, $J = 10.5$ Hz, 1H), 5.06 (s, 2H), 5.00 (s, 2H), 4.89 - 4.74 (m, 1H), 3.46 - 3.36 (m, 1H), 3.13 - 3.01 (m, 1H). MS (ESI) m/z 594.2 $[M+H]^+$.

Benzyl (1-(diphenoxyphosphoryl)-2-(4-hydroxyphenyl)ethyl)carbamate (21). General procedure D with benzyl (2-(4-(benzyloxy)phenyl)-1-(diphenoxyphosphoryl)ethyl)carbamate (**19**) (200 mg, 0.34 mmol) to yield benzyl (1-(diphenoxyphosphoryl)-2-(4-hydroxyphenyl)ethyl)carbamate (52 mg, 0.10 mmol, 31% yield) as a white solid. 1H NMR (400 MHz, $DMSO-d_6$) δ : 9.33 (s, 1H), 8.17 (d, $J = 9.5$ Hz, 1H), 7.48 - 7.09 (m, 18H), 6.73 (t, $J = 5.5$ Hz, 2H), 5.08 - 4.94 (m, 2H), 4.53 - 4.38 (m, 1H), 3.19 (dt, $J = 14.0, 3.5$ Hz, 1H), 2.98 - 2.85 (m, 1H). ^{13}C NMR (100 MHz, $DMSO-d_6$) δ : 157.0, 151.1, 150.7, 137.9, 131.0, 130.8, 129.2, 128.6, 128.1,

126.3, 126.2, 121.6, 121.3, 116.0, 66.3, 51.2 (d, C₁-P, $J_{CP} = 156.0$ Hz), 34.1. MS (ESI) m/z 504.2 [M+H]⁺. MP = 172-174 °C

Benzyl (1-(diphenoxyphosphoryl)-2-(3-hydroxyphenyl)ethyl)carbamate (22). General procedure **D** with benzyl (2-(3-(benzyloxy)phenyl)-1-(diphenoxyphosphoryl) ethyl)carbamate (**20**) (500 mg, 0.84 mmol), to yield benzyl (1-(diphenoxyphosphoryl)-2-(3-hydroxyphenyl)ethyl)carbamate (196 mg, 0.39 mmol, 46% yield) as a white solid. ¹H NMR (400 MHz, Acetone-*d*₆) δ : 8.33 (s, 1H), 7.50 - 7.20 (m, 14H), 7.16 (t, $J = 8.0$ Hz, 1H), 7.09 (d, $J = 10.0$ Hz, 1H), 6.96 - 6.84 (m, 2H), 6.78 (dd, $J = 8.0, 2.0$ Hz, 1H), 5.05 (s, 2H), 4.76 (dddd, $J = 13.5, 12.0, 10.0, 3.5$ Hz, 1H), 3.40 (ddd, $J = 14.0, 5.0, 3.5$ Hz, 1H), 3.09 (ddd, $J = 14.0, 12.0, 8.5$ Hz, 1H). ¹³C NMR (100 MHz, Acetone-*d*₆) δ : 158.0, 156.6, 151.4, 151.14, 139.2, 137.8, 130.2, 129.9, 128.9, 128.2, 128.0, 125.8, 125.6, 121.4, 121.1, 120.9, 116.7, 114.3, 66.6, 50.8 (d, C₁-P, $J_{CP} = 158.5$ Hz), 35.6. MS (ESI) m/z 504.2 [M+H]⁺. MP = 140-142 °C

Benzyl (1-((4-acetamidobenzyl)(4-acetamidophenoxy)phosphoryl)-2-phenylethyl) carbamate (23). General procedure **C** with phenylethanal (0.31 mL, 2.65 mmol) and tris(4-acetamidophenyl) phosphite (1.40 g, 2.91 mmol, 1.1 eq) to give benzyl (1-(bis(4-acetamidophenoxy)phosphoryl)-2-phenylethyl)carbamate (368 mg, 0.61 mmol, 23% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ : 10.00 (d, $J = 3.0$ Hz, 2H), 8.15 (d, $J = 9.5$ Hz, 1H), 7.61 - 7.50 (m, 4H), 7.39 - 7.20 (m, 8H), 7.19 - 7.05 (m, 6H), 5.01 - 4.77 (m, 2H), 4.47 (tdd, $J = 14.5, 9.5, 3.0$ Hz, 1H), 3.25 (dt, $J = 7.5, 3.5$ Hz, 1H), 2.98 (ddd, $J = 13.5, 12.5, 8.0$ Hz, 1H), 2.03 (s, 6H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ : 168.9, 155.9, 145.2, 145.0, 137.2, 136.5, 129.1, 128.2, 127.6, 127.2, 126.6, 120.8, 120.5, 120.1, 65.9, 49.8 (d, C₁-P, $J_{CP} = 157.5$ Hz), 34.1, 23.9. MS (ESI) m/z 602.2 [M+H]⁺. HRMS: Calc: 602.21 Found: 602.2054 [M+H]⁺.

Benzyl (1-(hydroxy(phenoxy)phosphoryl)-2-(4-hydroxyphenyl)ethyl)carbamate (24). KOH (58 mg, 0.99 mmol, 3 eq) was added to a solution of benzyl (1-(diphenoxyphosphoryl)-2-(4-hydroxyphenyl)ethyl)carbamate (**21**) (250 mg, 0.50 mmol) in H₂O (5 mL) and 1,4-dioxane (5 mL) and the resulting mixture was stirred at rt over 16 h. The crude reaction was evaporated and HCl (1N aq. sol.) was added to form the HCl salt. The residue was purified by reverse column chromatography (C18, MeOH in H₂O 0/100 to 100/0). The desired fractions were then collected and evaporated to yield benzyl (1-(hydroxy(phenoxy)phosphoryl)-2-(4-hydroxyphenyl)ethyl)carbamate hydrochloride (47 mg, 0.10 mmol, 20% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.38 - 7.27 (m, 5H), 7.25 - 7.15 (m, 5H), 7.10 (d, $J = 8.5$ Hz, 2H), 6.76 - 6.66 (m, 2H), 5.10 - 4.91 (m, 2H), 4.32 (dd, $J = 20.0, 7.5$ Hz, 1H), 3.26 - 3.16 (m, 1H), 2.89 - 2.77 (m, 1H). ¹³C NMR (100 MHz, Methanol-*d*₄) δ : 157.2, 155.8, 150.9, 136.9, 129.9, 129.3, 128.0, 127.4, 127.0, 124.4, 120.4, 114.8, 66.1, 50.4 (d, C₁-P, $J_{CP} = 155.0$ Hz), 34.2. MS (ESI) m/z 428.2 [M+H]⁺.

Benzyl (2-(4-hydroxyphenyl)-1-(methoxy(phenoxy)phosphoryl)ethyl)carbamate (25). NH₃ (7N in MeOH) (0.09 mL, 0.60 mmol) was added to the stirred solution of benzyl (1-(diphenoxyphosphoryl)-2-(4-

hydroxyphenyl)ethyl)carbamate (**21**) (200 mg, 0.40 mmol) and NH₄Cl (32 mg, 0.60 mmol) in MeOH (4 mL). The reaction mixture was stirred at rt and for 3 days. The reaction mixture was concentrated and purified by reverse column chromatography (C18, MeOH in H₂O, 0/100 to 100/0). The desired fractions were then collected and evaporated to yield benzyl (2-(4-hydroxyphenyl)-1-(methoxy(phenoxy)phosphoryl)ethyl)carbamate (26 mg, 0.06 mmol, 15% yield) as a colourless oil. ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.73 (dd, *J* = 9.5, 6.0 Hz, 1H), 7.41 - 7.32 (m, 2H), 7.32 - 7.24 (m, 3H), 7.24 - 7.13 (m, 5H), 7.12 - 6.96 (m, 2H), 6.71 (dd, *J* = 8.5, 3.5 Hz, 2H), 4.99 (ddd, *J* = 33.0, 12.5, 7.5 Hz, 2H), 4.52 - 4.31 (m, 1H), 3.91 - 3.72 (m, 3H), 3.23 - 3.09 (m, 1H), 2.91 - 2.71 (m, 1H). ¹³C NMR (100 MHz, Methanol-*d*₄) δ: 158.3, 157.4, 151.7, 138.2, 131.1, 130.9, 129.4, 128.8, 128.4, 126.4, 121.57, 116.3, 67.5, 54.7, 51.1 (d, C₁-P, *J*_{CP} = 158.0), 35.3. MS (ESI) *m/z* 442.1 [M+H]⁺.

(S)-2-(((Benzyloxy)carbonyl)amino)-3-phenylpropanoic acid (26). Procedure and characterization consistent with previously reported data.⁴⁸

(S)-Benzyl (1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)carbamate (27). General procedure **E** with ((benzyloxy)carbonyl)tyrosine (382 mg, 1.23 mmol) to yield (*S*)-benzyl (1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)carbamate (140 mg, 0.45 mmol, 37% yield). MS (ESI) *m/z* 315.1 [M+H]⁺.

(S)-Benzyl (1-amino-1-oxo-3-phenylpropan-2-yl)carbamate (28). General procedure **E** with (*S*)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanoic acid (**26**) (100 mg, 0.33 mmol) to yield (*S*)-benzyl (1-amino-1-oxo-3-phenylpropan-2-yl)carbamate (98 mg, 0.33 mmol, 98% yield). MS (ESI) *m/z* 299.1 [M+H]⁺.

Benzyl (S)-(1-cyano-2-(4-hydroxyphenyl)ethyl)carbamate (29). General procedure **F** with (*S*)-benzyl (1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)carbamate (**27**) (140 mg, 0.45 mmol) to yield (*S*)-benzyl (1-cyano-2-(4-hydroxyphenyl)ethyl)carbamate (113 mg, 0.38 mmol, 86% yield). ¹H NMR (400 MHz, Acetone-*d*₆) δ: 8.33 (s, 1H), 7.48 - 7.32 (m, 5H), 7.25 - 7.18 (m, 2H), 6.89 - 6.78 (m, 2H), 5.13 (m, 2H), 4.88 - 4.74 (m, 1H), 3.15 (d, *J* = 7.5 Hz, 2H). ¹³C NMR (100 MHz, Acetone-*d*₆) δ: 158.3, 157.0, 138.4, 132.2, 130.0, 129.6, 129.5, 127.8, 120.5, 117.0, 116.9, 68.0, 46.2, 39.3. MS (ESI) *m/z* 297.1 [M+H]⁺.

Benzyl (S)-(1-cyano-2-phenylethyl)carbamate (30). General procedure **F** with (*S*)-benzyl (1-amino-1-oxo-3-phenylpropan-2-yl)carbamate (**28**) (460 mg, 1.54 mmol) to yield (*S*)-benzyl (1-cyano-2-phenylethyl)carbamate (323 mg, 1.15 mmol, 75% yield). ¹H NMR (400 MHz, Acetone-*d*₆) δ: 7.43 - 7.25 (m, 10H), 5.13 (s, 2H), 4.90 (dt, *J* = 8.0, 5.5 Hz, 1H), 3.30 - 3.20 (m, 2H). ¹³C NMR (100 MHz, Acetone-*d*₆) δ: 157.0, 138.4, 137.2, 131.1, 130.2, 130.0, 129.6, 129.5, 128.9, 120.4, 68.1, 45.9, 40.0. MS (ESI) *m/z* 281.1 [M+H]⁺. MP = 132-134 °C. Characterization consistent with previously reported data.⁴⁹

Diphenyl (1-amino-2-(4-hydroxyphenyl)ethyl)phosphonate hydrobromide (31). Benzyl (2-(4-(benzyloxy)phenyl)-1-(diphenoxyphosphoryl)ethyl)carbamate (**19**) (1.00 g, 1.69 mmol) was dissolved in AcOH (2 mL) and then, 33% HBr/AcOH solution (1.22 mL, 6.74 mmol, 4 eq). The reaction was performed at rt for 6 h. Then, the reaction mixture was concentrated in vacuo. The crude was purified by reverse phase column chromatography (C18, MeOH in H₂O 0/100 to 60/40). The desired fractions were collected and concentrated to yield diphenyl (1-amino-2-(4-hydroxyphenyl)ethyl)phosphonate hydrobromide (374 mg, 49% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 7.41 - 7.35 (m, 4H), 7.26 - 7.16 (m, 6H), 7.11 (d, *J* = 8.5 Hz, 2H), 6.70 (d, *J* = 8.5 Hz, 2H), 3.49 (td, *J* = 10.0, 3.5 Hz, 1H), 3.16 - 3.07 (m, 1H), 2.70 (dt, *J* = 14.0, 10.5 Hz, 1H). MS (ESI) *m/z* 369.2 [M+H]⁺.

Benzyl ((2S)-1-((1-(diphenoxyphosphoryl)-2-(4-hydroxyphenyl)ethyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (32). To a stirred solution of *N*-carbobenzyloxy-*L*-leucine (71 mg, 0.27 mmol, 1.2 eq) in MeCN (3 mL) and DMF (1 mL), 1-hydroxybenzotriazolehydrate (37 mg, 0.24 mmol, 1.1 eq) and *N,N'*-dicyclohexylcarbodiimide (92 mg, 0.44 mmol, 2 eq) were added and the solution was stirred for 10 min at rt. Then, a solution of diphenyl (1-amino-2-(4-hydroxyphenyl)ethyl)phosphonate hydrobromide (**31**) (100 mg, 0.22 mmol) and Et₃N (0.03 mL, 0.22 mmol, 1 eq) in DCM (2 mL) at 0 °C and the mixture was left stirring at rt for 16 h. Then, the precipitate was filtrated off. The solvent was evaporated in vacuo from the filtrate and the crude was purified by flash column chromatography (SiO₂, EtOAc in heptane 0/100 to 100/0) to yield benzyl ((2S)-1-((1-(diphenoxyphosphoryl)-2-(4-hydroxyphenyl)ethyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**32**) (24 mg, 18% yield) as a colourless oil. ¹H NMR (400 MHz, CDCl₃) δ: 7.39 - 7.28 (m, 12H), 7.23 - 7.12 (m, 7H), 7.06 (d, *J* = 8.0 Hz, 2H), 5.28 (d, *J* = 8.5 Hz, 1H), 5.14 (s, 2H), 4.62 (td, *J* = 9.0, 4.5 Hz, 1H), 3.62 (td, *J* = 10.5, 3.0 Hz, 1H), 3.47 - 3.37 (m, 1H), 2.93 (dt, *J* = 14.0, 10.5 Hz, 1H), 1.87 - 1.77 (m, 2H), 1.73 - 1.65 (m, 1H), 1.03 - 1.00 (m, 6H). ¹³C NMR (100 MHz, CDCl₃) δ: 172.0, 156.1, 150.4, 149.5, 136.2, 135.3, 130.5, 129.9, 128.7, 128.4, 128.3, 125.4, 121.7, 120.7, 67.3, 52.8, 50.7 (d, C₁-P, *J*_{CP} = 157.5 Hz), 41.7, 37.2, 25.0, 23.0, 22.0. MS (ESI) *m/z* 617.3 [M+H]⁺.

***Tert*-butyl (1-(diphenoxyphosphoryl)-3-(methylthio)propyl)carbamate (33).** General procedure C with 4-thiapentanal (4.40 g, 42.20 mmol) and *O-tert*-butylcarbamate (4.95 g, 42.20 mmol) to give *tert*-butyl (1-(diphenoxyphosphoryl)-3-(methylthio)propyl)carbamate (5.480 g, 12.53 mmol, 30% yield). MS (ESI) *m/z* 438.2 [M+H]⁺.

Diphenyl (1-amino-3-(methylthio)propyl)phosphonate hydrochloride (34). General procedure G with *tert*-butyl (1-(diphenoxyphosphoryl)-3-(methylthio)propyl)carbamate (**33**) (500 mg, 1.14 mmol) to yield diphenyl (1-amino-3-(methylthio)propyl)phosphonate hydrochloride (425 mg, 1.14 mmol, 99% yield). MS (ESI) *m/z* 338.2 [M+H]⁺.

Benzyl tert-butyl ((5S)-6-((1-(diphenoxyphosphoryl)-3-(methylthio)propyl)amino)-6-oxohexane-1,5-diyl)dicarbamate (35). To a stirred solution of (*R*)-2-(((benzyloxy)carbonyl)amino)-6-((*tert*-butoxycarbonyl)amino)hexanoic acid (122 mg, 0.32 mmol) in MeCN (3 mL) and DMF (1 mL), 1-hydroxybenzotriazolehydrate (53 mg, 0.35 mmol) and *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (62 mg, 0.32 mmol) were added and the solution was stirred for 10 min at rt. Then, a solution of diphenyl (1-amino-3-(methylthio)propyl)phosphonate hydrochloride (**34**) (100 mg, 0.27 mmol) and Et₃N (0.08 mL, 0.59 mmol) in MeCN (3 mL) was added at 0 °C and the mixture was left stirring at rt for 16 h. Then, the precipitate was filtrated off. The solvent was evaporated in vacuo from the filtrate and the crude was purified by flash column chromatography (SiO₂, EtOAc in heptane 0/100 to 100/0). Desired fractions were collected and concentrated to yield benzyl tert-butyl ((5S)-6-((1-(diphenoxyphosphoryl)-3-(methylthio)propyl)amino)-6-oxohexane-1,5-diyl)dicarbamate (**35**) (220 mg, 0.252 mmol, 94% yield). MS (ESI) *m/z* 700.4 [M+H]⁺.

Benzyl ((2S)-6-amino-1-((1-(diphenoxyphosphoryl)-3-(methylthio)propyl)amino)-1-oxohexan-2-yl)carbamate hydrochloride (36). General procedure **G** with benzyl tert-butyl ((5S)-6-((1-(diphenoxyphosphoryl)-3-(methylthio)propyl)amino)-6-oxohexane-1,5-diyl)dicarbamate (**35**) (250 mg, 0.36 mmol) to yield benzyl ((2S)-6-amino-1-((1-(diphenoxyphosphoryl)-3-(methylthio)propyl)amino)-1-oxohexan-2-yl)carbamate hydrochloride (**36**) (123 mg, 0.19 mmol, 54% yield) as a colourless oil. ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.52 - 7.05 (m, 15H), 5.20 - 4.93 (m, 3H), 4.32 - 4.08 (m, 1H), 2.95 - 2.74 (m, 2H), 2.74 - 2.38 (m, 2H), 2.37 - 2.12 (m, 2H), 2.11 - 1.98 (m, 3H), 1.92 - 1.54 (m, 4H), 1.54 - 1.33 (m, 2H). ¹³C NMR (100 MHz, Methanol-*d*₄) δ: 175.0, 158.3, 151.4, 138.1, 131.1, 130.9, 129.5, 129.1, 128.9, 126.8, 121.8, 121.6, 67.7, 56.3, 46.3 (d, C₁-P, *J*_{CP} = 160.0 Hz), 40.4, 32.6, 31.1, 29.5, 28.1, 23.7, 15.3. MS (ESI) *m/z* 600.3 [M+H]⁺.

(S)-2-((*Tert*-butoxycarbonyl)amino)-3-(4-hydroxyphenyl)propanoic acid (37). Di-*tert*-butyldicarbonate (1.21 g, 5.52 mmol) was added to a solution of (*S*)-(-)-Tyrosine (1.00 g, 5.52 mmol) in a mixture of dioxane (5 mL), H₂O (2.5 mL) and NaOH (1 M, 5 mL) at 0 °C and the above mixture and stirred for 6 h at rt. Then the solution was concentrated in vacuum, cooled in an ice water bath, covered with a layer of EtOAc and acidified with a dilute solution of KHSO₄ such that the solution pH 2-3. The aqueous phase was extracted with EtOAc, dried (Na₂SO₄), filtered and solvents and evaporated in vacuo to yield (*S*)-2-((*tert*-butoxycarbonyl)amino)-3-(4-hydroxyphenyl)propanoic acid (1.55 g, 5.16 mmol, 94% yield). MS (ESI) *m/z* 304.2 [M+Na]⁺.

***Tert*-butyl (*S*)-(1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)carbamate (38).** General procedure **E** with (*S*)-2-((*tert*-butoxycarbonyl)amino)-3-(4-hydroxyphenyl)propanoic acid (**37**) (1.00 g, 3.55 mmol) to yield (*S*)-*tert*-butyl (1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)carbamate (1.09 g, 3.51 mmol, 99% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 9.13 (s, 1H), 7.02 (d, *J* = 8.5 Hz, 2H), 6.64 (d, *J* = 8.5 Hz, 2H), 3.98 (dd, *J* =

9.5, 5.0 Hz, 1H), 2.82 (dd, $J = 14.0$, 4.5 Hz, 1H), 2.61 (dd, $J = 14.0$, 10.0 Hz, 1H), 1.36 - 1.26 (m, 9H). MS (ESI) m/z 303.2 [M+Na]⁺.

(S)-2-Amino-3-(4-hydroxyphenyl)propanamide hydrochloride (39). General procedure **G** with (*S*)-*tert*-butyl (1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)carbamate (**38**) (1.09 g, 3.90 mmol) to yield (*S*)-2-amino-3-(4-hydroxyphenyl)propanamide hydrochloride (804 mg, 3.71 mmol, 95% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.18 - 7.11 (m, 2H), 6.82 - 6.74 (m, 2H), 4.06 (dd, $J = 8.0$, 6.0 Hz, 1H), 3.16 (dd, $J = 14.0$, 6.0 Hz, 1H), 2.98 (dd, $J = 14.0$, 8.0 Hz, 1H). MS (ESI) m/z 181.1 [M+H]⁺.

Benzyl ((S)-1-(((S)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (40). A solution of (*S*)-2-amino-3-(4-hydroxyphenyl)propanamide hydrochloride (**39**) (804 mg, 3.71 mmol) and *N,N*-diisopropylethylamine (0.65 mL, 3.71 mmol) in DCM (1 mL) was added dropwise to a solution of Z-Leu-OSu (1.61 g, 4.45 mmol) in DCM (10 mL) at 0 °C. The reaction mixture was stirred at rt for 16 h. The mixture was concentrated, dissolved in EtOAc, washed with NaHCO₃ sat. and HCl (1 M), dried (Na₂SO₄), filtered and solvents concentrated in vacuo to yield benzyl ((*S*)-1-(((*S*)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (608 mg, 1.34 mmol, 36% yield). MS (ESI) m/z 428.3 [M+H]⁺.

