CHEMICAL BIOLOGY APPROACHES TO DRUG DESIGN: ACTIVATED CYSTEINE-LIKE PROTEIN LIGATION, SARS-CoV-2 DRUG REPURPOSING AND HYPOXIA ACTIVATED PRODRUG TRIGGERS

A Dissertation

by

KACI CAITLIN KRATCH

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Wenshe R. Liu
Committee Members,	Frank M. Raushel
	Jonathan Sczepanski
	Pingwei Li
Head of Department,	Simon W. North

December 2021

Major Subject: Chemistry

ABSTRACT

Chemical biology creates a unique opportunity to reexamine and alter biological processes by means of chemical manipulations. To better understand biological systems, chemical functionalities can be used to generate selective probes that can detect, alter and measure biological reactions as well as other influencing factors and mechanisms. In order to incorporate desirable chemical functionalities into biological systems, peptide/protein synthesis is used as well as various ligation techniques. These ligation methods allow for incorporation of large fluorophores, tags, reactive species, as well as altering the terminal ends of proteins that are essential for their biological breakdown, resulting in longer half-lives. In chapter two, a new ligation method is explored that improves upon expressed protein ligation. This method employs a small molecule that selectively activates the c-terminal end of a protein, making a once unreactive carbonyl more electrophilic and labile towards substitution. This chapter shows the ease in which a small peptide or a large protein can be easily altered at its c-terminus without the need of an intein to undergo protein splicing. This method opens the field of drug discovery to altering proteins at their c-terminus and adding desirable functionalities to peptide therapeutics to treat and better understand biological systems

Chapter three focuses on the repurposing of a drug to treat severe acute respiratory syndrome coronavirus 2, (SARs-CoV-2). Sometimes drug discovery is re-evaluating biologically active compounds in new conditions. The FDA is a huge hurdle to

overcome when getting drugs approved for human use and in situations where time is of the essence, sometimes an effective alternative can be re-discovered. To repurpose a drug the structure and function of the virus has to be understood. Mapping a virus' genome and establishing a crystal structure is essential to identifying possible drug candidates. Because of SARs-CoV-2's sequence similarity and quick characterization of its crystal structure, many previous compounds are identifiable as possible candidates. These candidates are chosen by being previously approved by the FDA as well as their structure and fitting in the enzymes active site. our studies on main protease of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is described. The turmoil created by SARS-CoV-2 prompted us to study main protease (M^{pro}) which is one of the key enzymes in life cycle of SARS-CoV-2. Our studies began by docking selected FDA/EMA (Food Drug Administration/European Medical Agency) drugs to active site of M^{pro}. Next, the promising drugs were purchased and their IC_{50} values were determined. Next, three most promising drugs were tested in live-virus microneutralization assay. Bepridil (an antiginal medicine) showed complete eradication of SARS-CoV-2 replication in low-micromolar concentrations showing promise for animal studies.

In chapter four, a new drug is designed for the treatment of tumor cells. This drug is designed to be a biologically activated prodrug that is triggered in response to hypoxic conditions. Tumors generate a microenvironment around themselves typified by low oxygen concentrations. These environments make tumor cells resistant to

chemotherapies and radiation due to poor blood flow/vasculature and lack of reactive oxygen species, respectively. These Hypoxia activated prodrugs (HAPs) circumvent these obstacles by targeting the environments that have previously prevented therapies. The selectivity of these compounds comes from their nitro containing triggers. These Nitro compounds are selectively reduced enzymatically by 1 electron transfers with the ability to be back-oxidized to their parent compounds in the presence of oxygen. This back-oxidation assures that the therapeutic effector will only be released in low oxygen environments and not in normoxic, healthy cells. The purpose of this study was not only to develop an effective therapeutic but to also challenge the previous rationale in designing these types of prodrug triggers. Nitro aromatic reduction is well established in the sense of electrochemical reduction potential and mechanism, but these compounds have shown to behave much differently in biological environments and their rate of reduction shows to be influenced by many other factors unrelated to the trigger's overall reduction potential. Aqueous, enzymatic and *in vivo* reduction must be considered when designing these compounds. We tested various trigger moieties in aqueous conditions, then selected the best candidates to synthesize our fluorophore analogs. Our fluorescent compounds were then subjected to chemical reduction, deuterium kinetics isotope reduction, and reduction by two different enzymes.

ACKNOWLEDGEMENTS

In the past five years at Texas A&M I have experienced a wide range of emotions but none more than gratitude. I would like to first thank my advisor Dr. Liu, who allowed me to join and work in his group and witness his creative and direct methods of solving problems, “it should be simple!” will forever echo in my mind. Next I would like to recognize my committee members; Dr. Pingwei Li, Dr. Jonathan Sczepanski and Dr. Frank Raushel, all of which I had the opportunity to take classes from and experience their enthusiasm for each of their respective fields. I appreciate their time and commitment to teaching and putting up with my excessive questions and unprompted office visits.

I would also like to thank my group members, none of which are similar or comparable to the next. I’ve enjoyed working with and getting to know each of them. There’s greatness in the Liu Lab and I hope my lab mates see it in themselves and each other the way that I do.

I’d also like to thank my past chemistry friends; Greg, Tom M, Lauren, Chris, Luis, Tom O and David, all who have graduated but were still willing to proof read my papers, listen and critique my practice talks and share their own experiences and advice in efforts to make my life easier. My Austin friends; Victoria, Melissa, Divian, and Crista who had no problem distracting me from chemistry and grad school altogether.

I'd also like to thank Brittany Rupple for keeping me sane as best she could and handling all the chaos that I can bring with incredible patience and grace.

Finally, I would like to thank my family. My Father, Tim Kratch, for his unconditional positive outlook and unshakeable belief that when you "do the right things, the right things happen". My Mother, Debbie Kratch for her pragmaticism, reminding me to always weigh the possible outcomes and to think before I act. And my Brother, Joe Kratch, whose determination and confidence has always inspired me to be better than I was yesterday.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

Faculty Committee Recognition: This work was conducted under the direction of Professor Wenshe Liu of the Department of Chemistry and supervised by a thesis committee consisting of Professors Frank Raushel and Jonathan Sczepanksi of the Department of Chemistry and Professor Pingwei Li from the Department of Biochemistry and Biophysics.

Student/Collaborator Contributions: For the work described in Chapter 2, The RNAseH expression, hydrazide ligation reaction and hydrolysis assay were conducted by Dr. Ge Yu of the Liu Lab. The cloning and expression of H2A-K129C-6H protein, as well as the H2A-H2B dimer folding assay were performed by Dr. Wei Wang of the Liu Lab. The expression vector for SUMO protease and pET28a-SUMO cloning vector were kindly provided by Dr. Pingwei Li from the Department of Biochemistry & Biophysics at Texas A&M University. The cloning of Ub-G76C-6H and Ub-C-6H, and the cell stock preparation of Ub-KxC mutants were conducted by Dr. Xiaoyan Wang in the Liu Group.

Chapter 3, the screening assays were carried out by Dr. Erol Vatansever, the crystallization conditions and the crystal structure determination of the expressed and

purified N346D, C348S, Y384F PylRS (185-454) (*o*-ClFRS) protein was done by Dr. Kai S. Yang. The docking study in Chapter 4 was done by Dr. Wenshe R. Liu. IC_{50} assays and M^{pro} expression and purification described in Chapter 3 are done by Dr. Erol Vatansever with Dr. Kai S. Yang. Crystal structure determination was done by Dr. Kai S. Yang. Synthesis of FRET-Sub2 is done by PhD candidate Lauren R. Blankenship.