Benzyl ((S)-1-(((S)-1-cyano-2-(4-hydroxyphenyl)ethyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (41). General procedure **F** with benzyl ((*S*)-1-(((*S*)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (**40**) (508 mg, 1.19 mmol) to yield benzyl ((*S*)-1-(((*S*)-1-cyano-2-(4-hydroxyphenyl)ethyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (137 mg, 0.34 mmol, 28% yield). ¹H NMR (400 MHz, Acetone-*d*₆) δ : 8.32 (s, 1H), 8.04 (d, $J = 8.0$ Hz, 1H), 7.60 - 7.26 (m, 5H), 7.17 (d, $J = 8.5$ Hz, 2H), 6.79 (d, $J = 8.5$ Hz, 2H), 6.43 (m, 1H), 5.08 (q, $J = 12.5$ Hz, 2H), 4.99 (dt, $J = 7.5$, 5.5 Hz, 1H), 4.22 (dd, $J = 14.5$, 8.5 Hz, 1H), 3.05 (d, $J = 7.5$ Hz, 2H), 1.78 - 1.63 (m, 1H), 1.63 - 1.40 (m, 2H), 0.89 (dd, $J = 9.0$, 6.5 Hz, 6H). ¹³C NMR (100 MHz, Acetone-*d*₆) δ : 173.0, 157.6, 157.1, 138.1, 131.5, 129.2, 128.7, 126.9, 119.4, 116.2, 66.9, 54.3, 43.0, 41.7, 38.26, 25.3, 23.3, 21.8. MS (ESI) m/z 410.1 [M+H]⁺.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(2-(4-methoxyphenyl)acetamido)methyl)phosphonate 2,2,2-trifluoroacetate (42). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.33 (dt, $J = 16.0$, 8.0 Hz, 4H), 7.27-7.18 (m, 4H), 7.14 - 7.07 (m, 2H), 7.02 (dd, $J = 7.5$, 1.0 Hz, 2H), 6.87 - 6.79 (m, 2H), 4.77 (dd, $J = 18.5$, 6.5 Hz, 1H), 3.90 (t, $J = 14.5$ Hz, 2H), 3.74 (s, 3H), 3.54 (s, 2H), 3.19 - 3.02 (m, 2H), 2.39 (ddd, $J = 18.5$, 9.5, 5.5 Hz, 1H), 2.04 (d, $J = 13.0$ Hz, 2H), 1.49 (qd, $J = 13.0$, 4.0 Hz, 2H). MS (ESI) m/z 537.0 [M+H]⁺. Synthetic procedures in the supporting information.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(nicotinamido)methyl)phosphonate 2,2,2-trifluoroacetate (43). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 8.6-5.73 (m, 2H), 8.14 (d, $J = 8.0$ Hz, 1H), 7.56 (d, $J = 4.5$ Hz, 1H),

7.42 - 7.28 (m, 4H), 7.25 - 7.15 (m, 6H), 5.02 (dd, $J = 17.5, 8.0$ Hz, 1H), 3.96 (d, $J = 13.5$ Hz, 2H), 3.17 (ddd, $J = 21.5, 14.0, 2.5$ Hz, 2H), 2.54 (ddd, $J = 11.5, 8.0, 3.5$ Hz, 2H), 2.27 (d, $J = 13.0$ Hz, 1H), 2.09 (d, $J = 13.5$ Hz, 1H), 1.76 - 1.46 (m, 2H). MS (ESI) m/z 494.0 [M+H]⁺. Synthetic procedures in the supporting information.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(furan-2-carboxamido)methyl)phosphonate 2,2,2-trifluoroacetate (44). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.72 (dd, $J = 1.5, 0.5$ Hz, 1H), 7.45 - 7.27 (m, 4H), 7.26 - 7.07 (m, 7H), 6.62 (dd, $J = 3.5, 1.5$ Hz, 1H), 3.94 (d, $J = 13.5$ Hz, 2H), 3.15 (td, $J = 15.5, 2.5$ Hz, 2H), 2.52 (ddd, $J = 16.0, 9.5, 6.0$ Hz, 1H), 2.26 (d, $J = 13.0$ Hz, 1H), 2.08 (d, $J = 13.5$ Hz, 1H), 1.72 - 1.39 (m, 2H). MS (ESI) m/z 483.0 [M+H]⁺. Synthetic procedures in the supporting information.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(cinnamamido)methyl)phosphonate bis(2,2,2-trifluoroacetate) (45). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.62 (d, $J = 15.5$ Hz, 1H), 7.60 - 7.55 (m, 2H), 7.46 - 7.31 (m, 7H), 7.28 - 7.14 (m, 6H), 6.74 (d, $J = 15.5$ Hz, 1H), 4.96 (dd, $J = 18.5, 6.5$ Hz, 1H), 4.07 - 3.80 (m, 2H), 3.16 (td, $J = 15.5, 2.5$ Hz, 2H), 2.45 (ddd, $J = 18.5, 9.5, 5.5$ Hz, 1H), 2.12 (dd, $J = 9.5, 4.0$ Hz, 2H), 1.70 - 1.47 (m, 2H). MS (ESI) m/z 519.3 [M+H]⁺. Synthetic procedures in the supporting information.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(2-phenoxyethylsulfonamido)methyl)phosphonate 2,2,2-trifluoroacetate (46). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.35-7.30 (m, 4H), 7.25-7.17 (m, 6H), 7.13-7.07 (m, 2H), 6.97-6.91 (m, 1H), 6.90-6.84 (m, 2H), 4.42 (t, $J = 6.5$ Hz, 2H), 4.35 (dd, $J = 19.0, 5.4$ Hz, 1H), 3.98 (dd, $J = 10.5, 3.5$ Hz, 2H), 3.68 (t, $J = 6.5$ Hz, 2H), 3.14 (td, $J = 15.0, 2.5$ Hz, 2H), 2.52 - 2.34 (m, 1H), 2.21-1.98 (m, 2H), 1.87-1.56 (m, 2H). MS (ESI) m/z 573.2 [M+H]⁺. Synthetic procedures in the supporting information.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(3-(piperidin-4-yl)propanamido)methyl)phosphonate bis(2,2,2-trifluoroacetate) (47). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.44 - 7.33 (m, 4H), 7.29 - 7.19 (m, 4H), 7.15 - 7.09 (m, 2H), 4.82 (dd, $J = 18.0, 7.0$ Hz, 1H), 4.01 - 3.87 (m, 2H), 3.28 (d, $J = 2.5$ Hz, 2H), 3.14 (td, $J = 13.0, 2.0$ Hz, 2H), 2.84 - 2.70 (m, 2H), 2.48 - 2.29 (m, 3H), 2.16 - 2.02 (m, 2H), 1.94 - 1.82 (m, 2H), 1.68 - 1.45 (m, 5H), 1.41 - 1.25 (m, 2H). MS (ESI) m/z 528.3 [M+H]⁺. Synthetic procedures in the supporting information.

Diphenyl (*E*)-diphenyl ((3-(benzo[d][1,3]dioxol-5-yl)acrylamido)(1-carbamimidoylpiperidin-4-yl)methyl)phosphonate 2,2,2-trifluoroacetate (48). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.53 (d, $J = 15.5$ Hz, 1H), 7.36 (dd, $J = 17.0, 8.5$ Hz, 4H), 7.28 - 7.10 (m, 7H), 7.06 (dd, $J = 8.0, 1.5$ Hz, 1H), 6.86 (d, $J = 8.0$ Hz, 1H), 6.55 (d, $J = 15.5$ Hz, 1H), 6.01 (s, 2H), 5.04 - 4.90 (m, 1H), 3.94 (t, $J = 11.0$ Hz, 2H), 3.16 (dd, $J = 24.0, 13.0$ Hz, 2H), 2.56 - 2.32 (m, 1H), 2.20 - 1.99 (m, 2H), 1.70 - 1.48 (m, 2H). MS (ESI) m/z 563.2 [M+H]⁺. Synthetic procedures in the supporting information.

Diphenyl ((3-(benzo[d][1,3]dioxol-5-yl)propiolamido)(1-carbamimidoylpiperidin-4-yl)methyl)phosphonate (49). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.38 (td, *J* = 8.0, 3.0 Hz, 4H), 7.28 - 7.12 (m, 7H), 7.04 (d, *J* = 1.5 Hz, 1H), 6.89 (d, *J* = 8.0 Hz, 1H), 6.04 (s, 2H), 4.01 - 3.88 (m, 2H), 3.22 - 3.06 (m, 2H), 2.53 - 2.38 (m, 1H), 2.12 (t, *J* = 13.5 Hz, 2H), 1.69 - 1.48 (m, 2H). MS (ESI) *m/z* 561.2 [M+H]⁺. Synthetic procedures in the supporting information.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)((5-phenylpyrimidin-2-yl)amino)methyl)phosphonate 2,2,2-trifluoroacetate (50). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 8.79 - 8.59 (m, 2H), 8.07 (d, *J* = 10.0 Hz, 1H), 7.73 - 7.57 (m, 2H), 7.53 - 7.43 (m, 2H), 7.26 - 7.06 (m, 5H), 5.11 (ddd, *J* = 17.0, 10.5, 7.0 Hz, 1H), 7.36 (dt, *J* = 12.5, 4.0 Hz, 7H), 3.90 (t, *J* = 15.0 Hz, 2H), 3.19-2.95 (m, 2H), 2.49 - 2.42 (m, 1H), 2.00 (t, *J* = 10.5 Hz, 2H), 1.65-1.39 (m, 2H). MS (ESI) *m/z* 543.2 [M+H]⁺. Synthetic procedures in the supporting information.

(Z)-Diphenyl ((1-carbamimidoylpiperidin-4-yl)(3-phenylacrylamido)methyl)phosphonate 2,2,2-trifluoroacetate (51). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.52 (dd, *J* = 7.5, 1.5 Hz, 2H), 7.36 (t, *J* = 8.0 Hz, 4H), 7.27 - 7.18 (m, 5H), 7.17 - 7.11 (m, 4H), 6.87 (d, *J* = 12.5 Hz, 1H), 6.10 (dd, *J* = 12.5, 1.0 Hz, 1H), 4.92 - 4.88 (m, 1H), 3.91 (d, *J* = 14.0 Hz, 2H), 3.20 - 3.00 (m, 2H), 2.48 - 2.27 (m, 1H), 2.05 (dd, *J* = 25.5, 14.5 Hz, 2H), 1.65 - 1.36 (m, 2H). MS (ESI) *m/z* 519.3 [M+H]⁺. Synthetic procedures in the supporting information.

Methyl (1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethyl)carbamate (52). Procedure and characterization consistent with previously reported data.⁵⁰

Diphenyl (2-(4-guanidinophenyl)-1-((S)-2-((S)-3-hydroxy-2-(thiophene-2-carboxamido)propanamido)propanamido)ethyl)phosphonate (53). ¹H NMR (CDCl₃) δ: 7.8 - 7.1 (m, 17H), 5.1 (m, 1H), 4.2 - 4.3 (m, 2H), 3.9 (m, 2H), 3.4 (m, 2H), 1.3 (m, 3H). MS (ESI) *m/z* 679.3 [M+H]⁺, (100%). Procedure and characterization consistent with previously reported data.³⁵

Pent-4-yn-1-yl (1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethyl)carbamate (54). Procedure and characterization consistent with previously reported data.⁵¹

2-(2-Azidoethoxy)ethyl (1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethyl)carbamate (55). Procedure and characterization consistent with previously reported data.⁵¹

(Perfluorophenyl)methyl (1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethyl)carbamate 2,2,2-trifluoroacetate (56). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.5-7.0 (m, 14H), 5.1 (m, 2H), 4.60 (m, 1H), 4.57 (m, 1H), 3.4 (m, 1H), 3.08 (m, 1H). MS (ESI) *m/z* 635.1 [M+H]⁺. Synthetic procedures in the supporting information.

3,3,4,4,5,5,6,6,7,8,8,8-dodecafluoro-7-(trifluoromethyl)octyl 1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethylcarbamate 2,2,2-trifluoroacetate (57). ¹H NMR (400 MHz, CDCl₃) δ: 10.02 (s, 1H), 7.41-7.12 (m, 14H), 5.43 (m, 1H), 4.69 (m, 1H), 4.26 (m, 2H), 3.38 (m, 1H), 3.08 (m, 1H), 2.35 (m, 2H). MS (ESI) *m/z* 851.1 [M+H]⁺. Synthetic procedures in the supporting information.

Benzyl ((4-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (58). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 8.82 (d, *J* = 10.0 Hz, 1H), 7.48 (dd, *J* = 16.0, 9.0 Hz, 2H), 7.42 – 7.26 (m, 8H), 7.18 (dd, *J* = 15.5, 8.0 Hz, 2H), 7.09 – 7.02 (m, 2H), 6.96 (t, *J* = 8.0 Hz, 3H), 5.50 (dd, *J* = 22.0, 10.0 Hz, 1H), 5.09 (dd, *J* = 33.5, 12.5 Hz, 2H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ: 156.0, 155.9, 150.1, 149.8, 136.6, 131.0, 129.83, 129.78, 129.6, 129.5, 128.3, 127.9, 125.3, 125.2, 120.4, 120.35, 120.29, 120.2, 118.3, 66.1, 52.4 (d, C₁-P, *J*_{CP} = 159.0 Hz). HRMS: Calc: 489.16 Found: 489.1588 [M+H]⁺. Procedure and characterization consistent with previously reported data.⁵²

2-Phenoxyethyl ((4-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (59). Procedure and characterization consistent with previously reported data.⁵²

4,4,4-Trifluorobutyl ((4-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (60). Procedure and characterization consistent with previously reported data.⁵²

(Perfluorophenyl)methyl ((4-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (61). Procedure and characterization consistent with previously reported data.⁵²

Diphenyl ((4-guanidinophenyl)((4-(trifluoromethyl)phenyl)sulfonamido)methyl)phosphonate (62). Procedure and characterization consistent with previously reported data.⁵²

Diphenyl ((4-guanidinophenyl)(phenylsulfonamido)methyl)phosphonate (63). Procedure and characterization consistent with previously reported data.⁵²

(Perfluorophenyl)methyl ((diphenoxyphosphoryl)(4-guanidinophenyl)methyl)carbamate (64). Procedure and characterization consistent with previously reported data.⁵²

Methyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (65). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.70 (m, 1H), 7.55 (m, 2H), 7.41 -7.07 (m, 10H), 5.6 (d, 1H, *J* = 20.0 Hz), 3.75 (s, 3H). MS (ESI) *m/z* 531.1 [M+Na]⁺. Synthetic procedures in the supporting information.

Benzo[d][1,3]dioxol-5-ylmethyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (66). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.70-7.50 (m, 4H), 7.40-6.80 (m, 13H), 5.80 (m, 2H), 5.60 (d, *J* = 28.0 Hz, 1H), 4.90 (s, 2H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ: 156.6, 156.1, 150.0,

149.7, 147.3, 146.8, 136.2, 131.5, 131.1, 130.0, 129.9, 125.5, 125.4, 122.1, 121.7, 121.1, 120.4, 120.4, 120.4, 120.0, 119.2 – 113.6 (m, CF₃), 108.5, 108.1, 101.4, 100.8, 66.1, 52.4 (d, C₁-P, *J*_{CP} = 158.0 Hz). MS (ESI) *m/z* 629.2 [M+H]⁺, (96%). HRMS: Calc: 629.13 Found: 629.1301 [M+H]⁺. Synthetic procedures in the supporting information.

2-Aminoethyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate 2,2,2-trifluoroacetate (67). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.75-7.5 (m, 4H), 7.40-6.93 (m, 10H), 5.72 (m, 2H), 4.43 (m, 2H), 3.44 (m, 2H). MS (ESI) *m/z* 538.2 [M-H]⁻. Synthetic procedures in the supporting information.

Benzyl 2-(4-(3,3-dimethylureido)phenyl)-1-(diphenoxyphosphoryl)ethylcarbamate (68). ¹H NMR (400 MHz, CDCl₃) δ: 7.05 – 7.38 (m, 19H), 6.39 (s, 1H), 5.34 (d, *J* = 10.5 Hz, 1H), 4.95 – 5.10 (m, 2H), 4.69 – 4.84 (m, 1H), 3.36 (ddd, *J* = 4.5, 10.0, 14.5 Hz, 1H), 3.02 (s, 6H), 1.28 (s, 1H). MS (ESI) *m/z* 574.7 [M+H]⁺. Synthetic procedures in the supporting information.

Benzyl ((4-(2-aminoethoxy)phenyl)(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (69). ¹H NMR (400 MHz, CDCl₃) δ: 7.60-6.80 (m, 19H), 5.58 (d, *J* = 22.5 Hz, 1H), 5.15 (m, 2H), 4.25 (t, *J* = 5.0 Hz, 2H), 3.37 (t, *J* = 5.0 Hz, 2H). MS (ESI) *m/z* 533.1 [M+H]⁺. Synthetic procedures in the supporting information.

Benzyl (1-(diphenoxyphosphoryl)-3-(4-nitrophenyl)propyl)carbamate (70). General procedure C with 3-(4-nitrophenyl)propanal (420 mg, 2.34 mmol) to yield benzyl (1-(diphenoxyphosphoryl)-3-(4-nitrophenyl)propyl)carbamate (700 mg, 1.28 mmol, 55% yield). ¹H NMR (400 MHz, Acetone-*d*₆) δ: 8.24 - 8.05 (m, 2H), 7.57 - 7.44 (m, 2H), 7.44 - 7.27 (m, 9H), 7.26 - 7.05 (m, 6H), 5.22 - 5.08 (m, 2H), 4.58 - 4.25 (m, 1H), 3.06 (ddd, *J* = 14.0, 9.0, 5.0 Hz, 1H), 2.92 (ddd, *J* = 24.0, 15.0, 11.0 Hz, 1H), 2.46 - 2.17 (m, 2H). ¹³C NMR (100 MHz, Acetone-*d*₆) δ: 157.9, 152.3, 152.0, 150.8, 148.2, 138.7, 131.4, 131.3, 130.0, 129.9, 129.6, 129.3, 126.8, 126.7, 125.1, 122.3, 122.0, 68.1, 49.7 (d, C₁-P, *J*_{CP} = 159.0 Hz), 33.2, 32.4. MS (ESI) *m/z* 547.1 [M+H]⁺, (100%). HRMS: Calc: 547.16 Found: 547.1646 [M+H]⁺.

Methyl ((4-carbamimidoylphenyl)(diphenoxyphosphoryl)methyl)carbamate (71). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 9.45 (s, 2H), 9.22 (s, 1H), 8.80 (d, *J* = 20.0 Hz, 1H), 7.90 – 7.85 (m, 4H), 7.40 – 7.36 (m, 4H), 7.24 – 7.20 (m, 2H), 7.10 – 7.00 (m, 4H) 5.78 – 5.72 (m, 1H), 3.61 (s, 3H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ: 165.4, 156.7, 149.9, 149.6, 140.5, 130.1, 128.9, 127.9, 125.4, 120.4, 120.38, 120.32, 120.27, 120.23, 64.9, 52.3 (d, C₁-P, *J*_{CP} = 157.5 Hz). MS (ESI) *m/z* 440.2 [M+H]⁺, (100%). HRMS: Calc: 440.14 Found: 440.1369 [M+H]⁺. Procedure and characterization consistent with previously reported data.⁵⁰

Methyl ((diphenoxyphosphoryl)(5-nitronaphthalen-1-yl)methyl)carbamate (72). ^1H NMR (400 MHz, CDCl_3) δ : 8.62-8.51 (m, 1H), 8.46-8.31 (m, 1H), 8.24-8.17 (m, 1H), 8.03-7.91 (m, 1H), 7.74-7.58 (m, 2H), 7.35-7.01 (m, 10H), 6.5-6.35 (m, 1H), 6.11-6.01 (m, 1H). MS (ESI) m/z 493.1 $[\text{M}+\text{H}]^+$. Synthetic procedures in the supporting information.

Diphenyl ((1-carbamimidoylazetid-3-yl)(pyrimidin-2-ylamino)methyl)phosphonate 2,2,2-trifluoroacetate (73). ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ : 8.47-8.24 (m, 2H), 8.08 (d, $J = 9.5$ Hz, 1H), 7.37 (td, $J = 8.0, 3.0$ Hz, 4H), 7.29 (s, 3H), 7.21 (td, $J = 7.5, 3.5$ Hz, 2H), 7.11 (d, $J = 8.0$ Hz, 3H), 6.73 (t, $J = 5.0$ Hz, 1H), 5.37 (dt, $J = 15.5, 9.5$ Hz, 1H), 4.13 – 4.26 (m, 2H), 4.07 (ddd, $J = 9.5, 6.0, 3.5$ Hz, 2H), 3.62-3.44 (m, 1H). MS (ESI) m/z 439.2 $[\text{M}+\text{H}]^+$. Synthetic procedures in the supporting information.

Diphenyl ((3-(*N*-hydroxycarbamimidoyl)phenyl)(pyrimidin-2-ylamino)methyl)phosphonate (74). ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ : 9.67 (s, 1H), 8.47 (dd, $J = 10.5, 2.0$ Hz, 1H), 8.37 (d, $J = 4.5$ Hz, 2H), 8.05 (q, $J = 2.0$ Hz, 1H), 7.72 – 7.81 (m, 1H), 7.66 (dq, $J = 8.0, 1.5$ Hz, 1H), 7.41 (t, $J = 8.0$ Hz, 1H), 7.29 – 7.36 (m, 4H), 7.14 – 7.21 (m, 2H), 7.04 (dq, $J = 7.8, 1.2$ Hz, 2H), 6.98 (dq, $J = 8.0, 1.0$ Hz, 2H), 6.71 (t, $J = 5.0$ Hz, 1H), 6.26 (dd, $J = 22.5, 10.5$ Hz, 1H), 5.82 (s, 2H). No ionization found. Synthetic procedures in the supporting information.

2-(2-(Prop-2-yn-1-yloxy)ethoxy)ethyl ((diphenoxyphosphoryl)(4-(piperazin-1-yl)phenyl)methyl)carbamate 2,2,2-trifluoroacetate (75). ^1H NMR (400 MHz, CDCl_3) δ : 7.42 (dd, $J = 8.5, 1.5$ Hz, 2H), 7.36 – 7.28 (m, 2H), 7.26 – 7.16 (m, 3H), 7.16 – 7.05 (m, 3H), 6.96 – 6.78 (m, 4H), 6.25 (d, $J = 9.5$ Hz, 1H), 5.53 (dd, $J = 22.5, 9.5$ Hz, 1H), 4.22 (d, $J = 2.5$ Hz, 2H), 4.17 (d, $J = 2.5$ Hz, 2H), 3.76 - 3.73 (m, 2H), 3.71 – 3.60 (m, 4H), 3.40 – 3.25 (m, 4H), 3.03 (s, 2H), 2.96 – 2.92 (m, 2H), 2.42 (s, 1H). MS (ESI) m/z 594.8 $[\text{M}+\text{H}]^+$. Synthetic procedures in the supporting information.

Benzyl (1-(diphenoxyphosphoryl)-4-guanidinobutyl)carbamate (76). Procedure and characterization consistent with previously reported data.⁵³

Benzyl (benzo[d][1,3]dioxol-5-yl(diphenoxyphosphoryl)methyl)carbamate (77). Procedure and characterization consistent with previously reported data.⁴³

Benzyl ((4-(dimethylamino)phenyl)(diphenoxyphosphoryl)methyl)carbamate (78). ^1H NMR (400 MHz, CDCl_3) δ : 7.37 – 7.06 (m, 15H), 6.88 (d, $J = 9.0$ Hz, 2H), 6.69 (d, $J = 8.5$ Hz, 2H), 5.72 (d, $J = 8.0$ Hz, 1H), 5.47 (m, 1H), 5.09 (m, 2H), 2.94 (s, 6H, 2 CH_3). MS (ESI) m/z 517.3 $[\text{M}+\text{H}]^+$, (95%). HRMS: Calc: 517.19 Found: 517.1894 $[\text{M}+\text{H}]^+$. Procedure and characterization consistent with previously reported data.⁴³

Benzyl ((diphenoxyphosphoryl)(pyridin-3-yl)methyl)carbamate (79). Procedure and characterization consistent with previously reported data.⁵⁴

Benzyl ((diphenoxyphosphoryl)(3-nitrophenyl)methyl)carbamate (80). Procedure and characterization consistent with previously reported data.⁵⁵

Benzyl (1-(diphenoxyphosphoryl)ethyl)carbamate (81). Procedure and characterization consistent with previously reported data.⁴⁶

Benzyl (1-(diphenoxyphosphoryl)-3-methylbutyl)carbamate (82). Procedure and characterization consistent with previously reported data.⁴⁶

Benzyl ((diphenoxyphosphoryl)(4-guanidinophenyl)methyl)carbamate (83). Procedure and characterization consistent with previously reported data.⁵²

Benzyl (2-(benzyloxy)-1-(diphenoxyphosphoryl)ethyl)carbamate (84). Procedure and characterization consistent with previously reported data.⁵⁶

Benzyl ((3-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (85). ¹H NMR (400 MHz, CDCl₃) δ: 7.29 - 6.75 (m, 19H), 6.54 (s (b), 1H), 5.59 (dd, *J* = 22.5, 10.0 Hz, 1H), 5.05 (dd, *J* = 53.0, 12.2 Hz, 2H). ¹³C NMR (CDCl₃) δ: 155.6, 150.1, 146.5, 136.0, 135.2, 129.8, 129.7, 128.6, 128.3, 128.2, 125.4, 120.5, 118.6, 115.7, 115.0, 67.5, 52.9 (d, C₁-P, *J*_{CP} = 159.5 Hz). MS (ESI) *m/z* 489.2 [M+H]⁺, (95%). HRMS: Calc: 489.16 Found: 489.1582 [M+H]⁺. Procedure and characterization consistent with previously reported data.⁵⁵

Benzyl ((diphenoxyphosphoryl)(6-methoxypyridin-2-yl)methyl)carbamate (86). General procedure C with 6-methoxypicolinaldehyde (200 mg, 1.46 mmol) to give benzyl (diphenoxyphosphoryl)(6-methoxypyridin-2-yl)methylcarbamate (298 mg, 0.59 mmol, 41% yield). ¹H NMR (400 MHz, CDCl₃) δ: 7.57 (dd, *J* = 25.0, 18.0 Hz, 1H), 7.51 - 7.21 (m, 8H), 7.19 - 6.92 (m, 6H), 6.75 (br s, 1H), 6.36 (m, 1H), 5.72 (d, *J* = 13.0 Hz, 2H), 5.19 (dd, *J* = 37.0, 12.0 Hz, 2H), 3.88 (s, 3H). ¹³C NMR (100 MHz, CDCl₃) δ: 163.6, 155.7, 150.5, 150.4, 150.3, 150.2, 149.6, 139.4, 136.1, 129.7, 129.6, 128.7, 128.6, 128.3, 125.3, 125.2, 120.5, 120.3, 116.4, 111.0, 67.5, 56.3 (d, *J* = 156.5 Hz), 53.6. MS (ESI) *m/z* 505.6 [M+H]⁺, (100%). HRMS: Calc: 505.15 Found: 505.1506 [M+H]⁺.

Benzyl ((2-chloro-5-nitrophenyl)(diphenoxyphosphoryl)methyl)carbamate (87). Procedure and characterization consistent with previously reported data.⁵⁷

Benzyl ((4-carbamimidoylphenyl)(diphenoxyphosphoryl)methyl)carbamate (88). Procedure and characterization consistent with previously reported data.⁵⁰

Benzyl ((5-chloro-1H-indol-3-yl)(diphenoxyphosphoryl)methyl)carbamate (89). General procedure C with *tert*-butyl 5-chloro-3-formyl-1H-indole-1-carboxylate (500 mg, 1.78 mmol) to yield benzyl ((5-chloro-

1H-indol-3-yl) (diphenoxyphosphoryl)methyl carbamate (830 mg, 1.52 mmol, 85% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 11.43 (s, 1H), 8.70 (d, *J* = 9.5 Hz, 1H), 7.77 (s, 1H), 7.68 (t, *J* = 2.5 Hz, 1H), 7.46 – 7.25 (m, 10H), 7.24 – 7.05 (m, 6H), 6.97 (d, *J* = 8.0 Hz, 2H), 5.82 (dd, *J* = 21.0, 10.0 Hz, 1H), 5.22 – 4.97 (m, 2H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ: 156.1, 150.3, 150.0, 136.8, 134.4, 129.8, 128.4, 127.9, 127.4, 127.0, 127.0, 125.2, 123.9, 121.57, 120.4, 118.4, 113.3, 107.5, 66.1, 45.3 (d, C₁-P, *J*_{CP} = 165.5 Hz). MS (ESI) *m/z* 547.1 [M+H]⁺, (95%).

Benzyl ((6-carbamimidoylnaphthalen-2-yl)(diphenoxyphosphoryl)methyl)carbamate (90). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 8.80 – 8.60 (m, 1H), 8.46 (s, 1H), 8.22 (s, 1H), 8.15 – 8.08 (m, 2H), 8.00 – 7.65 (m, 4H), 7.64 – 7.58 (m, 1H), 7.45 – 7.12 (m, 10H), 7.08 – 6.95 (m, 5H), 6.00 – 5.78 (m, 1H), 5.25 – 5.11 (m, 2H). MS (ESI) *m/z* 566.2 [M+H]⁺, (100%). Procedure and characterization consistent with previously reported data.⁵⁸

Benzyl ((diphenoxyphosphoryl)(4-nitrophenyl)methyl)carbamate (91). Procedure and characterization consistent with previously reported data.⁵⁵

Benzyl ((3-carbamimidoylphenyl)(diphenoxyphosphoryl)methyl)carbamate (92). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 9.63 (s, 2H), 9.34 (s, 2H), 8.96 (d, *J* = 10.0 Hz, 1H), 8.04 (d, *J* = 7.5 Hz, 2H), 7.80 (d, *J* = 7.5 Hz, 1H), 7.67 (t, *J* = 8.0 Hz, 1H), 7.43 – 7.28 (m, 9H), 7.20 (t, *J* = 7.0 Hz, 2H), 7.03 (dd, *J* = 18.0, 8.0 Hz, 4H), 5.72 (dd, *J* = 22.5, 10.0 Hz, 1H), 5.11 (dd, *J* = 39.0, 12.5 Hz, 2H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ: 165.6, 158.4, 150.0, 149.6, 136.5, 135.4, 133.2, 129.9, 129.9, 129.5, 128.6, 128.4, 128.1, 128.0, 125.5, 125.4, 120.3, 120.3, 120.2, 120.1, 66.4, 50.7 (d, C₁-P, *J*_{CP} = 159.5 Hz). MS (ESI) *m/z* 516.4 [M+H]⁺, (100%). HRMS: Calc: 516.17 Found: 516.1703 [M+H]⁺. Procedure and characterization consistent with previously reported data.⁴⁷

Benzyl (1-(diphenoxyphosphoryl)-3-phenylpropyl)carbamate (93). ¹H NMR (400 MHz, CDCl₃) δ: 7.26 (s, 20H), 5.26 – 5.07 (m, 3H), 4.59 – 4.46 (m, 1H), 2.92 – 2.81 (m, 1H), 2.80 – 2.67 (m, 1H), 2.44 – 2.29 (m, 1H), 2.15 – 1.98 (m, 1H). ¹³C NMR (100 MHz, CDCl₃) δ: 156.0, 150.3, 150.1, 140.6, 136.1, 130.0, 129.9, 128.7, 128.7, 128.6, 128.5, 128.3, 126.4, 125.5, 120.7, 120.5, 67.6, 48.3 (d, C₁-P, *J*_{CP} = 158.0 Hz), 32.1, 32.0. MS (ESI) *m/z* 502.3 [M+H]⁺. Synthetic procedures in the supporting information.

Benzyl ((diphenoxyphosphoryl)(naphthalen-1-yl)methyl)carbamate (94). ¹H NMR (400 MHz, CDCl₃) δ: 8.24 (d, *J* = 8.5 Hz, 1H), 7.81 (dd, *J* = 14.0, 8.0 Hz, 3H), 7.55 – 6.89 (m, 14H), 6.60 (d, *J* = 8.0 Hz, 2H), 6.49 – 6.38 (m, 1H), 6.04 – 5.88 (m, 1H), 5.04 (m, 2H). ¹³C NMR (100 MHz, CDCl₃) δ: 155.7, 150.1, 150.0, 136.0, 134.0, 131.4, 130.8, 129.9, 129.6, 129.0, 128.7, 127.2, 126.6, 126.2, 125.3, 123.3, 120.7, 120.2, 67.7, 48.4 (d, C₁-P, *J*_{CP} = 161.0 Hz). MS (ESI) *m/z* 524.2 [M+H]⁺. Synthetic procedures in the supporting information.

Benzyl ((diphenoxyphosphoryl)(naphthalen-2-yl)methyl)carbamate (95). ¹H NMR (400 MHz, CDCl₃) δ: 7.89 – 7.67 (m, 4H), 7.56 – 6.93 (m, 16H), 6.80 (m, 2H), 5.96 (m, 1H), 5.68 (m, 1H), 5.02 (m, 2H). ¹³C NMR

(100 MHz, CDCl₃) δ : 156.0, 150.2, 136.0, 133.3, 131.7, 129.9, 129.8, 128.9, 128.7, 128.5, 128.3, 127.8, 126.7, 125.6, 125.5, 120.6, 120.5, 67.8, 53.1 (d, C₁-P, J_{CP} = 157.0 Hz). MS (ESI) m/z 524.2 [M+H]⁺. Synthetic procedures in the supporting information.

Benzyl ((3-cyanophenyl)(diphenoxyphosphoryl)methyl)carbamate (96). General procedure C with 3-cyanobenzaldehyde (447 mg, 3.31 mmol) to yield benzyl ((3-cyanophenyl)(diphenoxyphosphoryl)methyl)carbamate (998 mg, 2.00 mmol, 61% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ : 7.81-6.89 (m, 19H), 6.15-6.04 (m, 1H), 5.14 (dd, J = 23.0, 10.5 Hz, 1H). MS (ESI) m/z 499.5 [M+H]⁺. HRMS: Calc: 499.14 Found: 499.1431 [M+H]⁺.

Benzyl ((4-(dimethylamino)naphthalen-1-yl)(diphenoxyphosphoryl)methyl)carbamate (97). General procedure C with 4-dimethylamino-1-naphthaldehyde (200 mg, 1.004 mmol) to yield benzyl ((4-(dimethylamino)naphthalen-1-yl)(diphenoxyphosphoryl)methyl)carbamate (21 mg, 0.04 mmol, 4% yield). ¹H NMR (400 MHz, CDCl₃) δ : 8.18 (t, J = 8.0 Hz, 1H), 7.69 (dd, J = 8.0, 2.5 Hz, 1H), 7.51 - 7.38 (m, 1H), 7.28 - 6.85 (m, 7H), 6.56 (d, J = 8.1 Hz, 1H), 6.34 (dd, J = 22.5, 9.9 Hz, 1H), 5.98 - 5.93 (m, 1H), 5.00 (dt, J = 28.0, 12.0 Hz, 1H), 2.78 (s, 3H), 1.18 - 1.13 (m, 1H). ¹³C NMR (100 MHz, CDCl₃) δ : 155.7, 152.0, 150.6, 150.2, 136.1, 132.6, 129.8, 129.1, 129.1, 128.6, 128.3, 126.9, 125.4, 125.1, 124.8, 123.6, 120.7, 120.2, 113.5, 67.0, 48.2 (d, C₁-P, J_{CP} = 162.0 Hz), 44.9. MS (ESI) m/z 567.2 [M+H]⁺.

Benzyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (98). General procedure C with 2,2,2-trifluoro-*N*-(4-formylphenyl)acetamide (70 mg, 0.32 mmol) to yield benzyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (17 mg, 0.03 mmol, 9% yield). ¹H NMR (400 MHz, CDCl₃) δ : 8.00 (s, 1H), 7.55, 7.50 - 7.00 (m, 14 H), 5.75 (m, 1H), 5.50 (m, 1H), 5.10 (m, 2H). MS (ESI) m/z 585.2 [M+H]⁺.

***Tert*-butyl 4-(4-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl) methyl)phenyl)piperazine-1-carboxylate (99).** General procedure C with *tert*-butyl 4-(4-formylphenyl)piperazine-1-carboxylate (200 mg, 0.69 mmol) to yield 4-(4-((benzyloxycarbonylamino)(diphenoxyphosphoryl) methyl)phenyl)piperazine-1-carboxylate (80 mg, 0.12 mmol, 17% yield). ¹H NMR (400 MHz, CDCl₃) δ : 6.85-7.40 (m, 19H), 5.95 (s, 1H), 5.49 (m, 1H), 5.09 (m, 2H), 3.57 (t, J = 5.0 Hz, 4H), 3.12 (t, J = 5.0 Hz, 4 H), 1.48 (s, 9H). MS (ESI) m/z 680.2 [M+Na]⁺.

***Tert*-butyl 4-(4-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl)methyl)-2-cyanophenyl)piperazine-1-carboxylate (100).** General procedure C with *tert*-butyl 4-(2-cyano-4-formylphenyl)piperazine-1-carboxylate (500 mg, 1.59 mmol), to yield *tert*-butyl 4-(4-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl)methyl)-2-cyanophenyl)piperazine-1-carboxylate (457 mg, 0.67 mmol, 42% yield). MS (ESI) m/z 705.3 [M+Na]⁺.

***Tert*-butyl 4-((benzyloxycarbonylamino)(diphenoxyphosphoryl)methyl)piperidine-1-carboxylate (101).**

General procedure C with *tert*-butyl 4-formylpiperidine-1-carboxylate (120 mg, 0.56 mmol) to yield *tert*-butyl 4-((benzyloxycarbonylamino)(diphenoxyphosphoryl)methyl)piperidine-1-carboxylate (100 mg, 0.17 mmol, 30% yield). MS (ESI) m/z 603.1 [M+Na]⁺.

***Tert*-butyl 4-(2-(benzyloxycarbonylamino)-2-(diphenoxyphosphoryl)ethyl)piperidine-1-carboxylate (102).**

General procedure C with *tert*-butyl 4-(2-oxoethyl)piperidine-1-carboxylate (600 mg, 2.64 mmol) to give *tert*-butyl 4-(2-(benzyloxycarbonylamino)-2-(diphenoxyphosphoryl)ethyl)piperidine-1-carboxylate (250 mg, 0.42 mmol, 16% yield). MS (ESI) m/z 595.9 [M+H]⁺.

Benzyl ((4-(((*tert*-butoxycarbonyl)amino)methyl)cyclohexyl)(diphenoxyphosphoryl)methyl)carbamate (103).

General procedure C with *tert*-butyl ((1*r*,4*r*)-4-formylcyclohexyl)methylcarbamate (400 mg, 1.66 mmol) to yield benzyl ((4-(((*tert*-butoxycarbonyl)amino)methyl)cyclohexyl)(diphenoxyphosphoryl)methyl)carbamate (252 mg, 0.42 mmol, 25% yield). MS (ESI) m/z 609.6 [M+H]⁺.

***Tert*-butyl 3-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl)methylazetidine-1-carboxylate (104).**

General procedure C with *tert*-butyl 3-formylazetidine-1-carboxylate (613 mg, 3.31 mmol) to yield *tert*-butyl 3-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl) methylazetidine-1-carboxylate as a white solid (1.084 g, 1.96 mmol, 59% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ : 6.38 (dd, $J = 23.5, 10.5$ Hz, 1H), 6.74 (t, $J = 5.0$ Hz, 1H), 7.02 (ddq, $J = 16.0, 8.0, 1.0$ Hz, 4H), 7.15 - 7.24 (m, 2H), 7.28 - 7.40 (m, 4H), 7.63 (t, $J = 8.0$ Hz, 1H), 7.83 (dq, $J = 8.0, 1.5$ Hz, 1H), 8.04 - 8.15 (m, 1H), 8.27 (q, $J = 2.0$ Hz, 1H), 8.38 (d, $J = 5.0$ Hz, 2H), 8.65 (dd, $J = 10.5, 2.5$ Hz, 1H).

Benzyl (2-(4-(((*tert*-butoxycarbonyl)amino)phenyl)-1-(diphenoxyphosphoryl)ethyl)carbamate (105).

Procedure and characterization consistent with previously reported data.⁵⁰

Benzyl ((6-(((*tert*-butoxycarbonyl)amino)naphthalen-2-yl)(diphenoxyphosphoryl)methyl)carbamate (106).

General procedure C with *tert*-butyl 6-formylnaphthalen-2-ylcarbamate (300 mg, 1,106 mmol) to yield benzyl ((6-(((*tert*-butoxycarbonyl)amino)naphthalen-2-yl)(diphenoxyphosphoryl)methyl)carbamate (350 mg, 0.55 mmol, 50% yield). ¹H NMR (400 MHz, CDCl₃) δ : 9.57 (s, 1H), 8.94 (d, $J = 10.0$ Hz, 1H), 8.07 (br s, 1H), 7.98 (br s, 1H), 7.65 (d, $J = 8.5$ Hz, 1H), 7.28 (m, 9H), 7.47 (dd, $J = 2.0, 8.5$ Hz, 1H), 7.14 (t, $J = 6.0$ Hz, 2H), 7.01 (d, $J = 8.5$ Hz, 2H), 6.92 (d, $J = 8.5$ Hz, 2H), 5.65 (m, 1H), 5.09 (d, $J = 12.5$ Hz, 1H), 5.01 (d, $J = 12.5$ Hz, 1H), 1.47 (s, 9H). MS (ESI) m/z 639.1 [M+H]⁺

Benzyl ((diphenoxyphosphoryl)(4-(piperazin-1-yl)phenyl)methyl)carbamate 2,2,2-trifluoroacetate (107).

General procedure H with *tert*-butyl 4-(4-

((benzyloxycarbonylamino)(diphenoxyphosphoryl)methyl)phenyl)piperazine-1-carboxylate (**98**) (800 mg, 1.22 mmol) to afford benzyl ((diphenoxyphosphoryl)(4-(piperazin-1-yl)phenyl)methyl)carbamate 2,2,2-trifluoroacetate (75 mg, 0.11 mmol, 9% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 6.90-7.50 (m, 19H), 5.53 (d, *J* = 20.0 Hz, 1H), 5.12 (m, 2H), 3.48 (m, 4H), 3.37 (m, 4H). ¹³C NMR (100 MHz, Methanol-*d*₄) δ: 158.3, 158.2, 151.7, 151.4, 138.0, 130.8, 130.7, 129.5, 129.1, 129.0, 126.6, 121.8, 121.7, 121.6, 121.5, 117.9, 117.8, 68.2, 53.7 (d, C₁-P, *J*_{CP} = 159.5 Hz), 47.5, 44.7. MS (ESI) *m/z* 558.2 [M+H]⁺, (95%). HRMS: Calc: 558.22 Found: 558.2163 [M+H]⁺.

Benzyl ((3-cyano-4-(piperazin-1-yl)phenyl)(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (108). General procedure **H** with *tert*-butyl 4-(4-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl)methyl)-2-cyanophenyl)piperazine-1-carboxylate (**100**) (120 mg, 0.18 mmol) to yield benzyl ((3-cyano-4-(piperazin-1-yl)phenyl)(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (9 mg, 0.15 mmol, 9%) as a colourless oil. ¹H NMR (400 MHz, CDCl₃) δ: 7.59 (dd, *J* = 28.5, 19.5 Hz, 2H), 7.44 - 7.27 (m, 6H), 7.24 - 6.74 (m, 11H), 5.99 (s, 1H), 5.48 (dd, *J* = 22.5, 9.5 Hz, 1H), 5.24 - 4.95 (m, 2H), 3.31 - 3.11 (m, 4H), 3.06 (s, 4H). ¹³C NMR (100 MHz, CDCl₃) δ: 156.0, 155.1, 150.0, 135.6, 133.9, 133.9, 130.0, 129.7, 128.8, 128.6, 128.4, 127.7, 125.7, 120.5, 120.4, 120.4, 120.3, 119.2, 117.9, 115.5, 105.7, 67.9, 52.6, 51.0 (d, C₁-P, *J*_{CP} = 151.5 Hz), 45.61. MS (ESI) *m/z* 583.2 [M+H]⁺.

Benzyl ((diphenoxyphosphoryl)(piperidin-4-yl)methyl)carbamate 2,2,2-trifluoroacetate (109). General procedure **H** with *tert*-butyl 4-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl)methyl)piperidine-1-carboxylate (**101**) (900 mg, 1.55 mmol) to give benzyl ((diphenoxyphosphoryl)(piperidin-4-yl)methyl)carbamate 2,2,2-trifluoroacetate (520 mg, 0.87 mmol, 56% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 8.05 (d, *J* = 10.5 Hz, 1H), 7.34 (m, 9H), 7.23 (m, 2H), 7.14 (m, 4H), 5.14 (m, 2H), 4.46 (m, 1H), 3.45 (m, 2H), 2.40 (m, 1H), 2.31 (m, 2H), 2.21 (m, 2H), 1.71 (m, 2H). MS (ESI) *m/z* 481.7 [M+H]⁺.