Funding Sources

Graduate study was supported by a teaching assistantship from Departments of Chemistry at Texas A&M University and research assistantship from Dr. Wenshe R. Liu.

This work was also made possible in part by National Institute of Health under Grant Number R01GM127575 and R01GM121584, and Welch Foundation under Grant Number A-1715

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I INTRODUCTION: MEDICINAL CHEMISTRY AND DRUG DESIGN

1.1 Therapeutic Proteins and Peptides

Peptide therapeutics have played a notable role in medical practice since the advent of insulin therapy in the 1920s. Most peptide drugs are essentially replacement therapies that add back or supplement peptides in cases where endogenous levels are inadequate or absent. This is exemplified by the isolation and first therapeutic use of insulin in diabetics who did not produce sufficient quantities of the hormone[1]. There are three classes of therapeutic peptides; native, analogous, and heterologous. A native peptide has the same sequence as a peptide natural product and are produced synthetically or through recombinant expression. Limitations of native peptides include short half-lives and poor oral bioavailability. The short half- life of many peptide hormones is due to the presence of peptidases and natural excretory functions that work solely to inactivate and clear peptides. This lability is essential to maintaining homeostasis but is problematic for therapeutic purposes. Because of these limitations peptide analogs have been developed with improved pharmaceutical properties.

Peptide analogs are modified versions of native peptides with improved drug properties. These improved properties are obtained by backbone modification, incorporation of non-natural amino acids, and by conjugation of various moieties that improve half-life, specificity, or solubility. For example, conjugation to polyethylene glycol (PEG), lipids, and proteins such as Fc fragments has been used as a half-life extension. Conjugation

can also be used to deliver a cytotoxic payload or imaging agent to specific cell types targeted by the therapeutic peptide[1].

Heterologous peptides are peptides with desirable properties that are discovered and designed independently of the natural peptide, by means of synthetic library screening, phage display or computer assisted drug design (CADD).

Currently, the majority of peptide drugs on the market and in development are analogs that build on the intrinsic activity of native hormones with improved pharmaceutical properties. While the limitation of poor oral bioavailability is still an issue, injection is now viewed as an acceptable route of administration for certain indications, largely due to the development of longer-acting peptides resulting in fewer injections needed for therapeutic results.

1.1.1 Peptide ligation & Protein synthesis

While peptides and proteins are structurally the same, they are differentiated by their lengths. Peptides are considered to be amino acid chains made up of 2 to 50 amino acids where proteins are anything larger than 50 amino acids. In 1980 most therapeutic peptides were no larger than 10 amino acids[1], due to advances in chemical protein synthesis and peptide ligation techniques, proteins as large as 400 amino acids have been synthesized, for novelty more than utility, while most current, useful, therapeutic peptides average to the size of about 40 amino acids. Native peptides and proteins can be obtained via recombinant methods but these techniques are limited to the use of the 20 naturally occurring amino acids. Chemical protein synthesis offers an effective strategy

for accessing proteins that are difficult or impossible to obtain via gene expression systems, such as site-specific post-translationally modified proteins, mirror-image proteins for racemic crystallography or drug discovery, and proteins with unnatural elements.[2] The improvements in chemical peptide synthesis has led to the accessibility of larger peptides as well as the incorporation of desirable functional groups allowing for increased specificity, desirable functionalities and longer lasting therapeutic activity. These functionalities can be used to tune the functions and properties of therapeutic protein candidates, or in the use of proteins as tools for biological investigations through proteomic or imaging studies[2]. The ability to alter protein structure and function has great potential to enhance our understanding of proteins, generate new tools for biomedical research, and create novel therapeutic agents[3]. There have been numerous methods developed to generate altered peptides as well as ligating proteins to various moieties to either enhance functionality or to elucidate mechanisms of structure and function.

1.1.2 Solution Phase Peptide Synthesis

Initially, peptides were synthesized by means of wet chemistry. The first successful coupling of two amino acids was performed via acyl chlorides by Emil Fischer in 1903, but at that time no suitable amino-protecting group was available for synthesizing longer peptides[4]. The introduction of the benzyloxycarbonyl (Cbz) protecting group by Bergmann and Zervas [5], and other inventions, such as the development coupling reagents and the successful protection of the mercapto group of Cysteine by the benzyl

residue[6], as well as the removal of S-benzyl and tosyl groups with sodium in liquid ammonia[6] allowed for the synthesis of the nonapeptide hormone oxytocin[7], for which du Vigneaud was awarded the Nobel Prize in 1955. Unfortunately, for assembly of longer peptides or small proteins, the repetitive procedures of coupling, deprotection, isolation and purification of intermediates were found to be very tedious when carried out in solution. Moreover, solubility problems often prevented the elongation of the peptide chain. The method of Solid Phase Peptide Synthesis developed by R.B. Merrifield resolved these laborious techniques and had an enormous impact on the development of peptide synthesis for which he too won a Noble prize in 1984[8].

1.1.3 Solid Phase Peptide Synthesis

Solid phase peptide synthesis (SPPS) offers important advantages over the synthesis in solution, in that coupling reactions can be carried out more rapidly and nearly to completion using an excess of the activated amino acid which is removed at the end of the reaction by washing. Since its inception, many advances have been made to improve the efficiency of SPPS. Different resins[9] have been developed, modified by appropriate handles which enable anchoring of the protected C-terminal amino acid residue by the formation of ester or amide bonds, thus allowing the synthesis of peptide acids and peptide amides, respectively[10]. Various protecting groups have been developed to assist in desired modifications and orthogonalities of amino acid side chains, as well advances in coupling reagents to give quantitative yields and minimal by-products. As a result of these advances, many medium-sized 30–50-mer peptides can be easily synthesized by

manual or automated SPPS and even longer protein-like peptides can be synthesized by coupling protected segments through ligation methods.

1.1.4 Native Chemical Ligation (NCL)

In the very first report on Native Chemical Ligation (NCL) Kent demonstrated the total synthesis of human interleukin, a 72-amino acid protein[11]. In native chemical ligation, the thiolate of an N-terminal cysteine residue of one peptide attacks the C-terminal thioester of a second peptide to effect transthioesterification. An amide linkage forms after rapid S → N acyl transfer, shown in Fig 1. Since the inception of NCL, many other ligation technique have been developed that utilize this same S to N acyl transfer mechanism to support the synthesis and semi-synthesis of proteins.

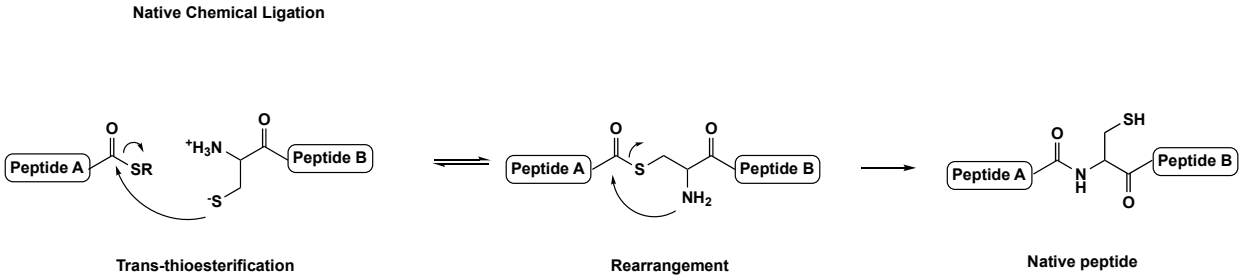


Figure 1 Native chemical Ligation

The practical utilization of NCL owes its effectiveness to the unique properties of thioesters, which are stable to hydroxide-catalyzed hydrolysis, but labile towards thiolysis and aminolysis. This controlled reactivity combination is so highly chemo-

selective that it occurs only at the N-terminal Cys residue, the thioesters formed at other Cysteine residues are reversible and unproductive. NCL is a defining point in establishing chemo-selective ligation as a general synthetic route for the chemical assembly of proteins. This approach is widely used because of the development of general access to peptide thioesters. Automation and advancements in SPPS essential for preparing the peptide fragments have strengthened the capabilities of researchers in achieving the success of NCL to an even greater extent.