Benzyl 1-(diphenoxyphosphoryl)-2-(piperidin-4-yl)ethylcarbamate 2,2,2-trifluoroacetate (110). General Procedure **H** with *tert*-butyl 4-(2-(benzyloxycarbonylamino)-2-(diphenoxyphosphoryl)ethyl)piperidine-1-carboxylate (**102**) (250 mg, 0.42 mmol) to give benzyl 1-(diphenoxyphosphoryl)-2-(piperidin-4-yl)ethylcarbamate 2,2,2-trifluoroacetate (200 mg, 0.33 mmol, 78% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.36 (m, 9H), 7.22 (m, 2H), 7.14 (m, 4H), 5.14 (q, *J* = 12.5 Hz, 2H), 4.53 (m, 1H), 3.60 (d, *J* = 12.5 Hz, 2H), 2.95 (dt, *J* = 12.5, 2.5 Hz, 1H), 2.85 (dt, *J* = 12.5, 2.5 Hz, 1H), 2.05 (m, 1H), 1.90 (m, 3H), 1.81 (m, 1H), 1.52 (m, 1H), 1.35 (m, 1H). MS (ESI) *m/z* 495.2 [M+H]⁺

Benzyl ((4-(aminomethyl)cyclohexyl)(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (111). General procedure **H** with benzyl ((4-(((*tert*-butoxycarbonyl)amino)methyl)cyclohexyl)(diphenoxyphosphoryl)methyl)carbamate (**103**) (252 mg,

0.42 mmol) to yield benzyl ((4-(aminomethyl)cyclohexyl)(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (180 mg, 0.35 mmol, 83% yield) as a colourless oil. ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.32 (m, 9H), 7.21 (m, 2H), 7.11 (m, 3H), 5.18 (d, *J* = 12.5 Hz, 1H), 5.10 (d, *J* = 12.5 Hz, 1H), 4.37 (m, 1H), 2.79 (d, *J* = 7.0 Hz, 2H), 2.10 (m, 3H), 1.90 (m, 2H), 1.60 (m, 2H), 1.30 (m, 2H), 1.11 (m, 2H). MS (ESI) *m/z* 509.1 [M+H]⁺.

Benzyl (azetidin-3-yl(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (112). General procedure **H** with *tert*-butyl 3-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl) methyl)azetidine-1-carboxylate (**104**) (1.084 g, 1.96 mmol) to yield benzyl (azetidin-3-yl(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (884 mg, 1.61 mmol, 82% yield). MS (ESI) *m/z* 453.4 [M+H]⁺.

Benzyl (2-(4-aminophenyl)-1-(diphenoxyphosphoryl)ethyl)carbamate 2,2,2-trifluoroacetate (113). Procedure and characterization consistent with previously reported data.⁵⁰

Benzyl ((6-aminonaphthalen-2-yl)(diphenoxyphosphoryl)methyl)carbamate (114)

General procedure **H** with benzyl ((6-((*tert*-butoxycarbonyl)amino)naphthalen-2-yl)(diphenoxyphosphoryl)methyl)carbamate (**106**) (350 mg, 0.55 mmol) to yield benzyl (6-aminonaphthalen-2-yl)(diphenoxyphosphoryl)methylcarbamate 2,2,2-trifluoroacetate (210 mg, 0.32 mmol, 59% yield) as a colourless oil. ¹H NMR (400 MHz, Methanol-*d*₄) δ: 9.76 (d, *J* = 10.0 Hz, 1H), 8.77 (s, 1H), 8.50 (m, 3H), 8.15 (m, 9H), 8.01 (m, 2H), 7.89 (m, 3H), 7.80 (d, *J* = 8.0 Hz, 2H), 6.49 (m, 1H), 5.96 (d, *J* = 12.5 Hz, 1H), 5.87 (d, *J* = 12.5 Hz, 1H). MS (ESI) *m/z* 539.9 [M+H]⁺.

Benzyl ((1-carbamimidoylazetidin-3-yl)(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (115). General procedure **I** followed by general procedure **H** with benzyl (azetidin-3-yl(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (**112**) (884 mg, 1.61 mmol) to yield benzyl ((1-carbamimidoylazetidin-3-yl)(diphenoxyphosphoryl)methyl) carbamate 2,2,2-trifluoroacetate (357 mg, 0.72 mmol, 45% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 8.35 (dd, *J* = 15.0, 10.0 Hz, 1H), 7.47 – 7.27 (m, 9H), 7.23 (dt, *J* = 11.5, 6.0 Hz, 2H), 7.15 (dd, *J* = 8.5, 7.0 Hz, 4H), 5.21 – 4.97 (m, 2H), 4.86 – 4.68 (m, 1H), 4.28 – 4.09 (m, 2H), 4.03 (dt, *J* = 17.5, 7.0 Hz, 2H), 3.42 – 3.27 (m, 1H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ: 156.3, 151.7, 149.6, 136.6, 130.0, 128.5, 128.0, 127.8, 125.5, 125.4, 120.6, 120.4, 66.2, 52.8, 49.7 (d, C₁-P, *J*_{CP} = 157.5 Hz), 28.0. MS (ESI) *m/z* 495.4 [M+H]⁺, (100%). HRMS: Calc: 495.18 Found: 495.1785 [M+H]⁺.

Benzyl 2-(4-acetamidophenyl)-1-(diphenoxyphosphoryl)ethylcarbamate (116). General procedure **J** with acetyl chloride (27 L, 0.38 mmol, 1.2 eq) and benzyl 2-(4-aminophenyl)-1-(diphenoxyphosphoryl)ethylcarbamate (**113**) (160 mg, 0.32 mmol) to yield benzyl 2-(4-acetamidophenyl)-1-(diphenoxyphosphoryl)ethylcarbamate (104 mg, 0.19 mmol, 60% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ:

7.55 – 7.45 (m, 2H), 7.40 – 7.30 (m, 4H), 7.30 – 7.17 (m, 9H), 7.15 (dd, $J = 5.5, 4.5$ Hz, 4H), 5.06 – 4.93 (m, 2H), 4.63 (qd, $J = 11.5, 5.5$ Hz, 1H), 3.39 – 3.32 (m, 1H), 3.01 (ddd, $J = 14.0, 12.0, 9.0$ Hz, 1H), 2.12 (s, 3H). ^{13}C NMR (100 MHz, Methanol- d_4) δ : 171.6, 159.0, 151.8, 151.4, 138.9, 138.1, 133.6, 131.00, 130.9, 130.7, 129.4, 128.9, 128.6, 126.8, 126.7, 121.8, 121.8, 121.7, 121.6, 121.1, 67.7, 51.4 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 158.5$ Hz), 35.6, 23.8. MS (ESI) m/z 545.7 $[\text{M}+\text{H}]^+$

Benzyl ((diphenoxyphosphoryl)(5-nitronaphthalen-1-yl)methyl)carbamate (117). General procedure **C** with 5-nitro-1-naphthaldehyde (200 mg, 0.99 mmol), to give benzyl ((diphenoxyphosphoryl)(5-nitronaphthalen-1-yl)methyl)carbamate (568 mg, 0.35 mmol, 35% yield). ^1H NMR (400 MHz, CDCl_3) δ : 8.62 (d, $J = 9.0$ Hz, 1H), 8.48 (d, $J = 9.0$ Hz, 1H), 8.20 (d, $J = 8.0$ Hz, 1H), 7.99 (s, 1H), 7.76 – 7.62 (m, 2H), 7.42 – 7.28 (m, 8H), 7.13 (m, 7H), 6.73 (d, $J = 8.5$ Hz, 2H), 6.44 (d, $J = 23.5$ Hz, 1H), 6.03 (s, 1H), 5.12 (dd, $J = 43.5, 12.0$ Hz, 2H). MS (ESI) m/z 569.3 $[\text{M}+\text{H}]^+$.

Benzyl ((4-(dimethylamino)-3-nitrophenyl)(diphenoxyphosphoryl)methyl)carbamate (118). General procedure **C** with 5-nitro-1-naphthaldehyde with 4-(dimethylamino)-3-nitrobenzaldehyde (500 mg, 2.57 mmol), to yield benzyl ((4-(dimethylamino)-3-nitrophenyl)(diphenoxyphosphoryl)methyl)carbamate (793 mg, 1.41 mmol, 55% yield). ^1H NMR (400 MHz, CDCl_3) δ : 7.91 (s, 1H), 7.54 (d, $J = 8.5$ Hz, 1H), 7.35 (s, 5H), 7.30 - 7.20 (m, 5H), 7.20 - 7.05 (m, 4H), 7.00 (t, $J = 9.0$ Hz, 3H), 6.00 (dd, $J = 9.0, 4.5$ Hz, 1H), 5.51 (dd, $J = 22.5, 9.5$ Hz, 1H), 5.21 - 4.99 (m, 2H), 2.91 (s, 6H). ^{13}C NMR (100 MHz, CDCl_3) δ : 155.7, 150.0, 145.9, 138.4, 135.8, 133.4, 129.8, 128.6, 128.4, 128.3, 126.5, 125.6, 123.8, 120.4, 120.3, 118.7, 67.7, 51.5 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 159.0$ Hz), 42.5. MS (ESI) m/z 562.2 $[\text{M}+\text{H}]^+$.

Benzyl ((5-aminonaphthalen-1-yl)(diphenoxyphosphoryl)methyl)carbamate (119). General procedure **K** with benzyl ((diphenoxyphosphoryl)(5-nitronaphthalen-1-yl)methyl)carbamate (**117**) (500 mg, 0.88 mmol) to give benzyl ((5-aminonaphthalen-1-yl)(diphenoxyphosphoryl)methyl)carbamate (440 mg, 0.82 mmol, 93% yield). MS (ESI) m/z 539.3 $[\text{M} + \text{Na}]^+$

Benzyl (3-(4-aminophenyl)-1-(diphenoxyphosphoryl)propyl)carbamate (120). General procedure **K** with benzyl (1-(diphenoxyphosphoryl)-3-(4-nitrophenyl)propyl)carbamate (**70**) (8.03 g, 14.69 mmol) to yield benzyl (3-(4-aminophenyl)-1-(diphenoxyphosphoryl)propyl)carbamate (4.24 g, 8.21 mmol, 56% yield). ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ : 8.12 (d, $J = 9.5$ Hz, 1H), 7.40 - 7.27 (m, 9H), 7.20 (q, $J = 7.0$ Hz, 2H), 7.08 (dd, $J = 12.0, 8.5$ Hz, 4H), 6.84 (d, $J = 8.5$ Hz, 2H), 6.51 (d, $J = 8.5$ Hz, 2H), 5.15 - 5.07 (m, 2H), 5.07 - 4.95 (m, 2H), 4.33 - 4.13 (m, 1H), 2.72 - 2.58 (m, 1H), 2.49 - 2.37 (m, 1H), 2.13 - 1.96 (m, 2H). ^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ : 156.1, 150.0, 149.8, 146.5, 137.0, 130.3, 129.8, 129.0, 128.4, 127.9, 127.8, 127.5, 125.2, 120.6, 120.4, 114.2, 65.8, 47.5 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 158.0$ Hz), 30.9, 30.3. MS (ESI) m/z 517.2 $[\text{M}+\text{H}]^+$. HRMS: Calc: 517.19 Found: 517.1871 $[\text{M}+\text{H}]^+$.

Benzyl ((3-amino-4-(dimethylamino)phenyl)(diphenoxyphosphoryl)methyl)carbamate (121). General procedure **K** with benzyl ((4-(dimethylamino)-3-nitrophenyl)(diphenoxyphosphoryl)methyl)carbamate (**118**) (793 mg, 1.41 mmol) to give benzyl ((3-amino-4-(dimethylamino)phenyl)(diphenoxyphosphoryl)methyl)carbamate (310 mg, 0.58 mmol, 41% yield). ¹H NMR (400 MHz, CDCl₃) δ: 7.40 - 7.28 (m, 4H), 7.25 - 7.04 (m, 9H), 6.95 (d, *J* = 8.0 Hz, 1H), 6.84 (d, *J* = 8.5 Hz, 4H), 5.88 (dd, *J* = 10.0, 3.0 Hz, 1H), 5.45 (dd, *J* = 22.0, 10.0 Hz, 1H), 5.09 (dd, *J* = 40.0, 12.0 Hz, 2H), 4.42 - 3.52 (m, 2H), 2.67 (s, 6H). ¹³C NMR (100 MHz, CDCl₃) δ: 155.6, 150.2, 141.6, 136.0, 129.9, 129.7, 129.6, 128.6, 128.3, 125.4, 125.3, 120.6, 120.5, 120.1, 119.6, 118.4, 115.0, 67.5, 52.6 (d, C₁-P, *J*_{CP} = 159.0 Hz), 43.6. MS (ESI) *m/z* 532.2 [M+H]⁺.

Benzyl ((4-acetamidophenyl)(diphenoxyphosphoryl)methyl)carbamate (122). General procedure **J** with aL, 0.39 mmol, \square cetyl chloride (28 1.2 eq) and benzyl (4-aminophenyl)(diphenoxyphosphoryl)methylcarbamate (**58**) (160 mg, 0.33 mmol) to yield benzyl ((4-acetamidophenyl)(diphenoxyphosphoryl)methyl)carbamate (110 mg, 0.20 mmol, 63% yield) as a white solid. ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.52 - 7.43 (m, 2H), 7.38 - 7.31 (m, 4H), 7.30 - 7.15 (m, 9H), 7.14 - 7.10 (m, 4H), 5.65 - 5.51 (m, 1H), 5.02 - 4.97 (m, 2H), 2.15 (s, 3H). MS (ESI) *m/z* 531.7 [M+H]⁺.

Benzyl ((4-(3,3-dimethylureido)phenyl)(diphenoxyphosphoryl)methyl)carbamate (123). General procedure **J** with dimethylcarbonylchloride (87 mg, 0.81 mmol, 2.2 eq) and benzyl (4-aminophenyl)(diphenoxyphosphoryl)methylcarbamate (**58**) (180 mg, 0.37 mmol) to yield benzyl ((4-(3,3-dimethylureido)phenyl)(diphenoxyphosphoryl)methyl)carbamate (30 mg, 0.05 mmol, 15% yield) as a white solid. ¹H NMR (400 MHz, CDCl₃) δ: 7.42 (s, 3H), 7.06 - 7.41 (m, 15H), 6.87 - 6.96 (m, 2H), 6.43 (s, 1H), 5.87 (d, *J* = 9.5 Hz, 1H), 5.55 (dd, *J* = 9.5, 22.0 Hz, 1H), 5.17 (d, *J* = 12.0 Hz, 1H), 5.08 (d, *J* = 12.0 Hz, 1H), 3.04 (s, 6H). MS (ESI) *m/z* 560.7 [M+H]⁺.

Benzyl ((diphenoxyphosphoryl)(3-guanidinophenyl)methyl)carbamate 2,2,2-trifluoroacetate (124). Procedure and characterization consistent with previously reported data.⁵⁹

Benzyl ((diphenoxyphosphoryl)(5-guanidinonaphthalen-1-yl)methyl)carbamate 2,2,2-trifluoroacetate (125). General procedure **I** followed by general procedure **H** with benzyl ((5-aminonaphthalen-1-yl)(diphenoxyphosphoryl)methyl)carbamate (**119**) (440 mg, 0.82 mmol) to yield benzyl ((diphenoxyphosphoryl)(5-guanidinonaphthalen-1-yl)methyl)carbamate 2,2,2-trifluoroacetate (69 mg, 0.10 mmol, 12% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 8.39 (d, *J* = 8.5 Hz, 1H), 8.13 - 8.01 (m, 2H), 7.72 - 7.65 (m, 2H), 7.58 (d, *J* = 7.0 Hz, 1H), 7.41 - 7.12 (m, 11H), 7.08 - 7.02 (m, 2H), 6.93 - 6.87 (m, 2H), 6.58 (d, *J* = 23.0 Hz, 1H), 5.14 (dd, *J* = 48.0, 12.5 Hz, 2H). MS (ESI) *m/z* 581.2 [M+H]⁺. HRMS: Calc: 581.20 Found: 581.1940 [M+H]⁺.

Benzyl (1-(diphenoxyphosphoryl)-3-(4-(methylsulfonamido)phenyl)propyl)carbamate (126). General procedure **J** with methanesulfonylchloride (0.42 mL, 5.43 mmol, 1.2 eq) and benzyl (3-(4-aminophenyl)-1-(diphenoxyphosphoryl)propyl)carbamate (**120**) (2.55 g, 4.94 mmol) to yield benzyl (1-(diphenoxyphosphoryl)-3-(4-(methylsulfonamido)phenyl)propyl)carbamate (1.62 g, 2.72 mmol, 55% yield) as an colourless foam. ^1H NMR (400 MHz, Acetone- d_6) δ : 8.49 (s, 1H), 7.58 - 7.13 (m, 19H), 7.08 (d, $J = 10.0$ Hz, 1H), 5.23 - 4.91 (m, 2H), 4.44 (dt, $J = 28.0, 25.0, 12.5$ Hz, 1H), 3.01 - 2.80 (m, 4H), 2.79 - 2.65 (m, 1H), 2.42 - 2.09 (m, 2H). ^{13}C NMR (100 MHz, Acetone- d_6) δ : 157.1, 151.6, 151.3, 138.1, 138.0, 137.5, 130.5, 130.4, 129.2, 128.8, 126.0, 125.9, 121.6, 121.3, 67.2, 49.0 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 159.0$ Hz), 39.2, 32.2, 31.9. MS (ESI) m/z 595.1 $[\text{M}+\text{H}]^+$. HRMS: Calc: 595.17 Found: 595.1647 $[\text{M}+\text{H}]^+$.

Benzyl ((4-(dimethylamino)-3-(methylsulfonamido)phenyl)(diphenoxyphosphoryl)methyl)carbamate (127). General procedure **J** with methanesulfonyl chloride (0.05 mL, 0.64 mmol, 1.2 eq) and benzyl ((3-amino-4-(dimethylamino)phenyl)(diphenoxyphosphoryl)methyl)carbamate (**121**) (310 mg, 0.58 mmol) to yield benzyl ((4-(dimethylamino)-3-(methylsulfonamido)phenyl)(diphenoxyphosphoryl)methyl)carbamate as a colourless foam. ^1H NMR (400 MHz, CDCl_3) δ : 7.75 (s, 1H), 7.59 (s, 1H), 7.22 - 6.95 (m, 15H), 6.91 - 6.77 (m, 2H), 6.06 (dd, $J = 10.0, 4.5$ Hz, 1H), 5.47 (dt, $J = 15.0, 7.5$ Hz, 1H), 5.13 - 4.91 (m, 2H), 2.79 (s, 3H), 2.52 (s, 6H). ^{13}C NMR (100 MHz, CDCl_3) δ : 155.6, 150.2, 142.85, 136.0, 133.6, 131.6, 129.9, 129.7, 128.6, 128.4, 128.3, 125.6, 125.4, 123.6, 121.7, 120.5, 120.4, 115.8, 67.6, 52.7 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 158.0$ Hz), 44.9, 39.3. MS (ESI) m/z 610.3 $[\text{M}+\text{H}]^+$.

Methyl ((diphenoxyphosphoryl)(3-(trifluoromethyl)phenyl)methyl)carbamate (128). General procedure **C** with 2-(4-(trifluoromethyl)phenyl)acetaldehyde (**7**) (290 mg, 1.54 mmol) and methyl carbamate (116 mg, 1.54 mmol, 1 eq) to give methyl (1-(diphenoxyphosphoryl)-2-(4-(trifluoromethyl)phenyl)ethyl)carbamate (150 mg, 0.31 mmol, 20% yield). ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ : 8.09 (d, $J = 9.5$ Hz, 1H), 7.68 (d, $J = 8.0$ Hz, 2H), 7.57 (t, $J = 9.5$ Hz, 2H), 7.44 - 7.36 (m, 5H), 7.22 (ddd, $J = 17.0, 8.0, 2.0$ Hz, 7H), 4.63 - 4.47 (m, 1H), 3.43 (s, 3H), 3.41 - 3.37 (m, 1H), 3.31 (s, 1H), 3.16 - 3.00 (m, 1H). ^{13}C NMR (100 MHz, $\text{DMSO-}d_6$) δ : 157.4, 150.9, 143.0, 131.0, 130.9, 128.3, 126.4, 126.0, 124.0, 121.4, 52.8, 50.6 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 159.0$ Hz), 35.0. MS (ESI) m/z 480.1 $[\text{M}+\text{H}]^+$. HRMS: Calc: 480.12 Found: 480.1179 $[\text{M}+\text{H}]^+$.

Benzyl (1-(bis(4-acetamidophenoxy)phosphoryl)-2-(4-(trifluoromethyl)phenyl)ethyl) carbamate (129). General procedure **C** with 2-(4-(trifluoromethyl)phenyl)acetaldehyde (**7**) (642 mg, 3.41 mmol), and tris(4-acetamidophenyl) phosphite (1.81 g, 3.75 mmol, 1.1 eq) to give benzyl (1-(bis(4-acetamidophenoxy)phosphoryl)-2-(4-(trifluoromethyl)phenyl)ethyl) carbamate (51 mg, 0.08 mmol, 2% yield). ^1H NMR (400 MHz, Acetone- d_6) δ : 9.22 (s, 2H), 7.63 (dt, $J = 12.5, 5.0$ Hz, 8H), 7.34 - 7.26 (m, 3H), 7.21 (dt, $J = 5.0, 4.0$ Hz, 2H), 7.18 - 7.10 (m, 4H), 5.07 - 4.88 (m, 2H), 4.82 - 4.64 (m, 1H), 3.58 - 3.44 (m, 1H), 3.21 (ddd, $J = 14.0, 12.0, 8.5$ Hz, 1H), 2.05 (d, $J = 2.0$ Hz, 6H). ^{13}C NMR (100 MHz, Acetone- d_6) δ : 168.7, 156.8,

146.7, 146.5, 142.9, 137.9, 137.84, 137.7, 130.9, 129.1, 128.6, 128.4, 126.1, 126.0, 124.2, 121.7, 121.4, 120.9, 66.9, 50.6 (d, C₁-P, $J_{CP} = 159.0$ Hz), 35.7, 24.6. MS (ESI) m/z 670.2 [M+H]⁺. HRMS: Calc: 670.19 Found: 670.1912 [M+H]⁺.