1.1.5 Expressed Protein Ligation (EPL)

Native chemical ligation relies on organic peptide synthesis to obtain the necessary C-terminal thioester fragments. The use of SPPS or solution phases peptide synthesis limits the size of proteins as well as produces significant amounts of waste. To overcome these disadvantages, a new ligation method, Expressed Protein Ligation, was develop in 1998 by Muir et. al. Expressed Protein ligation (EPL) exploits the natural process of protein splicing to ligate peptide fragments. Protein splicing is a process in which a protein undergoes a series of intramolecular rearrangements and reactions that result in precise removal of an inner segment, referred to as an intein, and ligation of the two flanking protein fragments, termed exteins. Protein splicing has no sequence requirements in either of the exteins. In contrast, inteins are characterized by several conserved sequence motifs. As shown in Scheme 2, the first step in expressed protein ligation mechanism involves an N→S (or N→O) acyl shift in which the N-extein unit is transferred to the side chain SH or OH group of a Cys/Ser residue, located at the immediate N terminus of the intein.

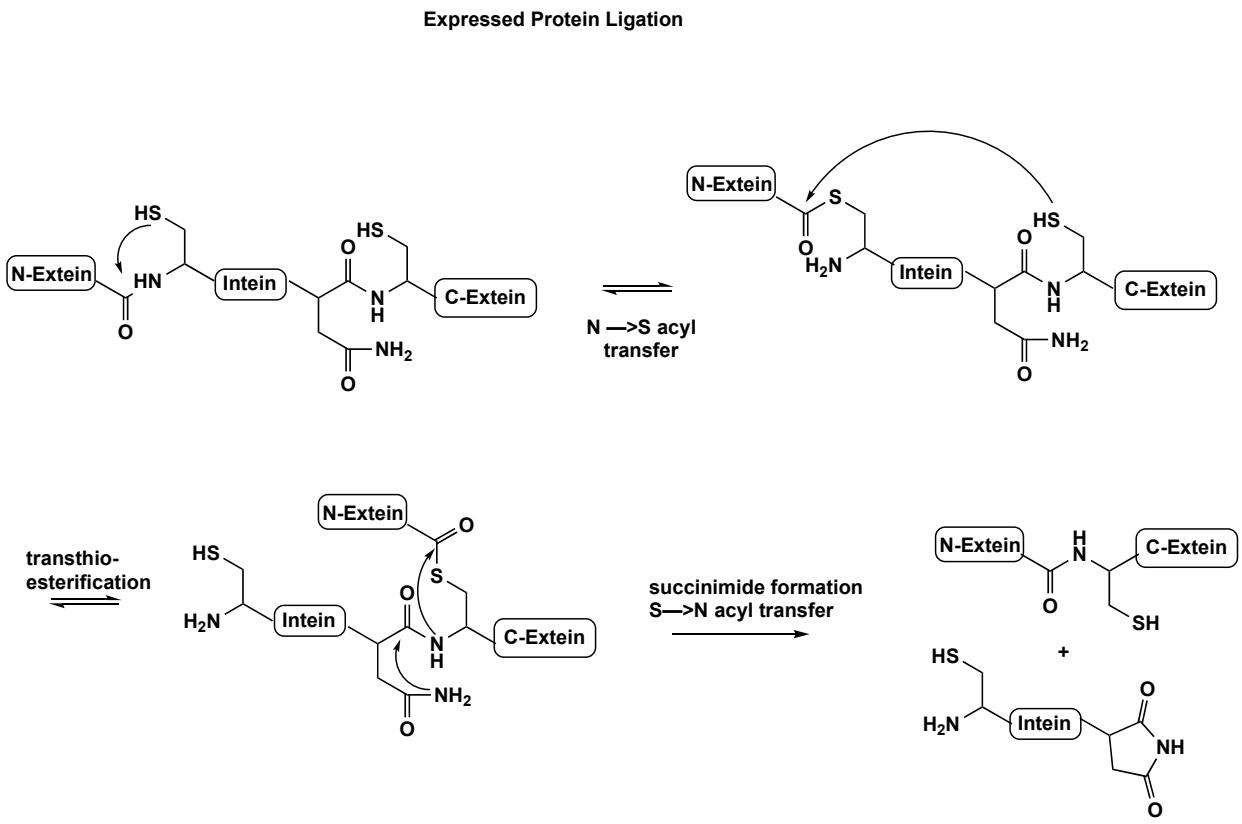


Figure 2 Expressed Protein Ligation

While this rearrangement is thermodynamically unfavorable, structural analysis indicates that the intein structure may catalyze this step in part, by twisting the scissile amide-bond into a higher energy conformation[12], pushing the equilibrium to the thioester side. In the next step in the process, the entire N-extein unit is transferred to a second conserved Cys/Ser/Thr residue at the intein-C-extein site in a transesterification step. The resulting intermediate is then resolved through a cyclization reaction involving a conserved asparagine residue at the C terminus of the intein. The intein is thus excised as

a C-terminal succinimide derivative. In the final step, an amide bond is formed between the two exteins as a result of an S→N (or O→N) acyl shift. The final step of protein splicing resembles the second step of native chemical ligation. Indeed, native chemical ligation provided a mechanistic framework for understanding the last step in the protein splicing mechanism[13]. Expressed protein ligation has achieved widespread use in protein structure-function analyses [14,15,16]. In particular, the semisynthesis of proteins that are phosphorylated[17,18], glycosylated[19-21], or labeled with isotopes[22] or fluorophores[23-25] on specific residues is leading to notable insights. From humble beginnings of dipeptide synthesis, the advances in ligation techniques of amino acids and peptides has led to the synthesis of numerous peptides and peptidomimetics that have been used as therapeutics as well as leading to revelations to the structure and function of various biological entities and processes.

1.2 SARS-CoV-2 Therapeutics

1.2.1 Life Cycle of Severe Acute Respiratory Syndrome Coronavirus 2

The pandemic caused by SARS-CoV-2 emerged in December 2019 and has surpassed 2 million global deaths as of February 2021.[26] SARS-CoV-2 is an mRNA virus. Its genome encodes for four structural proteins (SPs) named as spike (S), envelope (E), membrane (M) and nucleocapsid (N). Non-structural proteins (NSPs) are expressed as

two long polypeptides and digested by papain-like protease (PL^{pro}) and main protease (M^{pro}) to generate functional proteins.[27]

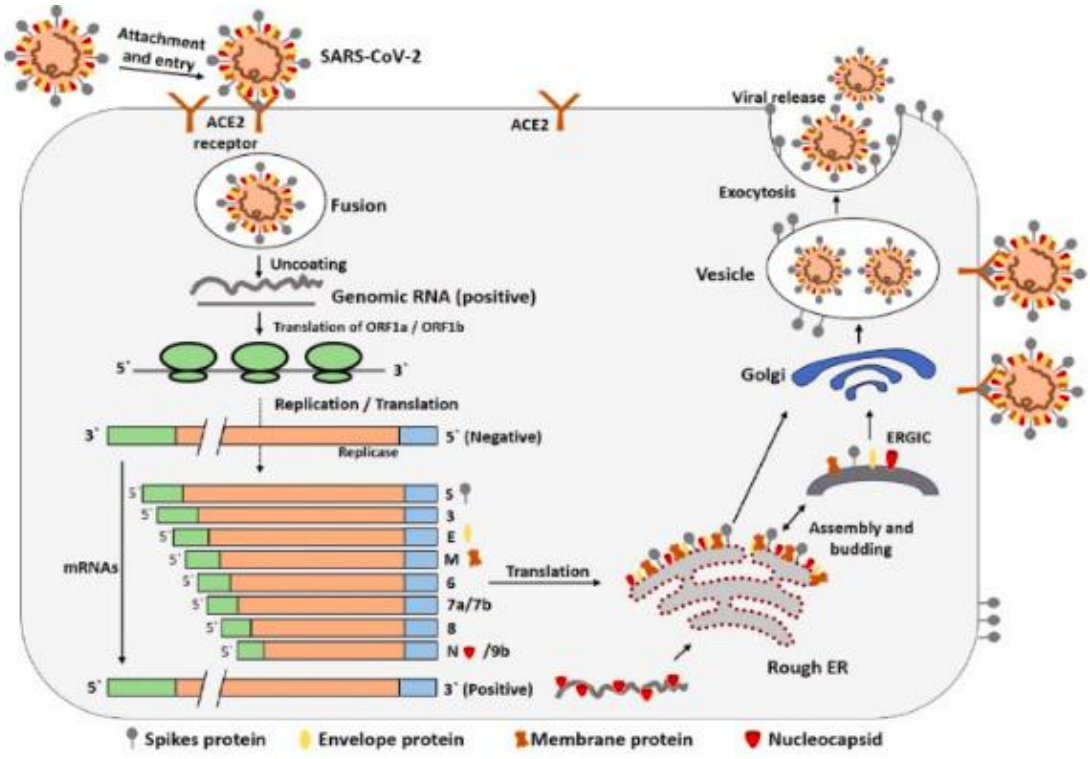


Figure 3 The life cycle of SARS-CoV-2 in host cells; begins its life cycle when Spike protein binds to the cellular receptor ACE2. After receptor binding SARS-CoV-2 releases RNA into the host cell. Genome RNA is translated into viral replicase polyproteins pp1a and 1ab, which are then cleaved into small products by viral proteinases. The polymerase produces a series of subgenomic mRNAs by discontinuous transcription and finally translated into viral proteins. Viral proteins and genome RNA

are subsequently assembled into virions in the ER and Golgi and then transported via vesicles and released out of the cell.

Shown in figure 3, the life cycle of SARS-CoV has four essential proteins that are necessary for its survival. Beginning with cell entry, angiotensin-converting enzyme 2 (ACE2) is responsible for the virus' entry into the host cell by binding with the virus' spike protein[28]. After entry, the virus is translated by the host ribosome into two large polyproteins that are processed to their functional components responsible for packing new virions[29]. There are two proteases responsible for processing these polyproteins into their functional parts; the main protease (M^{pro}) and papain like protease (PL^{pro}). The final essential protein is a RNA dependent RNA polymerase (RdRp) that is required to replicate the RNA genome. Because these four proteins, Spike, Mpro, PLpro, and RdRp, are essential for the RNA virus to survive and replicate, they are of high interest as drug targets when developing therapeutics.

1.2.2 M^{pro} as a Drug Target

M^{pro} is an essential enzyme for life cycle of SARS-CoV-2, generating functional proteins from long polypeptide chains making it necessary for the viruses reproduction. Because of this, M^{pro} is an ideal drug target. M^{pro} is a cysteine protease, it's active site contains a cysteine that is essential for its activity. This active site cysteine enables researchers to develop highly potent inhibitors. Moreover, substrate specificity of M^{pro} is different from human proteases which creates the opportunity to design very selective

M^{pro} inhibitors. To date, there are M^{pro} inhibitors which showed efficacy in eradicating SARS-CoV-2 replication in cell culture experiments and undergoing pre-clinical trials.[30]

1.2.3 Structure of M^{pro}

Through enormous effort by Chinese scientists, the first crystal structure of M^{pro} was reported within 4 months after the SARS-CoV-2 genome sequence data was released. Sequence data showed 96% similarity suggesting that the structural features, substrate preferences and inhibitors of M^{pro} is likely to be similar to SARS-CoV-M^{pro}.[29] Similar to SARS-CoV M^{pro}, M^{pro} also has a dissociation constant (K_d) 2.5 µM measured by analytical ultracentrifugation.[31] The the N-terminus residues come together by non-covalent interactions to for M^{pro} to for its' functional dimer. Notably, studies report that M^{pro} constructs having purification tags on N-terminus show significantly decreased activity.[32] M^{pro} crystal structure shows a tight dimer form (Figure 4A). M^{pro} active site consists of a catalytic dyad with Cys145 and His41 and substrate binding pockets (Figure 4B). One significant mutation in M^{pro} compared to SARS-CoV M^{pro} is T285A. Residue 285 is at the dimerization interface and mutating the residues S284, T285, I286 with alanine showed to 3.6 fold increase in activity.[33] This is presumably due to increasing the dimerization efficiency by reducing steric crowding at the dimer interface.

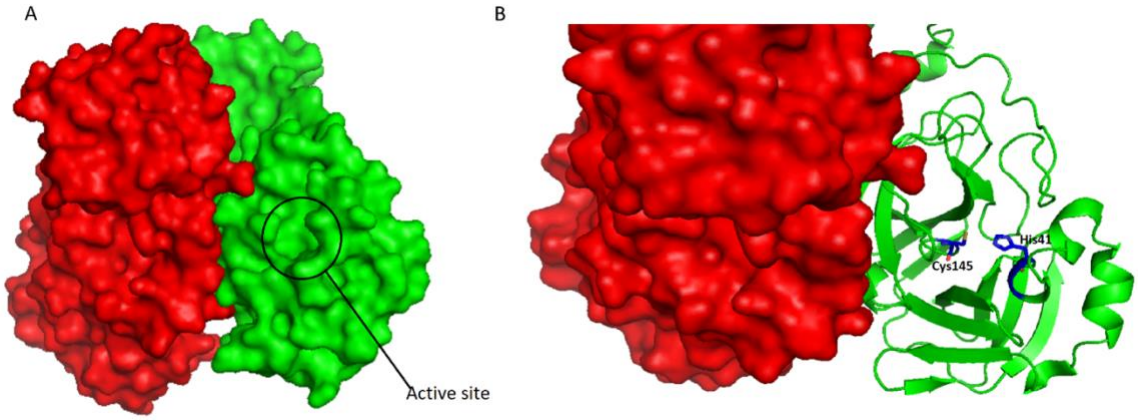


Figure 4 *crystal structure of M^{pro} (PDB ID: 6y2e) A) Dimer form of M^{pro} units are shown in red and green. B) Catalytic residues of M^{pro}*