Methyl (1-(bis(4-acetamidophenoxy)phosphoryl)-2-(4-(trifluoromethyl)phenyl)ethyl)carbamate (130). General procedure C with 2-(4-(trifluoromethyl)phenyl)acetaldehyde (**7**) (376 mg, 2.00 mmol), methyl carbamate (150 mg, 2.00 mmol, 1 eq) and tris(4-acetamidophenyl) phosphite (1.05 g, 2.20 mmol, 1.1 eq) to give methyl (1-(bis(4-acetamidophenoxy)phosphoryl)-2-(4-(trifluoromethyl)phenyl)ethyl)carbamate (8 mg, 0.01 mmol, 1% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ : 9.04 (s, 2H), 7.62-7.49 (m, 6H), 7.42-7.12 (m, 4H), 7.11 – 6.94 (m, 2H), 6.73 (m, 1H), 4.73 - 4.59 (m, 1H), 3.43 (s, 3H), 3.52 - 3.40 (m, 1H), 3.28 (ddd, $J = 13.5, 12.5, 8.0$ Hz, 1H), 2.03 (s, 6H). MS (ESI) m/z 594.1 [M+H]⁺.

Methyl ((diphenoxyphosphoryl)(3-nitrophenyl)methyl)carbamate (131). General procedure C with 3-nitrobenzaldehyde (500 mg, 3.31 mmol) and methyl carbamate (248 mg, 3.31 mmol) to give methyl ((diphenoxyphosphoryl)(3-nitrophenyl)methyl)carbamate (1.04 g, 2.13 mmol, 64% yield). ¹H NMR (400 MHz, Acetone-*d*₆) δ : 8.61 (dd, $J = 4.0, 2.2$ Hz, 1H), 8.24 (dt, $J = 8.0, 2.5$ Hz, 1H), 8.13 (d, $J = 7.5$ Hz, 1H), 7.95 (d, $J = 8.5$ Hz, 1H), 7.73 (t, $J = 8.0$ Hz, 1H), 7.34 (ddd, $J = 8.0, 4.0, 1.5$ Hz, 4H), 7.23 - 7.14 (m, 4H), 7.14 - 7.09 (m, 2H), 5.87 (dd, $J = 23.5, 10.0$ Hz, 1H), 3.66 (s, 3H). MS (ESI) m/z 443.2 [M+H]⁺.

Benzo[d][1,3]dioxol-5-ylmethyl ((diphenoxyphosphoryl)(3-nitrophenyl)methyl)carbamate (132). General procedure C with 3-nitrobenzaldehyde (200 mg, 1.32 mmol) and benzo[d][1,3]dioxol-5-ylmethyl carbamate (258 mg, 1.32 mmol) to yield benzo[d][1,3]dioxol-5-ylmethyl ((diphenoxyphosphoryl)(3-nitrophenyl)methyl)carbamate (250 mg, 0.39 mmol, 30% yield). MS (ESI) m/z 585.2 [M+Na]⁺.

Benzyl ((bis(4-acetamidophenoxy)phosphoryl)(3-nitrophenyl)methyl)carbamate (133). General procedure C with 3-nitrobenzaldehyde (1.00 g, 6.62 mmol) and tris(4-acetamidophenyl) phosphite (7.46 g, 7.28 mmol) to yield benzyl ((bis(4-acetamidophenoxy)phosphoryl)(3-nitrophenyl)methyl)carbamate (134 mg, 0.21 mmol, 3% yield). MS (ESI) m/z 633.2 = [M+H]⁺.

Methyl ((bis(4-acetamidophenoxy)phosphoryl)(3-nitrophenyl)methyl)carbamate (134). General procedure C with 3-nitrobenzaldehyde (1.00 g, 6.62 mmol), methyl carbamate (497 mg, 6.62 mmol) and tris(4-acetamidophenyl) phosphite (3.50 g, 7.28 mmol) to give methyl ((bis(4-acetamidophenoxy)phosphoryl)(3-nitrophenyl)methyl)carbamate (480 mg, 0.73 mmol, 11% yield). MS (ESI) m/z 443.2 [M+H]⁺.

Methyl ((3-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (135). General procedure K with pent-4-yn-1-yl ((diphenoxyphosphoryl)(4-nitrophenyl)methyl)carbamate (**131**) (2.24 g, 4.54 mmol) to yield pent-4-yn-1-yl ((4-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (2.04 g, 4.40 mmol, 97% yield). ¹H

NMR (400 MHz, DMSO-*d*₆) δ : 8.09 (d, *J* = 9.5 Hz, 1H), 7.68 (d, *J* = 8.0 Hz, 2H), 7.57 (t, *J* = 9.5 Hz, 2H), 7.44 - 7.36 (m, 5H), 7.22 (ddd, *J* = 17.0, 8.0, 2.0 Hz, 7H), 4.63 - 4.47 (m, 1H), 3.43 (s, 3H), 3.41 - 3.37 (m, 1H), 3.31 (s, 1H), 3.16 - 3.00 (m, 1H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ : 157.4, 150.9, 143.0, 131.0, 130.9, 128.3, 126.4, 126.0, 124.0, 121.4, 52.8, 50.6 (d, C₁-P, *J*_{CP} = 159.0 Hz), 35.0. MS (ESI) *m/z* 480.1 [M+H]⁺. HRMS: Calc: 413.13 Found: 413.1256 [M+H]⁺.

Benzo[d][1,3]dioxol-5-ylmethyl ((3-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (136).

General procedure **K** with benzo[d][1,3]dioxol-5-ylmethyl ((diphenoxyphosphoryl)(3-nitrophenyl)methyl)carbamate (**132**) (250 mg, 0.44 mmol) to yield benzo[d][1,3]dioxol-5-ylmethyl ((3-aminophenyl)(diphenoxyphosphoryl)methyl)carbamate (40 mg, 0.08 mmol, 17% yield). ¹H NMR (400 MHz, Acetone-*d*₆) δ : 7.59 (dd, *J* = 58.0, 10.0 Hz, 1H), 7.37 - 7.23 (m, 4H), 7.22 - 6.97 (m, 7H), 6.96 - 6.84 (m, 4H), 6.84 - 6.76 (m, 1H), 6.67 - 6.60 (m, 1H), 5.99 (d, *J* = 1.0 Hz, 2H), 5.56 (ddd, *J* = 58.0, 22.0, 10.0 Hz, 1H), 5.11 - 4.90 (m, 2H), 4.71 (s, 2H). ¹³C NMR (100 MHz, Acetone-*d*₆) δ : 156.8, 153.0, 151.4, 149.5, 148.6, 148.4, 136.2, 131.6, 130.4, 130.0, 125.9, 122.8, 121.3, 120.2, 117.4, 115.1, 109.6, 108.7, 102.0, 67.3, 54.2 (d, C₁-P, *J*_{CP} = 158.0 Hz). MS (ESI) *m/z* 533.1 [M+H]⁺.

Benzyl ((3-aminophenyl)(bis(4-acetamidophenoxy)phosphoryl)methyl)carbamate (137).

General procedure **K** with benzyl ((bis(4-acetamidophenoxy) phosphoryl)(3-nitrophenyl)methyl)carbamate (**133**) (130 mg, 0.21 mmol) to yield benzyl ((3-aminophenyl)(bis(4-acetamidophenoxy)phosphoryl)methyl)carbamate (40 mg, 0.07 mmol, 32% yield). ¹H NMR (400 MHz, Acetone-*d*₆) δ : 9.22 (d, *J* = 4.0 Hz, 2H), 7.78-7.66 (m, 1H), 7.54 (dd, *J* = 12.5, 6.0 Hz, 4H), 7.39 - 7.28 (m, 6H), 7.04 (dt, *J* = 17.0, 8.0 Hz, 3H), 6.94 - 6.87 (m, 3H), 6.67 (d, *J* = 2.0 Hz, 1H), 6.65 (d, *J* = 7.5 Hz, 1H), 5.63-5.49 (m, 1H), 5.11 (ddd, *J* = 34.5, 12.4, 3.0 Hz, 2H), 4.72 (s, 1H), 2.02 (d, *J* = 2.9 Hz, 6H). ¹³C NMR (100 MHz, Acetone-*d*₆) δ : 168.9, 156.8, 153.0, 149.5, 146.7, 137.7, 136.3, 130.0, 129.9, 129.2, 128.8, 123.7, 121.5, 120.9, 120.4, 120.3, 117.5, 115.3, 115.1, 67.4, 54.8 (d, C₁-P, *J*_{CP} = 156.5 Hz), 53.4, 24.2. MS (ESI) *m/z* 603.2 [M+H]⁺.

Methyl ((3-aminophenyl)(bis(4-acetamidophenoxy)phosphoryl)methyl)carbamate (138).

General procedure **K** with methyl ((bis(4-acetamidophenoxy)phosphoryl)(3-nitrophenyl)methyl)carbamate (**134**) (480 mg, 0.86 mmol) to yield methyl ((3-aminophenyl)(bis(4-acetamidophenoxy)phosphoryl)methyl)carbamate (135 mg, 0.26 mmol, 30% yield). ¹H NMR (400 MHz, CD₃CN) δ : 8.38 (d, *J* = 6.5 Hz, 1H), 7.52 - 7.42 (m, 4H), 7.10 (td, *J* = 8.0, 1.0 Hz, 1H), 7.02 - 6.97 (m, 2H), 6.92 - 6.86 (m, 2H), 6.79 (s, 1H), 6.78 - 6.76 (m, 2H), 6.73 (m, 1H), 6.63 - 6.59 (m, 1H), 5.38 (dd, *J* = 22.5, 10.0 Hz, 1H), 3.63 (s, 3H), 2.01 (t, *J* = 3.5 Hz, 6H). ¹³C NMR (100 MHz, CD₃CN) δ : 169.5, 157.2, 149.3, 146.6, 137.5, 136.3, 130.40, 121.7, 121.35, 117.75, 115.4, 115.0, 53.9 (d, C₁-P, *J*_{CP} = 156.5 Hz), 53.2, 24.2. MS (ESI) *m/z* 527.2 [M+H]⁺.

Methyl ((3-cyanophenyl)(diphenoxyphosphoryl)methyl)carbamate (139). General procedure **C** with 3-cyanobenzaldehyde (800 mg, 6.10 mmol) and methyl carbamate (458 mg, 6.10 mmol), to yield methyl ((3-

cyanophenyl)(diphenoxyphosphoryl)methyl)carbamate (2.58 g, 6.11 mmol, 99% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ: 8.84 (d, *J* = 10.0 Hz, 1H), 8.14 (d, *J* = 1.5 Hz, 1H), 8.00 (d, *J* = 8.0 Hz, 1H), 7.87 - 7.82 (m, 1H), 7.63 (t, *J* = 8.0 Hz, 1H), 7.37 (dd, *J* = 16.0, 7.5 Hz, 4H), 7.21 (td, *J* = 7.5, 3.5 Hz, 2H), 7.08 (d, *J* = 8.5 Hz, 2H), 7.03 - 6.98 (m, 2H), 5.77 (dd, *J* = 23.0, 10.5 Hz, 1H), 3.61 (s, 3H). MS (ESI) *m/z* 423.2 [M+H]⁺.

Benzyl ((bis(4-acetamidophenoxy)phosphoryl)(3-cyanophenyl)methyl)carbamate (140). General procedure **C** with 3-cyanobenzaldehyde (3.00 g, 22.9 mmol) and tris(4-acetamidophenyl) phosphite (28.7 g, 25.2 mmol) to give benzyl ((bis(4-acetamidophenoxy)phosphoryl)(3-cyanophenyl)methyl)carbamate (4.13 g, 5.12 mmol, 22% yield). MS (ESI) *m/z* 613.3 = [M+H]⁺.

Methyl ((bis(4-acetamidophenoxy)phosphoryl)(3-cyanophenyl)methyl)carbamate (141). General procedure **C** with 3-cyanobenzaldehyde (3.00 g, 22.8 mmol), methyl carbamate (1.72 g, 22.8 mmol) and tris(4-acetamidophenyl) phosphite (30.1 g, 25.2 mmol) to yield methyl ((bis(4-acetamidophenoxy)phosphoryl)(3-cyanophenyl)methyl) carbamate (2.13 g, 3.96 mmol, 17% yield). MS (ESI) *m/z* 537.2 = [M+H]⁺.

Methyl ((3-(*N*-acetoxycarbamimidoyl)phenyl)(diphenoxyphosphoryl)methyl)carbamate (142). General procedure **C** with methyl ((3-cyanophenyl)(diphenoxyphosphoryl)methyl)carbamate (**139**) (1.00 g, 2.37 mmol) to yield methyl ((3-(*N*-acetoxycarbamimidoyl)phenyl)(diphenoxyphosphoryl)methyl)carbamate (2.34 g, 2.82 mmol). MS (ESI) *m/z* 498.2 [M+H]⁺

Benzyl ((3-(*N*-acetoxycarbamimidoyl)phenyl)(bis(4-acetamidophenoxy)phosphoryl) methyl)carbamate (143). General procedure **L** with benzyl ((bis(4-acetamidophenoxy)phosphoryl)(3-cyanophenyl)methyl) carbamate (**140**) (4.13 g, 5.12 mmol) to yield benzyl ((3-(*N*-acetoxycarbamimidoyl)phenyl)(bis(4-acetamidophenoxy)phosphoryl)methyl)carbamate (950 mg, 1.38 mmol, 27% yield). MS (ESI) *m/z* 688.4 [M+H]⁺.

Methyl ((3-(*N*-acetoxycarbamimidoyl)phenyl)(bis(4-acetamidophenoxy)phosphoryl) methyl)carbamate (144). General procedure **L** with methyl ((bis(4-acetamidophenoxy)phosphoryl)(3-cyanophenyl)methyl) carbamate (**141**) (2.13 g, 3.96 mmol) to yield methyl ((3-(*N*-acetoxycarbamimidoyl)phenyl)(bis(4-acetamidophenoxy)phosphoryl)methyl)carbamate (3.70 g, 3.75 mmol, 95% yield). MS (ESI) *m/z* 612.3 = [M+H]⁺.

Methyl ((3-carbamimidoylphenyl)(diphenoxyphosphoryl)methyl)carbamate (145). General procedure **M** with methyl ((3-(*N*-acetoxycarbamimidoyl)phenyl)(diphenoxyphosphoryl)methyl)carbamate (**142**) (2.24 g, 2.79 mmol) to yield methyl ((3-carbamimidoylphenyl)(diphenoxyphosphoryl)methyl)carbamate (88 mg, 0.20 mmol, 7% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.99 (d, *J* = 7.5 Hz, 1H), 7.95 (s, 1H), 7.80 (d, *J* = 24.5, 12.5 Hz, 1H), 7.69 (t, *J* = 8.0 Hz, 1H), 7.34 (td, *J* = 8.0, 3.5 Hz, 4H), 7.22 (t, *J* = 7.5 Hz, 2H), 7.05 (dd, *J* =

22.5, 8.5 Hz, 4H), 5.79 (d, $J = 23.5$ Hz, 1H), 3.71 (s, 3H). ^{13}C NMR (100 MHz, Methanol- d_4) δ : 168.2, 158.8, 151.4, 137.3, 134.7, 131.0, 130.4, 129.2, 129.0, 127.0, 121.5, 54.6 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 158.0$ Hz), 53.4. MS (ESI) m/z 440.4 $[\text{M}+\text{H}]^+$.

Benzyl ((bis(4-acetamidophenoxy)phosphoryl)(3-carbamimidoylphenyl)methyl) carbamate (146).

General procedure **M** with benzyl ((3-(*N*-acetoxycarbamimidoyl)phenyl)(bis(4-acetamidophenoxy)phosphoryl) methyl)carbamate (**143**) (950 mg, 1.38 mmol) to yield benzyl ((bis(4-acetamidophenoxy)phosphoryl)(3-carbamimidoylphenyl)methyl)carbamate (172 mg, 0.273 mmol, 20% yield). ^1H NMR (400 MHz, Methanol- d_4) δ : 7.96 (d, $J = 7.5$ Hz, 1H), 7.93 (s, 1H), 7.79 (d, $J = 7.5$ Hz, 1H), 7.66 (t, $J = 8.0$ Hz, 1H), 7.51 - 7.46 (m, 4H), 7.39 - 7.30 (m, 5H), 6.95 (dd, $J = 22.0, 8.5$ Hz, 4H), 5.78 (d, $J = 23.0$ Hz, 1H), 5.14 (dd, $J = 49.5, 12.5$ Hz, 2H), 2.10 (s, 6H). ^{13}C NMR (100 MHz, Methanol- d_4) δ : 171.6, 168.4, 158.2, 147.0), 137.8, 137.2, 134.6, 131.0, 130.6, 129.5, 129.3, 129.1, 129.0, 122.3, 121.7, 68.4, 53.7 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 158.5$ Hz), 23.7. MS (ESI) m/z 630.3 $[\text{M}+\text{H}]^+$.

Methyl ((bis(4-acetamidophenoxy)phosphoryl)(3-carbamimidoylphenyl)methyl) carbamate (147).

General procedure **M** with methyl ((3-(*N*-acetylcarbamimidoyl)phenyl)(bis(4-acetamidophenoxy)phosphoryl)methyl) carbamate (**144**) (3.70 g, 3.85 mmol) to yield methyl ((bis(4-acetamidophenoxy)phosphoryl)(3-carbamimidoylphenyl)methyl)carbamate (190 mg, 0.34 mmol, 9% yield). ^1H NMR (400 MHz, DMSO- d_6) δ : 10.17 (s, 3H), 9.44 (s, 1H), 9.29 (s, 1H), 8.80 (d, $J = 10.0$ Hz, 1H), 8.06 (s, 1H), 8.02 (d, $J = 7.5$ Hz, 1H), 7.83 (d, $J = 7.0$ Hz, 1H), 7.68 (t, $J = 8.0$ Hz, 1H), 7.60 - 7.46 (m, 4H), 6.98 (dd, $J = 26.5, 8.5$ Hz, 4H), 5.66 (dd, $J = 22.5, 10.0$ Hz, 1H), 3.61 (d, $J = 10.5$ Hz, 3H), 2.03 (d, $J = 0.5$ Hz, 6H). ^{13}C NMR (100 MHz, DMSO- d_6) δ : 168.3, 165.5, 156.5, 144.9, 136.7, 135.7, 133.4, 129.3, 128.4, 128.1, 128.0, 120.5, 120.2, 52.2 (d, $\text{C}_1\text{-P}$, $J_{\text{CP}} = 157.0$ Hz), 52.3, 23.9. MS (ESI) m/z 554.3 = $[\text{M}+\text{H}]^+$.

(*S*)-2-(((Benzyloxy)carbonyl)amino)-2-(3-nitrophenyl)acetic acid (148). Procedure and characterization consistent with previously reported data.⁶⁰

(*S*)-Benzyl (2-amino-1-(3-nitrophenyl)-2-oxoethyl)carbamate (149). Procedure and characterization consistent with previously reported data.⁶¹

(*S*)-Benzyl (cyano(3-nitrophenyl)methyl)carbamate (150). General procedure **F** with (*S*)-benzyl (2-amino-1-(3-nitrophenyl)-2-oxoethyl)carbamate (**149**) (1.92 g, 5.83 mmol) to yield (*S*)-benzyl (cyano(3-nitrophenyl)methyl)carbamate (245 mg, 0.60 mmol, 13% yield). MS (ESI) m/z 312.2 $[\text{M}+\text{H}]^+$.

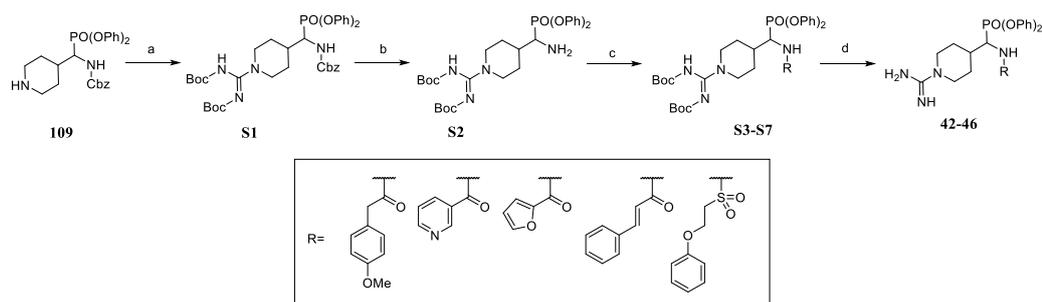
(*S*)-Benzyl ((3-aminophenyl)(cyano)methyl)carbamate (151). General procedure **K** with (*S*)-benzyl (cyano(3-nitrophenyl)methyl)carbamate (**150**) (245 mg, 0.79 mmol) to afford (*S*)-benzyl ((3-aminophenyl)(cyano)methyl)carbamate (215 mg, 0.78 mmol, 99% yield). ^1H NMR (400 MHz, CDCl_3) δ : 7.44

Appendix

- 7.31 (m, 5H), 7.23 - 7.11 (m, 1H), 6.83 (d, $J = 7.5$ Hz, 1H), 6.74 (d, $J = 11.0$ Hz, 1H), 6.69 (ddd, $J = 8.5, 5.0, 3.5$ Hz, 1H), 5.74 (d, $J = 8.5$ Hz, 1H), 5.41 - 5.28 (m, 1H), 5.16 (d, $J = 17.5$ Hz, 2H), 4.04 - 3.68 (m, 2H). ^{13}C NMR (100 MHz, CDCl_3) δ : 155.1, 147.5, 135.6, 134.2, 130.5, 128.8, 128.7, 128.5, 117.7, 116.7, 116.2, 113.1, 68.0, 46.7. MS (ESI) m/z 282.1 $[\text{M}+\text{H}]^+$.

Part 2. Synthetic procedure for UAMC library compounds – Chapter 2

A series of inactive novel UAMC library compounds were reported in the paper without specifying their synthetic procedures in order to maintain the manuscript coherence. Protocols followed to their synthesis can be found below. Every synthetic and characterization consideration, together with general procedure numbering will follow the specification of the main publication:



Scheme S1. Reagents and conditions. a) *N,N'*-bis-Boc-1-guanylpyrazole, Et₃N, DCM, rt, 48 h. b) Pd(II)/C (10%), H₂ gas, HCl (4 M in dioxane), MeOH, rt, 1 h. c) RCl, DIPEA, DCM, - 40 °C, 30 min. d) TFA, DCM, rt, 1 h.