1.2.4 Amino Acid Preference for M^{pro} Substrates

Stringent preference on residues P4, P3, P2, P1, P1`, P2` has been observed with the *in vitro* studies done with SARS-CoV M^{pro} and M^{pro}.[34, 35] The most stringent preference is observed for P1 position. Stringency decreases moving towards both N-terminus and C-terminus directions (Figure 5). Uniport orders the degree of selectivity of each position at the following order P1 > P2 > P1′ > P4 > P3.[30] Starting from the N-terminus site, P4 position favors amino acids with small hydrophobic side chains such as alanine and valine. P3 position favors amino acids such threonine, lysine and arginine. The P2 position is generally occupied by hydrophobic residues such as leucine, methionine, phenylalanine with leucine being the most preferred residue. There is a very strong preference for glutamine residue on P1 position. The only other residue that shows some activity at P1 position is histidine. P1` prefers residues with small side chains such as

serine, alanine, cysteine, glycine, threonine in the given order (serine showing the highest activity). P2` position favors small side chain amino acids like alanine, glycine, glutamic acid and threonine. SARS-CoV M^{pro} and M^{pro} have very similar sequence preferences.

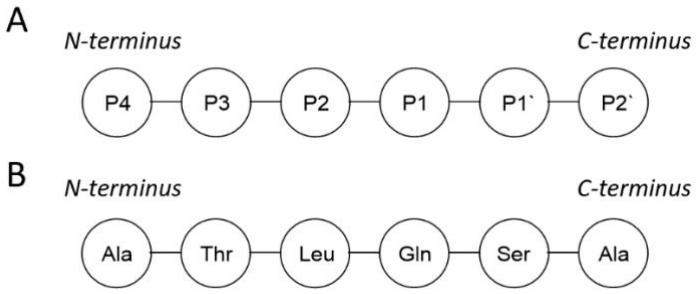


Figure 5 *A) Residue nomenclature B) Most preferred substrates for each position*

1.2.5 Substrates Used in M^{pro} Assays

Guided by the substrate amino acid preference of SARS-CoV M^{pro} , dozens of substrates were designed for measuring activity of SARS-CoV-M^{pro} and M^{pro}. The developed substrates can be grouped into three categories: Tag-less peptide substates, P1`-dye substrates and FRET-pair substrates.

Tag-less Peptide Substrates

In enzyme assays, the structural perturbation brought by tags is generally a point of concern. Therefore, assays without reporters or tags are desirable when applicable. Enzymatic reactions using tag-less peptide substrates can be tracked by high pressure liquid chromatography (HPLC). Because fluorescence-based methods include using

large dye molecules, quantification can be an issue due to quenching. Tag-less peptide substrate counterparts of fluorescent-based substrates can be used to overcome these challenges. The major drawback for using tag-less peptide substrates is the amount of time required for HPLC analysis. For example when doing inhibition analysis of a varying range of inhibitor concentrations or conducting a screening study of multiple drug candidates, high throughput screening would be more desirable. Tag-less peptide substrates are ideal when reporting molecules are either too large in size or are limited by fluorescence quenching.

P1`-dye Substrates

Fluorogenic assays that rely on turn-on fluorescence are desirable because of these compounds' inherent sensitivity. It's common to get clear signals from sub-micromolar concentrations of fluorogenic dyes. For M^{pro} assays, catalytic action selectively cleaves the amide bond between the P1 and P1` residues, this precision is used to develop substrates to characterize this cleavage. A widely used probe contains 7-amino-4-carbomylcoumarin (ACC) linked to P1 residue through an amide bond. In this case, only a very weak background fluorescence signal comes from the intact substrate due to diminished conjugation of the amine with the chromophore. However, upon hydrolysis the liberated amine group makes the chromophore highly fluorescent. However, as the substrate specificity studies demonstrate, M^{pro} is very selective at P1 and P1` residues and having a rigid bicyclic moiety like ACC at P1` position decreases the enzymatic activity significantly, necessitating using higher M^{pro} concentrations and resulting in inaccurate rates.

FRET-pair Substrates

FRET-pair substrates have two tags on both ends of the peptide that are acceptor and donor. Acceptor can serve to quench the light emitted from the donor or can be used to absorb the light coming from donor to emit at a different wavelength. FRET-pair substrates are preferred in measuring protease activities since they allow the tags to be far away from the catalytic residues. Because of their distanced proximity, (different from P1`-dye substrates) they allow the measurement of the cleavage of the native amide bond without disturbing catalytic activity. As a result, FRET-pair substrates can show enzyme activity at lower concentrations compared to P1`-dye substrate while maintaining the advantages of being a fluorescence-based assay. However, the kinetic parameters calculated with the FRET-pair substrates show large variation. Most likely due to quenching effects, accurate biochemical characterization with FRET-pairs can be troublesome.

1.2.6 Inhibitors of M^{pro}

In vitro virus inhibition assays show that M^{pro} inhibitors can inhibit SARS-CoV and SARS-CoV-2 replication with sub-micromolar concentrations. Moreover, M^{pro} has a wide active site with a cysteine residue as the key moiety for catalytic activity. These all together make M^{pro} a highly attractive target for inhibitor development. There is a vast amount of literature on inhibition of both SARS-CoV M^{pro} and M^{pro}. There are also clincal trials ongoing based on M^{pro} inhibitors that show potency in *in vitro* virus-based assays. The inhibitors developed for M^{pro} can be grouped into three categories: covalent inhibitors, non-covalent inhibitors and transition-metal based inhibitors.[36]

Covalent Inhibitors

Covalent inhibitors are divided into two categories: reversible covalent and irreversible covalent inhibitors. Reversible covalent inhibitors include inhibitors with warheads such as ketones (**6a**),[37] nitriles (**6b**),[38] aldehydes (**6c**),[39] bisulfite adducts (**6d**),[40] active esters (**6e**)[41] that can form covalent adducts in the active site in a reversible fashion (Figure 6 A-E). Irreversible inhibitor examples are chloroacetamide (**7a**),[42] epoxy ketones (**7b**),[43] Michael acceptors (**7c**),[44] phthalic hydrazide leaving group compounds[45] can be classified as irreversible covalent inhibitors (Figure 7 A-C). These inhibitor designs are guided by substrate preference of M^{pro}. They are peptide-based compounds that are mostly dipeptides, tripeptides and tetrapeptides having side chains that can bind to active site pockets such as P3, P2, P1 active site binding pockets. In addition to identity of the warhead, the side chain moieties play a significant role in determining the potency of the inhibitors. These side chains guide the inhibitor to the active site. This class of compounds (especially aldehyde containing inhibitors) have highly potent inhibitors that have IC_{50} values below 100 nM. However, there may be a trade-off between the higher potency of the covalent inhibitors and the side effects, toxicity and loss of selectivity. It has been reported that aldehyde-based inhibitors can target endogenous human proteases such as cathepsins as well. [40] Therefore, a search for non-covalent inhibitors is also a desirable direction in developing inhibitors for M^{pro}.

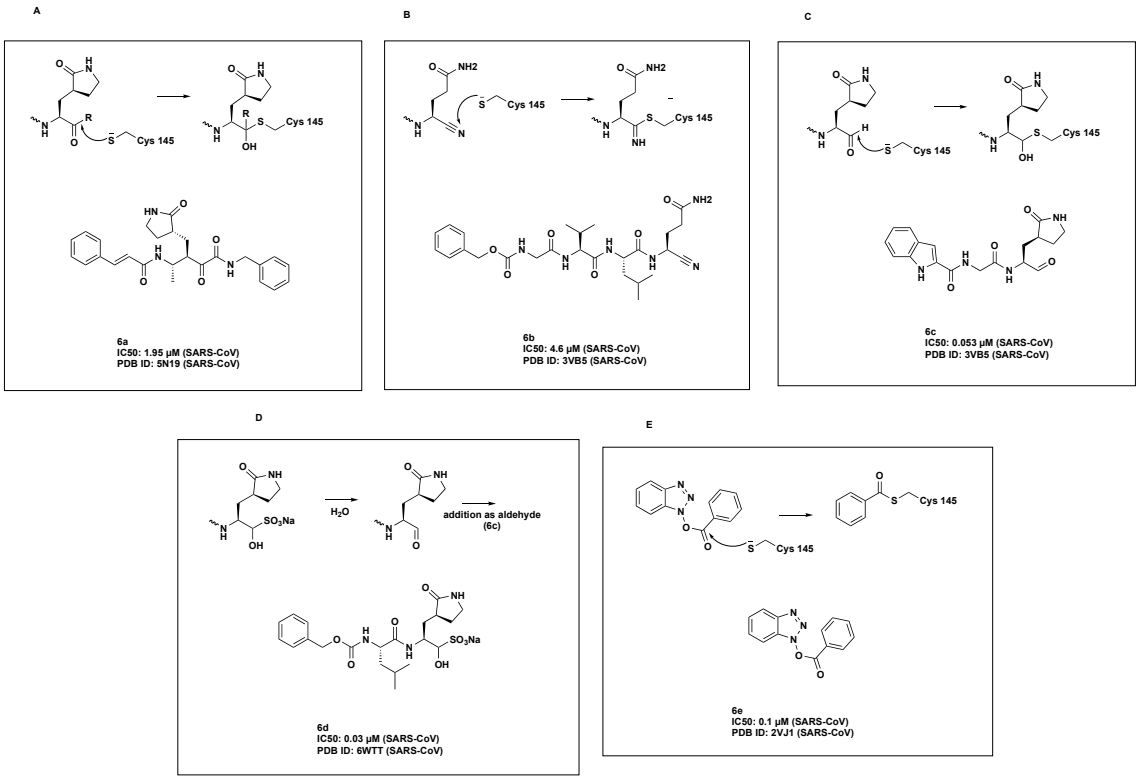


Figure 6 Examples of reversible covalent inhibitor warheads A) Ketone inhibitor B)Nitrile inhibitor C) Aldehyde inhibitor D) Bisulfite adduct inhibitor E) Active ester inhibitor

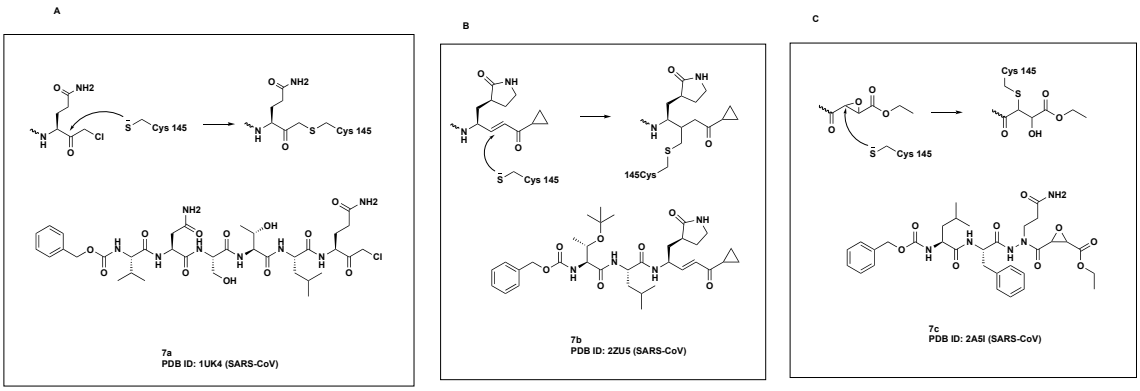


Figure 7 *Irreversible covalent inhibitor warheads A) Chloroacetamide inhibitor B) Michael acceptor inhibitor C) Epoxy ketone inhibitor*

Non-covalent Inhibitors

Non-covalent M^{pro} inhibitors are hydrophobic compounds with various different structural motifs. These compounds gain inhibition by out-competing natural substrates for the enzymes specificity, typically by filling the binding pocket of the active site. Some examples include benzatriazole (8a),[46] flavone (8b),[47] carmofur (8c),[48] isatin (8d),[49] pyridyl (8e)[50] groups. The IC_{50} or K_i (inhibition constant) values reported for non-covalent inhibitors are generally lower compared to covalent inhibitors but they may offer higher selectivity inside the cell. Moreover, the structural variety of non-covalent inhibitors is more diverse which creates more synthetic possibilities for structural optimization for improved pharmokinetics.