***Tert*-butyl (Z)-((4-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate (S1).** General Procedure I with benzyl ((diphenoxyphosphoryl)(piperidin-4-yl)methyl)carbamate 2,2,2-trifluoroacetate (**109**) (9.26 g, 19.27 mmol) to give *tert*-butyl (Z)-((4-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate (6.31 g, 8.73 mmol, 45% yield). ¹H NMR (400 MHz, CDCl₃) δ: 7.40 - 7.28 (m, 7H), 7.25 - 7.11 (m, 6H), 7.10 - 7.03 (m, 2H), 5.51 (dd, *J* = 10.5, 3.0 Hz, 1H), 5.23 - 4.99 (m, 2H), 4.49 (ddd, *J* = 20.0, 10.5, 4.5 Hz, 1H), 4.39 - 3.97 (m, 2H), 2.93 (dd, *J* = 28.5, 13.5 Hz, 2H), 2.28 (td, *J* = 8.0, 4.0 Hz, 1H), 2.09 - 1.94 (m, 1H), 1.85 (d, *J* = 13.0 Hz, 1H), 1.72 - 1.51 (m, 2H), 1.48 (s, 18H). MS (ESI) *m/z* 723.1 [M+H]⁺.

(Z)-*Tert*-butyl ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate hydrochloride (S2). Pd(II)/C (10%) (73 mg, 0.17 mmol, 0.05 eq) was added to a stirred solution of *tert*-butyl (Z)-((4-(((benzyloxy)carbonyl)amino)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate (**S1**) (2.5 g, 3.46 mmol) and HCl (4 M in dioxane) (0.86 mL,

3.46 mmol, 1 eq) in anhydrous MeOH (10 mL). The reaction mixture was hydrogenated for 1 h, by bubbling H₂ gas (1 atm). The mixture was filtered through a pad of celite, which was additionally rinsed with DCM. The filtrate was concentrated under reduced pressure and the crude was purified by flash column chromatography (SiO₂, MeOH in DCM, 0/100 to 12/88). The desired fractions were collected and concentrated to yield (*Z*)-*tert*-butyl ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate hydrochloride (900 mg, 1.14 mmol, 42% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.59-7.04 (m, 10H), 4.25-4.11 (m, 3H), 3.54-3.40 (m, 2H), 2.70-2.51 (m, 1H), 2.45-2.03 (m, 4H), 1.49 (s, 18H). MS (ESI) *m/z* 590.0 [M+H]⁺.

(*Z*)-*Tert*-butyl (((*tert*-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(2-(4-methoxyphenyl)acetamido)methyl)piperidin-1-yl)methylene)carbamate (S3). DIPEA (1 M in DCM) (0.68 mL, 0.68 mmol, 2.5 eq) was added portionwise to a stirred suspension of (*Z*)-*tert*-butyl ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate hydrochloride (**S2**) (170 mg, 0.27 mmol) in anhydrous DCM (5 mL) at -40 °C. The reaction mixture was stirred for 3 min followed by dropwise addition of a solution of 4-methoxyphenylacetyl chloride (50 mg, 0.27 mmol, 1 eq) in anhydrous DCM (0.5 mL). The reaction mixture was stirred for 30 min at -40 °C and then allowed to reach the rt. The reaction mixture was concentrated under reduced pressure. The crude was purified by flash column chromatography (SiO₂, EtOAc in heptane, 0/100 to 100/0). The desired fractions were collected and concentrated to yield (*Z*)-*tert*-butyl (((*tert*-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(2-(4-methoxyphenyl)acetamido)methyl)piperidin-1-yl)methylene)carbamate (107 mg, 0.15 mmol, 53% yield). ¹H NMR (400 MHz, CDCl₃) δ: 10.32 (s, 1H), 7.48 (dd, *J* = 14.5, 7.0 Hz, 4H), 7.35 (t, *J* = 7.5 Hz, 2H), 7.32 - 7.24 (m, 4H), 7.18 (d, *J* = 8.0 Hz, 2H), 7.04 (d, *J* = 8.5 Hz, 2H), 6.12 (s, 1H), 5.01 (ddd, *J* = 19.5, 10.5, 4.5 Hz, 1H), 4.66 - 4.62 (m, 1H), 4.74 - 4.08 (m, 2H), 3.96 (s, 3H), 3.68 (dd, *J* = 36.5, 16.0 Hz, 2H), 3.30 - 2.83 (m, 2H), 2.53 - 2.26 (m, 1H), 2.00 (d, *J* = 13.5 Hz, 2H), 1.67 (s, 18H). MS (ESI) *m/z* 737.2 [M+H]⁺.

(*Z*)-*Tert*-butyl (((*tert*-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(*nicotinamido*)acetamido)methyl)piperidin-1-yl)methylene)carbamate (S4). Procedure followed for **S3** with (*Z*)-*tert*-butyl ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate hydrochloride (**S2**) (170 mg, 0.27 mmol) and 3-pyridinecarbonylchloride (50 mg, 0.27 mmol, 1 eq) to yield (*Z*)-*tert*-butyl (((*tert*-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(2-(4-methoxyphenyl)acetamido)methyl)piperidin-1-yl)methylene)carbamate (111 mg, 0.16 mmol, 55% yield). ¹H NMR (400 MHz, CDCl₃) δ: 10.12 (s, 1H), 8.80 (d, *J* = 50.0 Hz, 2H), 7.96 (d, *J* = 8.0 Hz, 1H), 7.46 - 7.29 (m, 3H), 7.25 - 7.17 (m, 4H), 7.10 (dd, *J* = 12.0, 7.5 Hz, 3H), 6.74-6.60 (m, 1H), 5.10 (ddd, *J* = 19.5, 10.5, 5.0 Hz, 1H), 4.59-3.78 (m, 2H), 3.01-2.84 (m, 2H), 2.46-2.26 (m, 1H), 2.13-2.00 (m, 1H), 1.96 (d, *J* = 13.0 Hz, 1H), 1.46 (s, 18H). MS (ESI) *m/z* 694.1 [M+H]⁺.

(Z)-Tert-butyl (((tert-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(furan-2-carboxamido)methyl)piperidin-1-yl)methylene)carbamate (S5). Procedure followed for **S3** with *(Z)-tert-butyl* ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((tert-butoxycarbonyl)amino)methylene)carbamate hydrochloride (**S2**) (170 mg, 0.27 mmol) and of 2-furoylchloride (36 mg, 0.27 mmol, 1 eq) to yield *(Z)-tert-butyl* (((tert-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(furan-2-carboxamido)methyl)piperidin-1-yl)methylene)carbamate (115 mg, 0.17 mmol, 61% yield). ¹H NMR (400 MHz, CDCl₃) δ: 10.10 (s, 1H), 7.46 (d, *J* = 0.5 Hz, 1H), 7.36 - 7.28 (m, 2H), 7.23 (d, *J* = 7.5 Hz, 2H), 7.21 - 7.10 (m, 7H), 6.67 (d, *J* = 10.5 Hz, 1H), 6.52 (dd, *J* = 3.5, 1.5 Hz, 1H), 4.99 (ddd, *J* = 19.5, 10.5, 5.0 Hz, 1H), 4.77 - 3.81 (m, 2H), 2.96 (dd, *J* = 27.5, 13.5 Hz, 2H), 2.53 - 2.25 (m, 1H), 2.09 (d, *J* = 13.0 Hz, 1H), 1.93 (d, *J* = 13.5 Hz, 1H), 1.74 - 1.54 (m, 2H), 1.46 (s, 18H).

Tert-butyl ((Z)-((tert-butoxycarbonyl)amino)(4-(cinnamamido(diphenoxyphosphoryl)methyl)piperidin-1-yl)methylene)carbamate (S6). Procedure followed for **S3** with *(Z)-tert-butyl* ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((tert-butoxycarbonyl)amino)methylene)carbamate hydrochloride (**S2**) (2.45 g, 3.92 mmol) and 2-furoylchloride (653 mg, 3.92 mmol, 1 eq) in anhydrous DCM (0.5 mL) to yield *(Z)-tert-butyl* (((tert-butoxycarbonyl)amino)(4-(cinnamamido(diphenoxyphosphoryl)methyl)piperidin-1-yl)methylene)carbamate (710 mg, 0.99 mmol, 25% yield). ¹H NMR (400 MHz, CDCl₃) δ: 10.06 (s, 1H), 7.63 (d, *J* = 15.5 Hz, 1H), 7.50 - 7.41 (m, 2H), 7.34 (dd, *J* = 9.5, 5.5 Hz, 5H), 7.21 (dd, *J* = 13.5, 7.0 Hz, 5H), 7.12 - 7.01 (m, 3H), 6.52 (d, *J* = 8.0 Hz, 1H), 6.43 (d, *J* = 15.0 Hz, 1H), 5.07 (ddd, *J* = 20.0, 10.5, 4.5 Hz, 1H), 4.48 - 3.91 (m, 2H), 2.93 (dt, *J* = 25.0, 12.0 Hz, 2H), 2.34 (dd, *J* = 7.5, 4.0 Hz, 1H), 2.06 (d, *J* = 16.0 Hz, 1H), 1.89 (d, *J* = 13.5 Hz, 1H), 1.69 - 1.54 (m, 2H), 1.46 (s, 18H). MS (ESI) *m/z* 719.5 [M+H]⁺.

(Z)-Tert-butyl (((tert-butoxycarbonyl)imino)(4-((diphenoxyphosphoryl)(2-phenoxyethylsulfonamido)methyl)piperidin-1-yl)methyl)carbamate (S7). Procedure followed for **S3** with *(Z)-tert-butyl* ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((tert-butoxycarbonyl)amino)methylene)carbamate hydrochloride (**S2**) (400 mg, 0.64 mmol) and 2-Phenoxyethanesulfonyl chloride (282 mg, 1.28 mmol, 2 eq) to yield *(Z)-tert-butyl* (((tert-butoxycarbonyl)imino)(4-((diphenoxyphosphoryl)(2-phenoxyethylsulfonamido)methyl)piperidin-1-yl)methyl)carbamate (50 mg, 0.06 mmol, 10% yield). MS (ESI) *m/z* 773.4 [M+H]⁺.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(2-(4-methoxyphenyl)acetamido)methyl)phosphonate 2,2,2-trifluoroacetate (42). General procedure **H** with *(Z)-tert-butyl* (((tert-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(2-(4-methoxyphenyl)acetamido)methyl)piperidin-1-yl)methylene)carbamate (**S3**) (107 mg, 0.15 mmol) to yield diphenyl ((1-carbamimidoylpiperidin-4-yl)((5-phenylpyrimidin-2-yl)amino)methyl)phosphonate 2,2,2-trifluoroacetate (35 mg, 0.05 mmol, 37% yield). ¹H NMR (400 MHz,

Methanol-*d*₄) δ : 7.33 (dt, $J = 16.0, 8.0$ Hz, 4H), 7.27-7.18 (m, 4H), 7.14 - 7.07 (m, 2H), 7.02 (dd, $J = 7.5, 1.0$ Hz, 2H), 6.87 - 6.79 (m, 2H), 4.77 (dd, $J = 18.5, 6.5$ Hz, 1H), 3.90 (t, $J = 14.5$ Hz, 2H), 3.74 (s, 3H), 3.54 (s, 2H), 3.19 - 3.02 (m, 2H), 2.39 (ddd, $J = 18.5, 9.5, 5.5$ Hz, 1H), 2.04 (d, $J = 13.0$ Hz, 2H), 1.49 (qd, $J = 13.0, 4.0$ Hz, 2H). MS (ESI) m/z 537.0 [M+H]⁺.

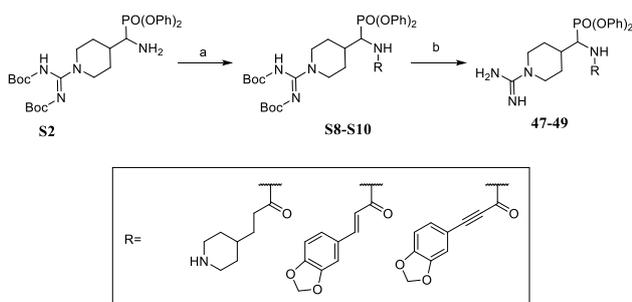
Diphenyl ((1-carbamimidoylpiperidin-4-yl)(nicotinamido)methyl)phosphonate 2,2,2-trifluoroacetate (43) General procedure **H** with (*Z*)-*tert*-butyl (((*tert*-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(nicotinamido)methyl)piperidin-1-yl)methylene)carbamate (**S4**) (111 mg, 0.16 mmol) to yield diphenyl ((1-carbamimidoylpiperidin-4-yl)(nicotinamido)methyl)phosphonate 2,2,2-trifluoroacetate (17 mg, 0.03 mmol, 17% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 8.6-5.73 (m, 2H), 8.14 (d, $J = 8.0$ Hz, 1H), 7.56 (d, $J = 4.5$ Hz, 1H), 7.42 - 7.28 (m, 4H), 7.25 - 7.15 (m, 6H), 5.02 (dd, $J = 17.5, 8.0$ Hz, 1H), 3.96 (d, $J = 13.5$ Hz, 2H), 3.17 (ddd, $J = 21.5, 14.0, 2.5$ Hz, 2H), 2.54 (ddd, $J = 11.5, 8.0, 3.5$ Hz, 2H), 2.27 (d, $J = 13.0$ Hz, 1H), 2.09 (d, $J = 13.5$ Hz, 1H), 1.76 - 1.46 (m, 2H). MS (ESI) m/z 494.0 [M+H]⁺.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(furan-2-carboxamido)methyl)phosphonate 2,2,2-trifluoroacetate (44). General procedure **H** with (*Z*)-*tert*-butyl (((*tert*-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)(furan-2-carboxamido)methyl)piperidin-1-yl)methylene)carbamate (**S5**) (115 mg, 0.17 mmol) to yield diphenyl ((1-carbamimidoylpiperidin-4-yl)(furan-2-carboxamido)methyl)phosphonate 2,2,2-trifluoroacetate (62 mg, 0.10 mmol, 62% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.72 (dd, $J = 1.5, 0.5$ Hz, 1H), 7.45 - 7.27 (m, 4H), 7.26 - 7.07 (m, 7H), 6.62 (dd, $J = 3.5, 1.5$ Hz, 1H), 3.94 (d, $J = 13.5$ Hz, 2H), 3.15 (td, $J = 15.5, 2.5$ Hz, 2H), 2.52 (ddd, $J = 16.0, 9.5, 6.0$ Hz, 1H), 2.26 (d, $J = 13.0$ Hz, 1H), 2.08 (d, $J = 13.5$ Hz, 1H), 1.72 - 1.39 (m, 2H). MS (ESI) m/z 483.0 [M+H]⁺.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(cinnamamido)methyl)phosphonate bis(2,2,2-trifluoroacetate) (45). General procedure **H** with *tert*-butyl ((*Z*)-((*tert*-butoxycarbonyl)imino)(4-(cinnamamido(diphenoxyphosphoryl)methyl)piperidin-1-yl)methyl)carbamate (**S6**) (172 mg, 0.22 mmol) to yield (*E*)-diphenyl ((1-carbamimidoylpiperidin-4-yl)(cinnamamido)methyl)phosphonate 2,2,2-trifluoroacetate (13 mg, 0.04 mmol, 16% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.62 (d, $J = 15.5$ Hz, 1H), 7.60 - 7.55 (m, 2H), 7.46 - 7.31 (m, 7H), 7.28 - 7.14 (m, 6H), 6.74 (d, $J = 15.5$ Hz, 1H), 4.96 (dd, $J = 18.5, 6.5$ Hz, 1H), 4.07 - 3.80 (m, 2H), 3.16 (td, $J = 15.5, 2.5$ Hz, 2H), 2.45 (ddd, $J = 18.5, 9.5, 5.5$ Hz, 1H), 2.12 (dd, $J = 9.5, 4.0$ Hz, 2H), 1.70 - 1.47 (m, 2H). MS (ESI) m/z 519.3 [M+H]⁺.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(2-phenoxyethylsulfonamido)methyl)phosphonate 2,2,2-trifluoroacetate (46). General procedure **H** with (*Z*)-*tert*-butyl (((*tert*-butoxycarbonyl)imino)(4-((diphenoxyphosphoryl)(2-phenoxyethylsulfonamido)methyl)piperidin-1-yl)methyl)carbamate (**S7**) (50 mg, 0.07 mmol) to yield diphenyl ((1-carbamimidoylpiperidin-4-yl)(2-phenoxyethylsulfonamido)methyl)phosphonate 2,2,2-trifluoroacetate (19 mg, 0.03 mmol, 43% yield) as a

white solid. ^1H NMR (400 MHz, Methanol- d_4) δ : 7.35-7.30 (m, 4H), 7.25-7.17 (m, 6H), 7.13-7.07 (m, 2H), 6.97-6.91 (m, 1H), 6.90-6.84 (m, 2H), 4.42 (t, $J = 6.5$ Hz, 2H), 4.35 (dd, $J = 19.0, 5.4$ Hz, 1H), 3.98 (dd, $J = 10.5, 3.5$ Hz, 2H), 3.68 (t, $J = 6.5$ Hz, 2H), 3.14 (td, $J = 15.0, 2.5$ Hz, 2H), 2.52 - 2.34 (m, 1H), 2.21-1.98 (m, 2H), 1.87-1.56 (m, 2H). MS (ESI) m/z 573.2 $[\text{M}+\text{H}]^+$.



Scheme S2. Reagents and conditions. a) HATU, ROH, DIPEA, DMF, rt, 4 h. b) TFA, DCM, rt, 1 h.

(Z)-Tert-butyl 4-(3-(((1-(*N,N'*-bis(*tert*-butoxycarbonyl)carbamimidoyl)piperidin-4-yl)(diphenoxyphosphoryl)methyl)amino)-3-oxopropyl)piperidine-1-carboxylate (S8). HATU (316 mg, 0.83 mmol, 2 eq) was added in one portion to a stirred solution of 3-(1-(*tert*-butoxycarbonyl)piperidin-4-yl)propanoic acid (214 mg, 0.83 mmol, 2 eq) in anhydrous DMF (2 mL). The reaction mixture was stirred at rt for 30 min. After this period of time, (*Z*)-*tert*-butyl ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)imino)methyl)carbamate hydrochloride (**S2**) (260 mg, 0.42 mmol) was added, followed by addition of *N,N*-Di-iso-propylethylamine (0.29 mL, 1.66 mmol). The reaction mixture was stirred at rt for 4 h. The reaction mixture was diluted with water, extracted with EtOAc, combined organic fractions were washed with water and brine, dried over Na_2SO_4 , filtered and concentrated under reduced pressure. The crude was purified by flash column chromatography (SiO_2 , EtOAc in heptane, 0/100 to 100/0). The desired fractions were collected and concentrated to yield (*Z*)-*tert*-butyl 4-(3-(((1-(*N,N'*-bis(*tert*-butoxycarbonyl)carbamimidoyl)piperidin-4-yl)(diphenoxyphosphoryl)methyl)amino)-3-oxopropyl)piperidine-1-carboxylate (257 mg, 0.31 mmol, 74% yield). ^1H NMR (400 MHz, CDCl_3) δ : 7.37 – 7.26 (m, 4H), 7.22 – 7.13 (m, 4H), 7.12 – 7.07 (m, 2H), 6.22 (d, $J = 10.0$ Hz, 1H), 4.92 (ddd, $J = 20.0, 10.5, 4.5$ Hz, 1H), 4.44 – 3.85 (m, 4H), 3.01 – 2.87 (m, 2H), 2.61 (dd, $J = 31.5, 17.5$ Hz, 2H), 2.22 (dddt, $J = 35.5, 22.5, 15.0, 7.5$ Hz, 3H), 2.03-2.01 (m, 1H), 1.80 (d, $J = 13.5$ Hz, 1H), 1.65 – 1.39 (m, 33H), 1.15 – 0.93 (m, 2H). MS (ESI) m/z 828.5 $[\text{M}+\text{H}]^+$.

(Z)-Tert-butyl (((4-(((*E*)-3-(benzo[d][1,3]dioxol-5-yl)acrylamido)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate (S9). Procedure followed for **S8** with (*Z*)-*tert*-butyl ((4-

(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate hydrochloride (**S2**) (300 mg, 0.48 mmol) and (*E*)-3-(benzo[d][1,3]dioxol-5-yl)acrylic acid (184 mg, 0.96 mmol, 2 eq) to yield (*Z*)-*tert*-butyl ((4-(((*E*)-3-(benzo[d][1,3]dioxol-5-yl)acrylamido)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate (253 mg, 0.33 mmol, 69% yield). ¹H NMR (400 MHz, CDCl₃) δ: 10.12 (s, 1H), 7.54 (d, *J* = 15.5 Hz, 1H), 7.34 (tt, *J* = 4.5, 2.0 Hz, 2H), 7.28 - 7.16 (m, 5H), 7.10 (t, *J* = 7.5 Hz, 3H), 6.97 (d, *J* = 6.5 Hz, 2H), 6.84 - 6.73 (m, 1H), 6.33 (s, 1H), 6.25 (d, *J* = 15.5 Hz, 1H), 5.99 (d, *J* = 7.0 Hz, 2H), 5.06 (ddd, *J* = 20.0, 10.5, 4.5 Hz, 1H), 4.59 - 3.91 (m, 2H), 3.08 - 2.84 (m, 2H), 2.43 - 2.27 (m, 1H), 2.11 - 2.03 (m, 1H), 1.88 (d, *J* = 13.0 Hz, 1H), 1.70 - 1.54 (m, 2H), 1.53 - 1.38 (m, 18H). MS (ESI) *m/z* 763.4 [M+H]⁺.

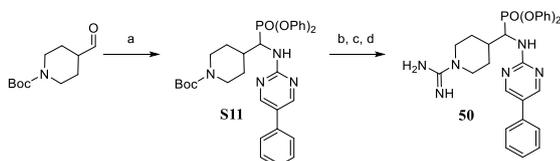
(Z)-Tert-butyl (E)-((4-((3-(benzo[d][1,3]dioxol-5-yl)propiolamido)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((tert-butoxycarbonyl)amino)methylene)carbamate (S10). Procedure followed for **S8** with (*Z*)-*tert*-butyl ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate hydrochloride (**S2**) (329 mg, 0.52 mmol) and 3-(benzo[d][1,3]dioxol-5-yl)propiolic acid (100 mg, 0.52 mmol, 1 eq) to yield (*Z*)-*tert*-butyl ((4-((3-(benzo[d][1,3]dioxol-5-yl)propiolamido)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate (137 mg, 0.18 mmol, 34% yield). ¹H NMR (400 MHz, CDCl₃) δ: 10.14 (s, 1H), 7.37 - 7.27 (m, 4H), 7.23 - 7.07 (m, 7H), 6.97 (d, *J* = 1.5 Hz, 1H), 6.82 - 6.77 (m, 1H), 6.37 (dd, *J* = 10.5, 2.5 Hz, 1H), 6.01 (s, 2H), 4.93 (ddd, *J* = 19.5, 10.5, 4.5 Hz, 1H), 4.39 - 3.96 (m, 2H), 3.09 - 2.78 (m, 2H), 2.41 - 2.25 (m, 1H), 2.10 - 2.00 (m, 1H), 1.88 (d, *J* = 13.0 Hz, 1H), 1.70 - 1.57 (m, 2H), 1.58 - 1.41 (m, 2H). MS (ESI) *m/z* 761.4 [M+H]⁺.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)(3-(piperidin-4-yl)propanamido)methyl)phosphonate bis(2,2,2-trifluoroacetate) (47). General procedure **H** with (*Z*)-*tert*-butyl 4-(3-(((1-(*N,N'*-bis(*tert*-butoxycarbonyl)carbamimidoyl)piperidin-4-yl)(diphenoxyphosphoryl)methyl)amino)-3-oxopropyl)piperidine-1-carboxylate (**S9**) (172 mg, 0.21 mmol) to yield diphenyl ((1-carbamimidoylpiperidin-4-yl)(3-(piperidin-4-yl)propanamido)methyl)phosphonate bis(2,2,2-trifluoroacetate) (47 mg, 0.06 mmol, 30% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.44 - 7.33 (m, 4H), 7.29 - 7.19 (m, 4H), 7.15 - 7.09 (m, 2H), 4.82 (dd, *J* = 18.0, 7.0 Hz, 1H), 4.01 - 3.87 (m, 2H), 3.28 (d, *J* = 2.5 Hz, 2H), 3.14 (td, *J* = 13.0, 2.0 Hz, 2H), 2.84 - 2.70 (m, 2H), 2.48 - 2.29 (m, 3H), 2.16 - 2.02 (m, 2H), 1.94 - 1.82 (m, 2H), 1.68 - 1.45 (m, 5H), 1.41 - 1.25 (m, 2H). MS (ESI) *m/z* 528.3 [M+H]⁺.

Diphenyl (E)-diphenyl ((3-(benzo[d][1,3]dioxol-5-yl)acrylamido)(1-carbamimidoylpiperidin-4-yl)methyl)phosphonate 2,2,2-trifluoroacetate (48). General procedure **H** with (*Z*)-*tert*-butyl ((4-(((*E*)-3-

(benzo[d][1,3]dioxol-5-yl)acrylamido)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate (253 mg, 0.33 mmol) to yield diphenyl (*E*)-diphenyl ((3-(benzo[d][1,3]dioxol-5-yl)acrylamido)(1-carbamimidoylpiperidin-4-yl)methyl)phosphonate 2,2,2-trifluoroacetate (6 mg, 9.10 μ mol, 3% yield). $^1\text{H NMR}$ (400 MHz, Methanol- d_4) δ : 7.53 (d, $J = 15.5$ Hz, 1H), 7.36 (dd, $J = 17.0, 8.5$ Hz, 4H), 7.28 - 7.10 (m, 7H), 7.06 (dd, $J = 8.0, 1.5$ Hz, 1H), 6.86 (d, $J = 8.0$ Hz, 1H), 6.55 (d, $J = 15.5$ Hz, 1H), 6.01 (s, 2H), 5.04 - 4.90 (m, 1H), 3.94 (t, $J = 11.0$ Hz, 2H), 3.16 (dd, $J = 24.0, 13.0$ Hz, 2H), 2.56 - 2.32 (m, 1H), 2.20 - 1.99 (m, 2H), 1.70 - 1.48 (m, 2H). MS (ESI) m/z 563.2 $[\text{M}+\text{H}]^+$.

Diphenyl ((3-(benzo[d][1,3]dioxol-5-yl)propiolamido)(1-carbamimidoylpiperidin-4-yl)methyl)phosphonate (49). General procedure **H** with (*Z*)-*tert*-butyl ((4-((3-(benzo[d][1,3]dioxol-5-yl)propiolamido)(diphenoxyphosphoryl)methyl)piperidin-1-yl)((*tert*-butoxycarbonyl)amino)methylene)carbamate (**S10**) (135 mg, 0.18 mmol) to yield diphenyl ((3-(benzo[d][1,3]dioxol-5-yl)propiolamido)(1-carbamimidoylpiperidin-4-yl)methyl)phosphonate (9 mg, 0.01 mmol, 8% yield). $^1\text{H NMR}$ (400 MHz, Methanol- d_4) δ : 7.38 (td, $J = 8.0, 3.0$ Hz, 4H), 7.28 - 7.12 (m, 7H), 7.04 (d, $J = 1.5$ Hz, 1H), 6.89 (d, $J = 8.0$ Hz, 1H), 6.04 (s, 2H), 4.01 - 3.88 (m, 2H), 3.22 - 3.06 (m, 2H), 2.53 - 2.38 (m, 1H), 2.12 (t, $J = 13.5$ Hz, 2H), 1.69 - 1.48 (m, 2H). MS (ESI) m/z 561.2 $[\text{M}+\text{H}]^+$.

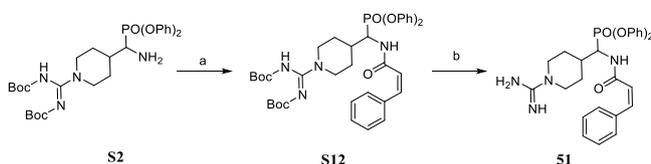


Scheme S3. Reagents and conditions. a) 5-phenylpyrimidin-2-amine, $\text{P}(\text{OPh})_3$, $\text{Cu}(\text{OTf})_2$, DCM, rt, 16 h. b) TFA, DCM, rt, 1 h. c) *N,N'*-bis-Boc-1-guanylpiperazole, Et_3N , DCM, rt, 48 h. d) TFA, DCM, rt, 1 h.

***Tert*-butyl 4-((diphenoxyphosphoryl)((5-phenylpyrimidin-2-yl)amino)methyl)piperidine-1-carboxylate (S11).** General procedure **C** with *tert*-butyl 4-formylpiperidine-1-carboxylate (706 mg, 3.31 mmol) and 5-phenylpyrimidin-2-amine (567 mg, 3.31 mmol, 1 eq) to give *tert*-butyl 4-((diphenoxyphosphoryl)((5-phenylpyrimidin-2-yl)amino)methyl)piperidine-1-carboxylate (610 mg, 1.02 mmol, 31% yield) as a white solid. MS (ESI) m/z 623.2 $[\text{M}+\text{Na}]^+$.

Diphenyl ((1-carbamimidoylpiperidin-4-yl)((5-phenylpyrimidin-2-yl)amino)methyl)phosphonate 2,2,2-trifluoroacetate (50). General procedure **H**, general procedure **I** and general procedure **H** again with *tert*-butyl 4-((diphenoxyphosphoryl)((5-phenylpyrimidin-2-yl)amino)methyl)piperidine-1-carboxylate (**S11**) (600 mg, 1.00 mmol) to yield diphenyl ((1-carbamimidoylpiperidin-4-yl)((5-phenylpyrimidin-2-yl)amino)methyl)phosphonate 2,2,2-trifluoroacetate (261 mg, 0.48 mmol, 48% yield). $^1\text{H NMR}$ (400 MHz,

DMSO-*d*₆) δ : 8.79 – 8.59 (m, 2H), 8.07 (d, $J = 10.0$ Hz, 1H), 7.73 – 7.57 (m, 2H), 7.53 – 7.43 (m, 2H), 7.26 – 7.06 (m, 5H), 5.11 (ddd, $J = 17.0, 10.5, 7.0$ Hz, 1H), 7.36 (dt, $J = 12.5, 4.0$ Hz, 7H), 3.90 (t, $J = 15.0$ Hz, 2H), 3.19–2.95 (m, 2H), 2.49 – 2.42 (m, 1H), 2.00 (t, $J = 10.5$ Hz, 2H), 1.65–1.39 (m, 2H), MS (ESI) m/z 543.2 $[M+H]^+$.

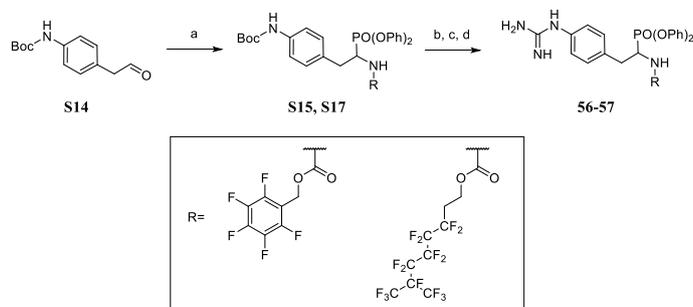


Scheme S4. Reagents and conditions. a) Oxalyl chloride, (*Z*)-3-phenylacrylic acid, DIPEA, DCM, 0-25 °C, 1 h. b) TFA, DCM, rt, 1 h.

(*Z*)-*Tert*-butyl (((*tert*-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)((*Z*)-3-phenylacrylamido)methyl)piperidin-1-yl)methylene)carbamate (S12). A solution of oxalyl chloride (0.12 mL, 1.40 mmol, 2 eq) in anhydrous DCM (1 mL) was added portion-wise to a stirred solution (*Z*)-3-phenylacrylic acid (100 mg, 0.68 mmol, 1 eq) in anhydrous DCM (2 mL) at 0 °C. The resulting mixture was stirred at 0 °C for 1 h. The solvents were evaporated and the crude was dissolved in DCM (2 mL). This mixture was added portionwise to a stirred solution of DIPEA (0.59 mL, 3.37 mmol, 5 eq) and (*Z*)-*tert*-butyl ((4-(amino(diphenoxyphosphoryl)methyl)piperidin-1-yl))((*tert*-butoxycarbonyl)amino)methylene)carbamate hydrochloride (**S2**) (422 mg, 0.68 mmol) in anhydrous DCM (5 mL) at 0 °C. The resulting mixture was stirred at rt for 1 h. The reaction mixture concentrated, dissolved in water, extracted with EtOAc, combined organic fractions were washed with water and brine, dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude was purified by flash column chromatography (SiO₂, EtOAc in heptane, 0/100 to 100/0). The desired fractions were collected and concentrated to yield (*Z*)-*tert*-butyl (((*tert*-butoxycarbonyl)amino)(4-((diphenoxyphosphoryl)((*Z*)-3-phenylacrylamido)methyl)piperidin-1-yl)methylene)carbamate (120 mg, 0.17 mmol, 25% yield). ¹H NMR (400 MHz, CDCl₃) δ : 10.17 (s, 1H), 7.49 – 7.40 (m, 2H), 7.36 – 7.24 (m, 7H), 7.17 (dtd, $J = 8.5, 8.0, 1.0$ Hz, 4H), 7.08 – 7.00 (m, 2H), 6.88 (d, $J = 12.5$ Hz, 1H), 5.98 (d, $J = 7.5$ Hz, 1H), 5.95 (d, $J = 12.0$ Hz, 1H), 4.91 (ddd, $J = 20.0, 10.5, 4.0$ Hz, 1H), 4.47 – 3.87 (m, 2H), 3.04 – 2.72 (m, 2H), 2.29 – 2.08 (m, 1H), 1.89 – 1.70 (m, 2H), 1.62 – 1.40 (m, 20H). MS (ESI) m/z 719.4 $[M+H]^+$.

(*Z*)-Diphenyl ((1-carbamimidoylpiperidin-4-yl)(3-phenylacrylamido)methyl)phosphonate 2,2,2-trifluoroacetate (51). General procedure **H** with (*Z*)-*tert*-butyl ((4-(((*E*)-3-(benzo[d][1,3]dioxol-5-yl)acrylamido)(diphenoxyphosphoryl)methyl)piperidin-1-yl))((*tert*-butoxycarbonyl)amino)methylene)carbamate (**S12**) (200 mg, 0.28 mmol) to yield (*Z*)-diphenyl ((1-carbamimidoylpiperidin-4-yl)(3-phenylacrylamido)methyl)phosphonate 2,2,2-trifluoroacetate (51 mg, 0.08 mmol, 30% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ : 7.52 (dd, $J = 7.5, 1.5$ Hz, 2H), 7.36 (t, $J = 8.0$ Hz,

4H), 7.27 – 7.18 (m, 5H), 7.17 – 7.11 (m, 4H), 6.87 (d, $J = 12.5$ Hz, 1H), 6.10 (dd, $J = 12.5, 1.0$ Hz, 1H), 4.92 – 4.88 (m, 1H), 3.91 (d, $J = 14.0$ Hz, 2H), 3.20 – 3.00 (m, 2H), 2.48 – 2.27 (m, 1H), 2.05 (dd, $J = 25.5, 14.5$ Hz, 2H), 1.65 – 1.36 (m, 2H). MS (ESI) m/z 519.3 $[M+H]^+$.



Scheme S5. Reagents and conditions. a) RNH_2 , $P(OPh)_3$, $Cu(OTf)_2$, DCM, rt, 16 h. b) TFA, DCM, rt, 1 h. c) N,N' -bis-Boc-1-guanylpyrazole, Et_3N , DCM, rt, 48 h. d) TFA, DCM, rt, 1 h.

Perfluorobenzyl carbamate (S13). Procedure and characterization consistent with previously reported data.¹

***Tert*-butyl 4-(2-oxoethyl)phenylcarbamate (S14).** Procedure and characterization consistent with previously reported data.²

(Perfluorophenyl)methyl (1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethyl)carbamate (S15).

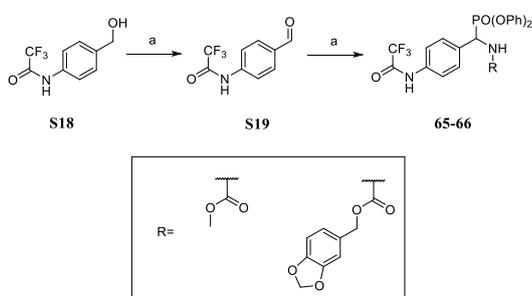
General procedure C with *tert*-butyl 4-(2-oxoethyl)phenylcarbamate (S14) (488 mg, 2.07 mmol) and perfluorobenzyl carbamate (S13) (500 mg, 2.07 mmol, 1 eq) to give (perfluorophenyl)methyl (2-(4-((*tert*-butoxycarbonyl)amino)phenyl)-1-(diphenoxyphosphoryl)ethyl)carbamate (260 mg, 0.38 mmol, 18% yield). 1H NMR (400 MHz, $CDCl_3$) δ : 7.30-7.00 (m, 14H), 6.41 (s, 1H), 5.09 (s, 2H), 5.02 (m, 1H), 4.70 (m, 1H), 3.33 (m, 1H), 2.97 (m, 1H), 1.52 (s, 9H). MS (ESI) m/z 715.2 $[M + Na]^+$

3,3,4,4,5,5,6,6,7,8,8,8-Dodecafluoro-7-(trifluoromethyl)octyl carbamate (S16). 1H,1H,2H,2H-Perfluoro-7-methyloctan-1-ol (0.60 mL, 2.42 mmol) was dissolved in anhydrous DCM (10 mL) and trichloroacetylisocyanate (0.35 mL, 2.90 mmol, 1.2 eq) was added at 0 °C and the mixture stirred for 2 h at this temperature. Then, solvent was evaporated and the mixture was re-dissolved in MeOH (30 mL) and H_2O (3 mL). Then, K_2CO_3 (1.00 g, 7.26 mmol, 3 eq) was added and reacted at rt for 16 h. The solvent was evaporated and H_2O (30 mL) was added. The aqueous layer was extracted with EtOAc, the combined organic layers washed with brine, dried over Na_2SO_4 , filtered and solvent evaporated under reduced pressure to yield 3,3,4,4,5,5,6,6,7,8,8,8-dodecafluoro-7-(trifluoromethyl)octyl carbamate (900 mg, 1.97 mmol, 82% yield) was formed. 1H NMR (400 MHz, $CDCl_3$) δ : 4.35 (t, $J = 6.5$ Hz, 2H), 2.59 (m, 2H). MS (ESI) m/z 479.9 $[M+Na]^+$.

***Tert*-butyl (4-(2-(diphenoxyphosphoryl)-2-(((3,3,4,4,5,5,6,6,7,8,8,8-dodecafluoro-7-(trifluoromethyl)octyl)oxy)carbonyl)amino)ethyl)phenyl)carbamate (S17).** General Procedure C with *tert*-butyl 4-(2-oxoethyl)phenylcarbamate (S14) (257 mg, 1.09 mmol) and 3,3,4,4,5,5,6,6,7,8,8,8-dodecafluoro-7-(trifluoromethyl)octyl carbamate (S16) (500 mg, 1.09 mmol, 1 eq) to give *tert*-butyl (4-(2-(diphenoxyphosphoryl)-2-(((3,3,4,4,5,5,6,6,7,8,8,8-dodecafluoro-7-(trifluoromethyl)octyl)oxy)carbonyl)amino)ethyl)phenyl)carbamate (300 mg, 0.33 mmol, 30% yield). ¹H NMR (400 MHz, CDCl₃) δ: 7.41-7.05 (m, 14H), 6.41 (s, 1H), 5.03 (m, 1H), 4.70 (m, 1H), 4.25 (t, *J* = 6.5 Hz, 2H), 3.37 (m, 1H), 2.98 (m, 1H), 2.32 (m, 2H), 1.56 (s, 9H). MS (ESI) *m/z* 931.3 [M + Na]⁺

(Perfluorophenyl)methyl (1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethyl)carbamate 2,2,2-trifluoroacetate (56). General procedure H, general procedure I and general procedure H again with (perfluorophenyl)methyl (2-(4-((*tert*-butoxycarbonyl)amino)phenyl)-1-(diphenoxyphosphoryl)ethyl)carbamate (S15) (260 mg, 0.38 mmol) to yield (perfluorophenyl)methyl (1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethyl)carbamate 2,2,2-trifluoroacetate (160 mg, 0.21 mmol, 55% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.5-7.0 (m, 14H), 5.1 (m, 2H), 4.60 (m, 1H), 4.57 (m, 1H), 3.4 (m, 1H), 3.08 (m, 1H). MS (ESI) *m/z* 635.1 [M+H]⁺.

3,3,4,4,5,5,6,6,7,8,8,8-Dodecafluoro-7-(trifluoromethyl)octyl 1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethylcarbamate 2,2,2-trifluoroacetate (57). General procedure H, general procedure I and general procedure H again *tert*-butyl (4-(2-(diphenoxyphosphoryl)-2-(((3,3,4,4,5,5,6,6,7,8,8,8-dodecafluoro-7-(trifluoromethyl)octyl)oxy)carbonyl)amino)ethyl)phenyl)carbamate (S17) (300 mg, 0.33 mmol) to yield 3,3,4,4,5,5,6,6,7,8,8,8-dodecafluoro-7-(trifluoromethyl)octyl 1-(diphenoxyphosphoryl)-2-(4-guanidinophenyl)ethylcarbamate 2,2,2-trifluoroacetate (71 mg, 0.07 mmol, 21% yield). ¹H NMR (400 MHz, CDCl₃) δ: 10.02 (s, 1H), 7.41-7.12 (m, 14H), 5.43 (m, 1H), 4.69 (m, 1H), 4.26 (m, 2H), 3.38 (m, 1H), 3.08 (m, 1H), 2.35 (m, 2H). MS (ESI) *m/z* 851.1 [M+H]⁺



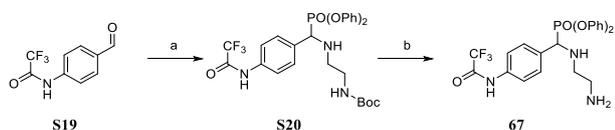
Scheme S6. Reagents and conditions. a) Dess-Martin periodinane, DCM, 0-25 °C, 2 h. b) RNH₂, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h.

2,2,2-Trifluoro-*N*-(4-(hydroxymethyl)phenyl)acetamide (S18). Procedure and characterization consistent with previously reported data.³

2,2,2-Trifluoro-*N*-(4-formylphenyl)acetamide (S19). General procedure **B** with 2,2,2-trifluoro-*N*-(4-(hydroxymethyl)phenyl)acetamide (**S18**) (1.00 g, 4.56 mmol) to yield 2,2,2-trifluoro-*N*-(4-formylphenyl)acetamide (450 mg, 2.07 mmol, 45% yield) as a colourless oil. ¹H NMR (400 MHz, Methanol-*d*₄) δ: 9.96 (s, 1H), 8.01-7.95 (m, 4H). No ionization found.

Methyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (65). General procedure **C** with 2,2,2-trifluoro-*N*-(4-formylphenyl)acetamide (**S19**) (70 mg, 0.32 mmol), and methyl carbamate (24 mg, 0.32 mmol) to give methyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (56 mg, 0.11 mmol, 34% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.70 (m, 1H), 7.55 (m, 2H), 7.41 -7.07 (m, 10H), 5.6 (d, 1H, *J* = 20.0 Hz), 3.75 (s, 3H). MS (ESI) *m/z* 531.1 [M+Na]⁺.

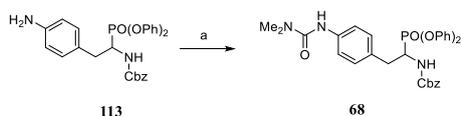
Benzo[d][1,3]dioxol-5-ylmethyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (66). General procedure **C** with 2,2,2-trifluoro-*N*-(4-formylphenyl)acetamide (**S19**) (70 mg, 0.32 mmol) and benzo[d][1,3]dioxol-5-ylmethyl carbamate (90 mg, 0.46 mmol). To give benzo[d][1,3]dioxol-5-ylmethyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (69 mg, 0.11 mmol, 34% yield). ¹H NMR (400 MHz, Methanol-*d*₄) δ: 7.70-7.50 (m, 4H), 7.40-6.80 (m, 13H), 5.80 (m, 2H), 5.60 (d, *J* = 28.0 Hz, 1H), 4.90 (s, 2H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ: 156.6, 156.1, 150.0, 149.7, 147.3, 146.8, 136.2, 131.5, 131.1, 130.0, 129.9, 125.5, 125.4, 122.1, 121.7, 121.1, 120.4, 120.4, 120.4, 120.0, 119.2 – 113.6 (m, CF₃), 108.5, 108.1, 101.4, 100.8, 66.1, 52.4 (d, C₁-P, *J* = 158.0 Hz). MS (ESI) *m/z* 629.2 [M+H]⁺.



Scheme S7. Reagents and conditions. a) RNH₂, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h. b) TFA, DCM, rt, 1 h.