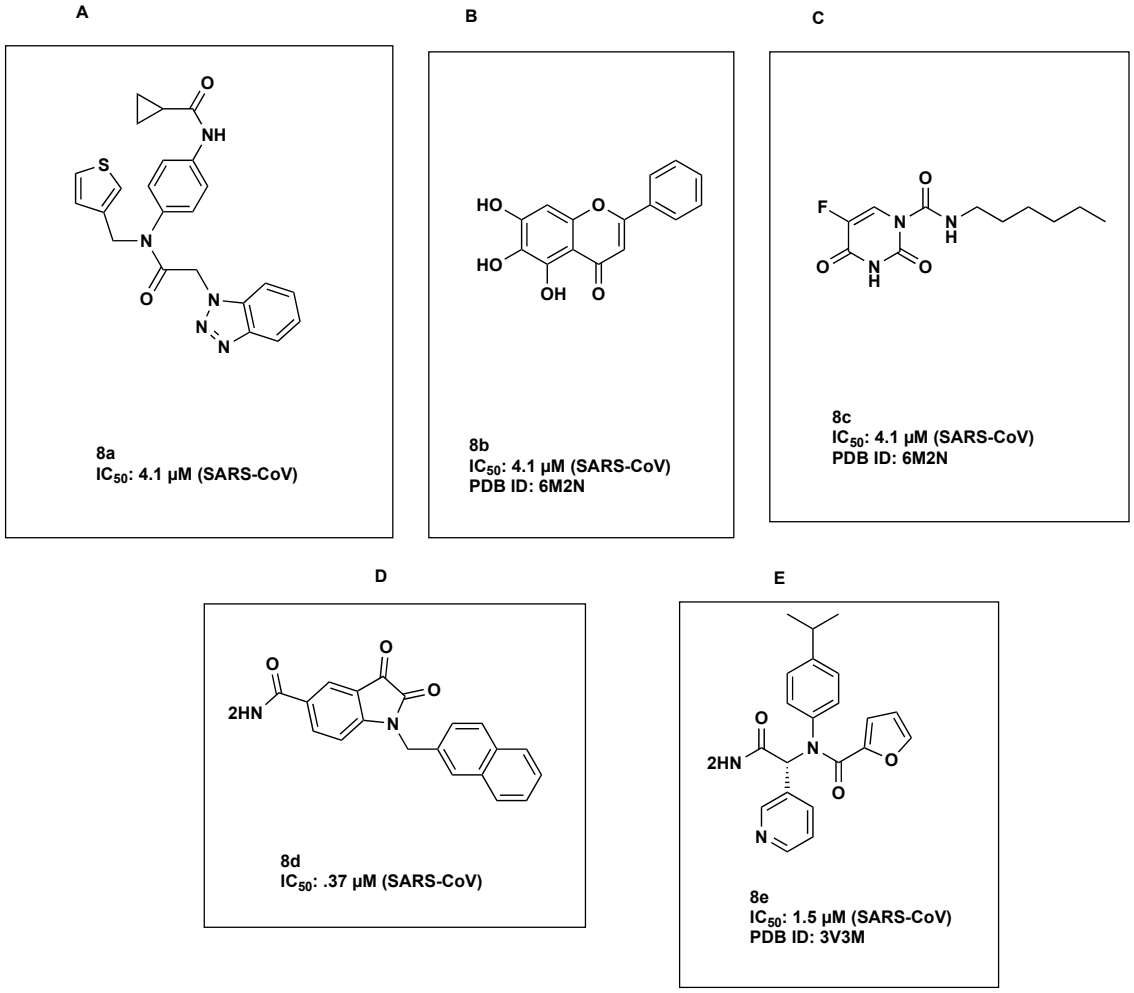


Figure 8 Examples of non-covalent M^{pro} inhibitors

Repurposed Drugs

There are many treatment options proposed and employed based on current understanding of the life cycle of SARS-CoV-2. Among some of the most prominent treatments, the FDA-approved drug Remdesivir inhibits the RNA-dependent polymerase and prevents replication of the mRNA.[51] Hydroxychloroquine is proposed to be

inhibiting endosome formation by raising the endosomal pH and through inactivating M^{pro}.[52] GC376 is a broad spectrum protease inhibitor that can inactivate M^{pro} in coronaviruses and disrupts polyprotein digestion into functional proteins for viral mRNA replication.[53] Favirapir is an antiviral that is previously known to be effective against influenza, acts as a substrate of viral RNA-dependent RNA polymerase and prevents further expansion of the RNA strand.[54] However, Remdesivir, Hydroxychloroquine, and Lopinavir-Ritonavir (previously proposed to inactivate M^{pro}) treatment options did not seem to decrease mortality rates significantly.[55] Therefore, the hunt for an effective small molecule treatment options is still ongoing.

1.3 Hypoxia Activated Prodrugs (HAPs)

1.3.1 Tumors and Hypoxic Environments

The treatment of tumor cells has proven to be difficult due to their chaotic microenvironment. Solid tumors form organ-like structures; like normal tissues, blood vessels nourish cells residing in tumors, tumor vasculature however, has abnormal organization, blood flow and function.[56] Blood flow for a tumor is heterogeneous and inconsistent throughout the solid tumor. This irregular blood flow makes it difficult for blood to carry therapeutics to and through tumors with uniform dispersion. The lack of functional lymphatic vessels inside tumors and the vascular hyperpermeability leads to increased interstitial fluid pressure in the tumor's extracellular matrix. This elevated pressure leads to increased fluid flow from the tumor to the outer tumor area and may facilitate metastasis[57]. These abnormal microcirculations lead to hostile

microenvironments typified by hypoxia and acidosis.[58,59] The oxygen concentration in normal tissue is in the range of pO_2 40−60mmHg (5−8%, 50−80 μM), while that in tumor tissue is often below pO_2 10mmHg (1.3%, 13 μM). This decrease in oxygen makes tumor cells three times more resistant to radiotherapy and chemotherapy, due to lack of catalytic oxygen and upregulated enzymes related to drug resistance, respectively.[60,61]

1.3.2 Hypoxia activated prodrugs; Structure and Mechanism of Action

While tumor hypoxia is known to promote resistance to both chemotherapy and radiotherapy and is strongly associated with poor clinical outcome, it is also recognized as a high-priority target. A group of small molecules, collectively known as hypoxia-activated prodrugs (HAPs), have been designed to treat these "difficult to reach" cells. HAP's were established over 40 years ago, but none have yet to move past clinical trials. These prodrugs' therapeutic efficacy is limited by insufficient activation in tumor micro environments[61], lack of specificity[56], and their limiting doses based on the well-known toxicity associated with nitro compounds.[62]

Toxicity is linked to the electron transfer mechanism (shown in figure 9) and its by-products. *In vivo* cycling with oxygen can give rise to oxidative stress through generation of reactive oxygen species. Nitroaromatic compounds are closely linked to carcinogenicity, hepatotoxicity, mutagenicity, and bone marrow suppression. These toxicity issues and a poor understanding of their bio-activation pathways, have impeded

success for such compounds as therapeutic agents.[63] In order for these compound to be considered effective, they need to be able to be administered in doses that are efficient and at the same time relatively safe for consumption. By better understanding these prodrugs' degradation and metabolic pathway, invaluable therapeutics for the treatment of tumor cells could be logically designed.

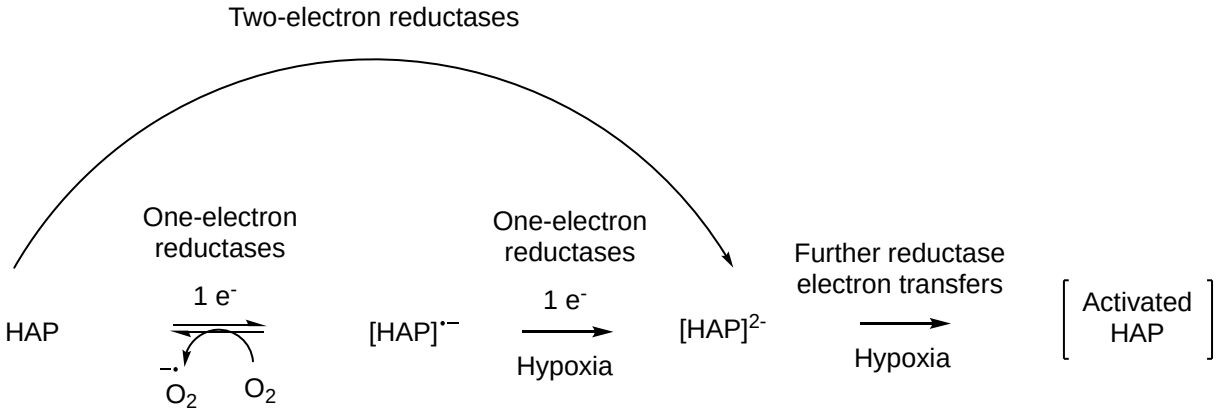


Figure 9 Enzymatic Reduction pathway

Hypoxia activated prodrugs, by design, are inactive prodrugs that require enzymatic activation to generate cytotoxic species with selectivity for hypoxic cells. The ideal HAP should possess the following properties: (1) the ability to penetrate from a blood vessel to hypoxic cells within its lifespan; (2) preferential activation in hypoxic conditions (<10mmHg) and not in normal tissues; (3) the released compound is lethal to hypoxic cells; (4) the released compound displays a 'by-standard effect'[56].

Currently, there are five different classes of HAPs: nitro (hetero)-cyclic compounds,

aromatic N-oxides, aliphatic N-oxides, quinone and transition metal complexes. Examples of recent clinical trial nitro HAPs are shown in figure 10.