2-((*Tert*-butoxycarbonyl)amino)ethyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (S20). General procedure **C** with 2,2,2-trifluoro-*N*-(4-formylphenyl)acetamide (**S19**) (70 mg, 0.32 mmol) and *tert*-butyl (2-(carbamoyloxy)ethyl)carbamate (66 mg, 0.32 mmol, 1 eq) to give 2-((*tert*-butoxycarbonyl)amino)ethyl((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (60 mg, 0.09 mmol, 29% yield). MS (ESI) *m/z* 638.3 [M-H]⁻

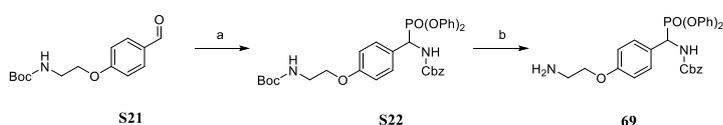
2-Aminoethyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate 2,2,2-trifluoroacetate (67). General procedure H with 2-((*tert*-butoxycarbonyl)amino)ethyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate (250 mg, 0.39 mmol) to yield 2-aminoethyl ((diphenoxyphosphoryl)(4-(2,2,2-trifluoroacetamido)phenyl)methyl)carbamate 2,2,2-trifluoroacetate (190 mg, 0.29 mmol, 74% yield). $^1\text{H NMR}$ (400 MHz, Methanol- d_4) δ : 7.75-7.5 (m, 4H), 7.40-6.93 (m, 10H), 5.72 (m, 2H), 4.43 (m, 2H), 3.44 (m, 2H). MS (ESI) m/z 538.2 $[\text{M-H}]^-$.



Scheme S8. Reagents and conditions. a) Dimethylcarbonylchloride, DIPEA, DCM, rt, 48 h

Benzyl 2-(4-(3,3-dimethylureido)phenyl)-1-(diphenoxyphosphoryl)ethylcarbamate (68).

Dimethylcarbonyl chloride ($\text{L} \square 93$, 1.02 mmol, 3.2 eq) was added to a mixture of benzyl 2-(4-aminophenyl)-1-(diphenoxyphosphoryl)ethylcarbamate (**113**) (160 mg, 0.32 mmol), DIPEA (0.08 mL, 0.48 mmol, 1.5 eq) in anhydrous DCM (15 mL) and the reaction mixture was stirred for 48 h at rt. Then, the solvent was evaporated and the crude was then purified by column chromatography (SiO_2 , EtOAc in heptane: 0/100 to 100/0). The desired fractions were collected and concentrated to yield benzyl 2-(4-(3,3-dimethylureido)phenyl)-1-(diphenoxyphosphoryl)ethylcarbamate (60 mg, 0.11 mmol, 33% yield) as a white solid. $^1\text{H NMR}$ (400 MHz, CDCl_3) δ : 7.05 – 7.38 (m, 19H), 6.39 (s, 1H), 5.34 (d, $J = 10.5$ Hz, 1H), 4.95 – 5.10 (m, 2H), 4.69 – 4.84 (m, 1H), 3.36 (ddd, $J = 4.5, 10.0, 14.5$ Hz, 1H), 3.02 (s, 6H), 1.28 (s, 1H). MS (ESI) m/z 574.7 $[\text{M+H}]^+$.



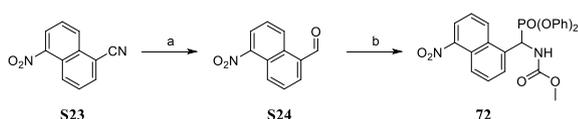
Scheme S9. Reagents and conditions. a) *Tert*-butyl (2-(4-formylphenoxy)ethyl)carbamate, P(OPh)_3 , Cu(OTf)_2 , DCM, rt, 16 h. b) TFA, DCM, rt, 1 h.

***Tert*-butyl (2-(4-formylphenoxy)ethyl)carbamate (S21).** Procedure and characterization consistent with previously reported data.⁴

Benzyl ((4-(2-((*tert*-butoxycarbonyl)amino)ethoxy)phenyl)(diphenoxyphosphoryl)methyl)carbamate (S22). General procedure C with *tert*-butyl 2-(4-formylphenoxy)ethylcarbamate (650 mg, 2.45 mmol) to yield

benzyl ((4-(2-((*tert*-butoxycarbonyl)amino)ethoxy)phenyl)(diphenoxyphosphoryl)methyl)carbamate (120 mg, 0.18 mmol, 7% yield). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ : 7.52-6.60 (m, 19 H), 5.90 (br t, 1H), 5.56 (m, 1H), 5.12 (m, 2H), 4.01 (m, 2H), 3.51 (m, 2H), 1.54 (9H). MS (ESI) m/z 655.2 $[\text{M}+\text{Na}]^+$.

Benzy ((4-(2-aminoethoxy)phenyl)(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (69). General procedure **H** with benzyl ((4-(2-((*tert*-butoxycarbonyl)amino)ethoxy)phenyl)(diphenoxyphosphoryl)methyl)carbamate (**S22**) (120 mg, 0.19 mmol) to yield benzyl ((4-(2-aminoethoxy)phenyl)(diphenoxyphosphoryl)methyl)carbamate 2,2,2-trifluoroacetate (20 mg, 0.031 mmol, 16% yield) as a colourless oil. $^1\text{H NMR}$ (400 MHz, CDCl_3) δ : 7.60-6.80 (m, 19H), 5.58 (d, $J = 22.5$ Hz, 1H), 5.15 (m, 2H), 4.25 (t, $J = 5.0$ Hz, 2H), 3.37 (t, $J = 5.0$ Hz, 2H). MS (ESI) m/z 533.1 $[\text{M}+\text{H}]^+$



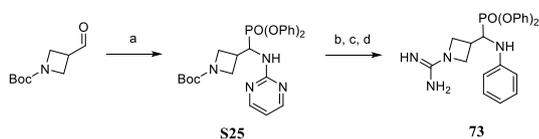
Scheme S10. Reagents and conditions. a) DIBAL-H, DCM, 0-25 °C, 16 h. b) Methyl carbamate, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h.

5-Nitro-1-naphthonitrile (S23). Procedure and characterization consistent with previously reported data.⁵

5-Nitro-1-naphthaldehyde (S24). DIBAL-H (2.31 mL, 2.78 mmol, 1.1 eq) was added dropwise to a suspension of 5-nitro-1-naphthonitrile (**S23**) (500 mg, 2.52 mmol) in anhydrous DCM (20 mL) at -5 °C. After the addition the reaction was left warm to rt and was stirred at this temperature for 16 h. The mixture was poured into HCl 2N sol and extracted with DCM. The combined organic phases were washed with brine, dried over Na₂SO₄ and evaporated to obtain the crude aldehyde. The crude was purified by column chromatography (SiO₂, EtOAc in heptane: 15/85 to 30/70). The desired fractions were collected and concentrated to yield 5-nitro-1-naphthaldehyde (200 mg, 0.99 mmol, 39% yield) as a yellow oil. $^1\text{H NMR}$ (400 MHz, CDCl_3) δ : 10.42 (s, 1H), 9.67 (dd, $J = 8.5, 0.9$ Hz, 1H), 8.80 (dt, $J = 8.5, 1.0$ Hz, 1H), 8.29 (dd, $J = 7.5, 1.0$ Hz, 1H), 8.15 (dd, $J = 7.0, 1.0$ Hz, 1H), 7.91 (dd, $J = 8.5, 7.0$ Hz, 1H), 7.79 (dd, $J = 8.5, 7.5$ Hz, 1H).

Methyl ((diphenoxyphosphoryl)(5-nitronaphthalen-1-yl)methyl)carbamate (72). General procedure **C** with 5-nitro-1-naphthaldehyde (**S24**) (500 mg, 2.48 mmol) and methyl carbamate (187 mg, 2.49 mmol, 1 eq) to give methyl ((diphenoxyphosphoryl)(5-nitronaphthalen-1-yl)methyl)carbamate (600 mg, 1.22 mmol, 49%

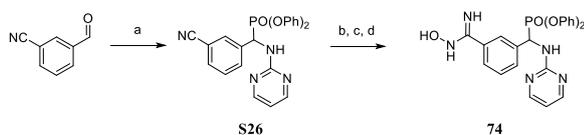
yield). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ : 8.62-8.51 (m, 1H), 8.46-8.31 (m, 1H), 8.24-8.17 (m, 1H), 8.03-7.91 (m, 1H), 7.74-7.58 (m, 2H), 7.35-7.01 (m, 10H), 6.5-6.35 (m, 1H), 6.11-6.01 (m, 1H). MS (ESI) m/z 493.1 $[\text{M}+\text{H}]^+$.



Scheme S11. Reagents and conditions. a) 2-aminopyrimidine, $\text{P}(\text{OPh})_3$, $\text{Cu}(\text{OTf})_2$, DCM, rt, 16 h. b) TFA, DCM, rt, 1 h. c) *N,N'*-bis-Boc-1-guanylpyrazole, Et_3N , DCM, rt, 48 h. d) TFA, DCM, rt, 1 h.

***Tert*-butyl 3-((diphenoxyphosphoryl)(pyrimidin-2-ylamino)methyl)azetidine-1-carboxylate (S25).** General procedure C with *tert*-butyl 3-formylazetidine-1-carboxylate (613 mg, 3.31 mmol) and 2-aminopyrimidine (315 mg, 3.31 mmol, 1 eq) to give *tert*-butyl 3-((diphenoxyphosphoryl)(pyrimidin-2-ylamino)methyl)azetidine-1-carboxylate (1.25 g, 2.52 mmol, 76% yield). MS (ESI) m/z 497.4 $[\text{M}+\text{H}]^+$.

Diphenyl ((1-carbamimidoylazetidin-3-yl)(pyrimidin-2-ylamino)methyl)phosphonate 2,2,2-trifluoroacetate (73). General procedure H, general procedure I and general procedure H again with *tert*-butyl 3-((diphenoxyphosphoryl)(pyrimidin-2-ylamino)methyl)azetidine-1-carboxylate (S25) (1.08 g, 1.96 mmol) to yield diphenyl ((1-carbamimidoylazetidin-3-yl)(pyrimidin-2-ylamino)methyl)phosphonate 2,2,2-trifluoroacetate (29 mg, 0.07 mmol, 4% yield). $^1\text{H NMR}$ (400 MHz, $\text{DMSO}-d_6$) δ : 8.47–8.24 (m, 2H), 8.08 (d, $J = 9.5$ Hz, 1H), 7.37 (td, $J = 8.0, 3.0$ Hz, 4H), 7.29 (s, 3H), 7.21 (td, $J = 7.5, 3.5$ Hz, 2H), 7.11 (d, $J = 8.0$ Hz, 3H), 6.73 (t, $J = 5.0$ Hz, 1H), 5.37 (dt, $J = 15.5, 9.5$ Hz, 1H), 4.13 – 4.26 (m, 2H), 4.07 (ddd, $J = 9.5, 6.0, 3.5$ Hz, 2H), 3.62-3.44 (m, 1H). MS (ESI) m/z 439.2 $[\text{M}+\text{H}]^+$.

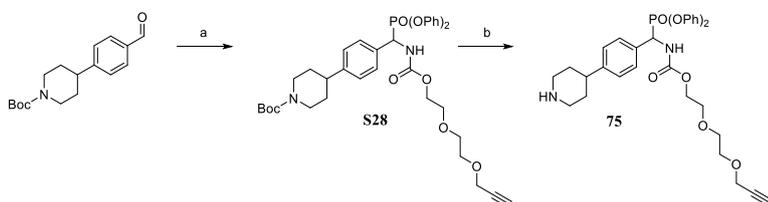


Scheme S12. Reagents and conditions. a) 2-aminopyrimidine, $\text{P}(\text{OPh})_3$, $\text{Cu}(\text{OTf})_2$, DCM, rt, 16 h. b) $\text{NH}_2\text{OH}\cdot\text{HCl}$, DIPEA, MeCN, 85 °C, 48 h.

Diphenyl ((3-cyanophenyl)(pyrimidin-2-ylamino)methyl)phosphonate (S26). General procedure C with 2-aminopyrimidine (1.05 g, 11.0 mmol) and 3-cyanobenzaldehyde (1.44 g, 11.0 mmol). To yield diphenyl ((3-cyanophenyl)(pyrimidin-2-ylamino)methyl)phosphonate (3.08 g, 6.96 mmol, 63% yield). $^1\text{H NMR}$ (400 MHz, $\text{DMSO}-d_6$) δ : 8.67 (d, $J = 8.5$ Hz, 1H), 8.36 (bs, 2H), 8.26 (s, 1H), 8.08 (d, $J = 7.5$ Hz, 1H), 7.82 (d, $J = 7.5$ Hz, 7H), 7.62 (t, $J = 7.5$ Hz, 1H), 7.37 – 7.27 (m, 4H), 7.16 (ddd, $J = 8.5, 7.5, 3.5$ Hz, 2H), 7.07 – 6.94 (m,

4H), 6.79 – 6.69 (m, 1H), 6.37 (dd, $J = 23.5, 9.5$ Hz, 1H). ^{13}C NMR (100 MHz, $\text{DMSO-}d_6$) δ : 161.2, 158.2, 150.0, 149.8, 136.9, 133.6, 132.2, 131.9, 129.9, 129.4, 125.4, 120.2, 118.9, 118.6, 115.2, 111.5, 51.7 (d, (d, $\text{C}_1\text{-P}$, $J = 155.5$ Hz).

Diphenyl ((3-(*N*-hydroxycarbamimidoyl)phenyl)(pyrimidin-2-ylamino)methyl)phosphonate (74). A mixture of diphenyl ((3-cyanophenyl)(pyrimidin-2-ylamino)methyl)phosphonate (**S26**) (1.02 g, 2.26 mmol), hydroxylammonium chloride (312 mg, 4.52 mmol, 2 eq) and DIPEA (0.79 mL, 4.52 mmol, 2 eq) in MeCN (10 mL) was heated to 85 °C. The reaction was stopped after 48 h. The crude was filtrated, the filtrate was evaporated and the crude was purified by reverse phase column chromatography (C18, MeOH in H_2O 0/100 to 100/0). The desired fractions were collected and concentrated to yield diphenyl ((3-(*N*-hydroxycarbamimidoyl)phenyl)(pyrimidin-2-ylamino)methyl)phosphonate (526 mg, 1.11 mmol, 49% yield). ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ : 9.67 (s, 1H), 8.47 (dd, $J = 10.5, 2.0$ Hz, 1H), 8.37 (d, $J = 4.5$ Hz, 2H), 8.05 (q, $J = 2.0$ Hz, 1H), 7.72 – 7.81 (m, 1H), 7.66 (dq, $J = 8.0, 1.5$ Hz, 1H), 7.41 (t, $J = 8.0$ Hz, 1H), 7.29 – 7.36 (m, 4H), 7.14 – 7.21 (m, 2H), 7.04 (dq, $J = 7.8, 1.2$ Hz, 2H), 6.98 (dq, $J = 8.0, 1.0$ Hz, 2H), 6.71 (t, $J = 5.0$ Hz, 1H), 6.26 (dd, $J = 22.5, 10.5$ Hz, 1H), 5.82 (s, 2H). No ionization found.

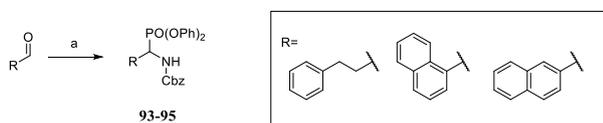


Scheme S13. Reagents and conditions. a) 2-(2-(Prop-2-ynyloxy)ethoxy)ethyl carbamate, P(OPh)_3 , Cu(OTf)_2 , DCM, rt, 16 h. b) TFA, DCM, rt, 1 h.

2-(2-(Prop-2-ynyloxy)ethoxy)ethyl carbamate (S27). Procedure and characterization consistent with previously reported data.¹

***Tert*-butyl 4-(2-(benzyloxycarbonylamino)-2-(diphenoxyphosphoryl)ethyl)piperidine-1-carboxylate (S28).** General procedure C with *tert*-butyl 4-(4-formylphenyl)piperazine-1-carboxylate (300 mg, 1.03 mmol) and 2-(2-(prop-2-ynyloxy)ethoxy)ethyl carbamate (**S27**) (193 mg, 1.03 mmol, 1 eq) to give *tert*-butyl 4-(2-(benzyloxycarbonylamino)-2-(diphenoxyphosphoryl)ethyl)piperidine-1-carboxylate (130 mg, 0.19 mmol, 18% yield). MS (ESI) m/z 595.9 $[\text{M}+\text{H}]^+$.

2-(2-(Prop-2-ynyloxy)ethoxy)ethyl (diphenoxyphosphoryl)(4-(piperazin-1-yl)phenyl)methylcarbamate 2,2,2-trifluoroacetate (75). General procedure **H** with *tert*-butyl 4-(2-(benzyloxycarbonylamino)-2-(diphenoxyphosphoryl)ethyl)piperidine-1-carboxylate (**S28**) (800 mg, 1.15 mmol) to give 2-(2-(prop-2-ynyloxy)ethoxy)ethyl (diphenoxyphosphoryl)(4-(piperazin-1-yl)phenyl)methylcarbamate 2,2,2-trifluoroacetate (340 mg, 0.57 mmol, 50% yield). ¹H NMR (400 MHz, CDCl₃) δ: 7.42 (dd, *J* = 8.5, 1.5 Hz, 2H), 7.36 – 7.28 (m, 2H), 7.26 – 7.16 (m, 3H), 7.16 – 7.05 (m, 3H), 6.96 – 6.78 (m, 4H), 6.25 (d, *J* = 9.5 Hz, 1H), 5.53 (dd, *J* = 22.5, 9.5 Hz, 1H), 4.22 (d, *J* = 2.5 Hz, 2H), 4.17 (d, *J* = 2.5 Hz, 2H), 3.76 - 3.73 (m, 2H), 3.71 – 3.60 (m, 4H), 3.40 – 3.25 (m, 4H), 3.03 (s, 2H), 2.96 – 2.92 (m, 2H), 2.42 (s, 1H). MS (ESI) *m/z* 594.8 [M+H]⁺.



Scheme S14. Reagents and conditions. a) CbzNH₂, P(OPh)₃, Cu(OTf)₂, DCM, rt, 16 h.

Benzyl (1-(diphenoxyphosphoryl)-3-phenylpropyl)carbamate (93). General procedure **C** with 3-phenylpropionaldehyde (0.495 mL, 3.73 mmol) to yield benzyl (1-(diphenoxyphosphoryl)-3-phenylpropyl)carbamate (509 mg, 1.01 mmol, 27% yield). ¹H NMR (400 MHz, CDCl₃) δ: 7.26 (s, 20H), 5.26 - 5.07 (m, 3H), 4.59 - 4.46 (m, 1H), 2.92 - 2.81 (m, 1H), 2.80 - 2.67 (m, 1H), 2.44 - 2.29 (m, 1H), 2.15 - 1.98 (m, 1H). ¹³C NMR (100 MHz, CDCl₃) δ: 156.0, 150.3, 150.1, 140.6, 136.1, 130.0, 129.9, 128.7, 128.7, 128.6, 128.5, 128.3, 126.4, 125.5, 120.7, 120.5, 67.6, 48.3 (d, C₁-P, *J* = 158.0 Hz), 32.1, 32.0. MS (ESI) *m/z* 502.3 [M+H]⁺.

Benzyl ((diphenoxyphosphoryl)(naphthalen-1-yl)methyl)carbamate (94). General procedure **C** with 1-naphthaldehyde (0.17 mL, 1.28 mmol) to yield benzyl ((diphenoxyphosphoryl)(naphthalen-1-yl)methyl)carbamate (405 mg, 0.77 mmol, 60 % yield). ¹H NMR (400 MHz, CDCl₃) δ: 8.24 (d, *J* = 8.5 Hz, 1H), 7.81 (dd, *J* = 14.0, 8.0 Hz, 3H), 7.55 - 6.89 (m, 14H), 6.60 (d, *J* = 8.0 Hz, 2H), 6.49 - 6.38 (m, 1H), 6.04 - 5.88 (m, 1H), 5.04 (m, 2H). ¹³C NMR (100 MHz, CDCl₃) δ: 155.7, 150.1, 150.0, 136.0, 134.0, 131.4, 130.8, 129.9, 129.6, 129.0, 128.7, 127.2, 126.6, 126.2, 125.3, 123.3, 120.7, 120.2, 67.7, 48.4 (d, C₁-P, *J* = 161.0 Hz). MS (ESI) *m/z* 524.2 [M+H]⁺.

Benzyl ((diphenoxyphosphoryl)(naphthalen-2-yl)methyl)carbamate (95). General procedure **C** with 2-naphthaldehyde (200 mg, 1.28 mmol), to yield benzyl ((diphenoxyphosphoryl)(naphthalen-2-

Appendix

yl)methyl)carbamate (434 mg, 0.83 mmol, 65% yield). ¹H NMR (400 MHz, CDCl₃) δ: 7.89 - 7.67 (m, 4H), 7.56 - 6.93 (m, 16H), 6.80 (m, 2H), 5.96 (m, 1H), 5.68 (m, 1H), 5.02 (m, 2H). ¹³C NMR (100 MHz, CDCl₃) δ: 156.0, 150.2, 136.0, 133.3, 131.7, 129.9, 129.8, 128.9, 128.7, 128.5, 128.3, 127.8, 126.7, 125.6, 125.5, 120.6, 120.5, 67.8, 53.1 (d, C₁-P, *J* = 157.0 Hz). MS (ESI) *m/z* 524.2 [M+H]⁺.

List of publications

α -Amino diphenyl phosphonates as inhibitors of *Escherichia coli* ClpP protease

Carlos Moreno-Cinos*, Elisa Sassetti*, Irene G.Salado, Gesa Witt, Siham Benramdane, Laura Reinhardt, Cristina Durante Cruz Jurgen Joossens, Pieter Van der Veken, Heike Brötz-Oesterhelt, Paivi Tammela, Mathias Winterhalter, Philip Gribbon, Björn Windshügel, and Koen Augustyns

*Shared first author

Journal of Medicinal Chemistry. 2019, 62, 2, 774-797

DOI: 10.1021/acs.jmedchem.8b01466

Identification and characterization of approved drugs and drug-like compounds as covalent *Escherichia coli* ClpP inhibitors

Elisa Sassetti, Cristina Durante Cruz, Paivi Tammela, Mathias Winterhalter, Koen Augustyns, Philip Gribbon, and Björn Windshügel

Frontiers in Microbiology, in revision.

Identification of a small molecule blocking the *E. coli* outer membrane protein TolC

Alessia Gilardi, Ursula Bilitewski, Mark Broenstrup, Klaas Pos, Adelia Razeto, Elisa Sassetti, Mathias Winterhalter, Philip Gribbon, and Björn Windshügel

Scientific Reports, in revision.