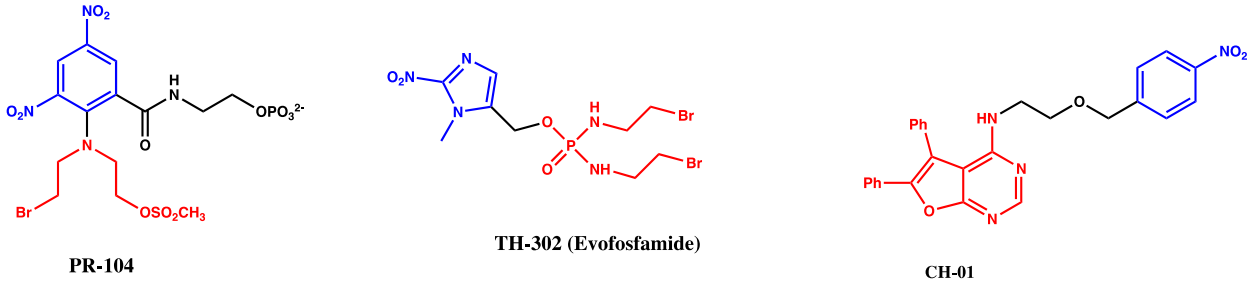


Figure 10 Recent HAPs in clinical trials; nitroaromatic (PR-104), nitroimidazole (TH-302) and nitroaromatic (CH-01).

While they vary in structure, every HAP is made of the same necessary three components: a trigger to catalyze the release of the effector, a linker to deactivate the effector, and an effector (shown in blue black and red, respectively in Figure 10). Of these three components, the trigger is the most critical. The design of the trigger determines the compound's activation rate and selectivity. Triggers must be reduced enzymatically- either by a one or two electron mechanism, most commonly carried out by Cytochrome p450 oxidoreductase(POR)[64]. POR is a membrane-bound enzyme required for electron transfer from NADPH to cytochrome p450 in the endoplasmic reticulum.

In one electron reduction, the HAP's specificity comes from the nitro functional group's ability to be re-oxidized in the presence of oxygen, reforming the parent compound. For

two electron reductions this is not possible. In a two electron reduction selectivity is largely determined by the presence of elevated levels of enzyme in tumor tissue.[64] The mechanism of reduction is shown in Figure 11.

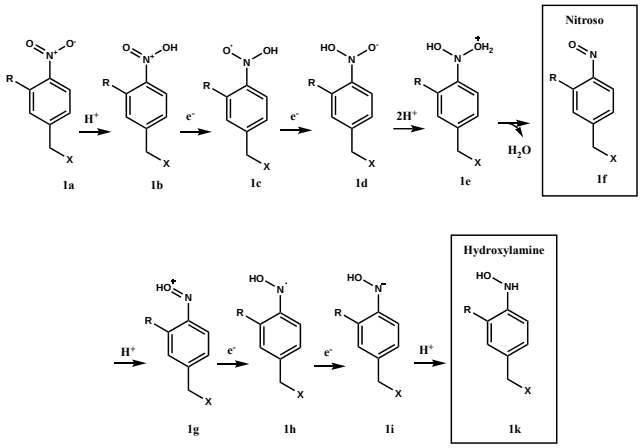


Figure 11 The one electron reduction mechanism from the starting nitroaromatic trigger to its therapeutic releasing hydroxylamine. R represents the added substituent and X represents the cytotoxic agent.

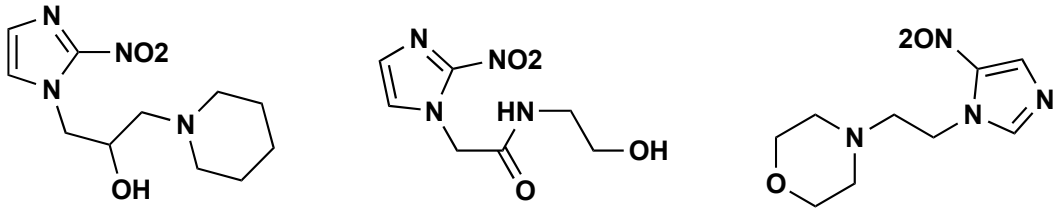
The most notable current limitations for these compounds are their dose limiting toxicities due to their metabolic by-products and their intended use to be administered with other therapeutics, as well as their short pharmacokinetic half-lives and low activation rates[56]. PR-104, shown in Figure 10, is currently regarded as “the gold standard” of HAPs due to its promising clinical I and II results and is currently under phase III clinical trials.

1.3.3 HAPs as Radiosensitizers

Normal tissue oxygen levels, often described as "physoxia," vary from 1% to 8% oxygen, with an average of approximately 5%[65]. Oxygenation in untreated tumors is significantly lower, with median oxygen levels of <2% and a range of <0.1% to 4.2% oxygen[65]. In addition, hypoxic cytotoxins are ideal to combine with ionizing radiation because they produce a profile of toxicity as a function of distance from active blood vessels that is the opposite to that produced by ionizing radiation [66]. Therefore, HAPs in combination with radiation therapy could present a unique therapeutic combination whereby the two should yield complementary killing, as hypoxia-activated drugs kill the tumor cells that are resistant to ionizing radiation. HAPs release stable and diffusible cytotoxins capable of killing the surrounding tumor cells at a higher oxygen concentration (which may not themselves be capable of prodrug activation)[67] This is known as the bystander effect and has been suggested to be an important factor in the ability of HAPs to overcome the radioresistance of hypoxic tumors[67] . Nitro-based HAPs were amongst the first to be shown to undergo the oxygen-sensitive redox cycling that is characteristic of bioreductive prodrugs.[68] Early members of this class, metronidazole and misonidazole, mimic the radiosensitization caused by oxygen in normoxic tissues, in tumor cell lines, and in preclinical animal models.[69, 70] Therefore, these compounds were thought to hold much promise in improving cancer therapy. However, clinical results were disappointing, and neither compound resulted in a statistically significant increase in survival compared with traditional radiation therapy alone.[70, 71] An important factor in their failure was thought to be the low radiosensitizing concentrations

achievable with the tolerable dose of the drugs.[70]

Since then, other HAPs have been designed as oxygen mimetic radiosensitizers, most notably, pimonidazole, etanidazole, and nimorazole.



Pimonidazole **Etanidazole** **Nimorazole**

Figure 12 Hypoxia activated prodrugs as radiosensitizers

Nimorazole specifically, significantly improved the effect of radiotherapeutic management of tumors of the head and neck, without major side effects[72]. This HAP has been carefully studied in clinical trials that measured the candidates level of tumor hypoxia and determined which candidates would best benefit from the bioreductive drug. This determination was guided by 15-gene hypoxic signature (DAHANCA 30)[73]. Nimorazole has benefited from careful study design with a large cohort and tight controls, giving greater statistical power. It is now standard of care for the treatment of HNSCC in Denmark [72]

1.3.4 Limitations

The failure to identify patients who are most likely to benefit from allocation to HAP therapy has been a huge disadvantage to these small molecules in clinical trials. Tumor

oxygenation can be directly measured using needle electrodes, and this technique was key to early work proving the association between hypoxia and treatment response[74-77]. However, this technique is limited to accessible tumors, and the availability of the equipment is limited to very few centers.

Patients can also be tested for specific hypoxia responsive genes, like CAIX, GLUT-1 and LOX[78, 79] However, an extensive review of endogenous markers of hypoxia suggested that no individual gene could be considered a definitive prognostic hypoxia marker, and instead, the use of multiple gene expressions would give more accurate and specific hypoxia information.[78] It follows that, for any meaningful patient stratification, future clinical trials must incorporate the imaging of tumor hypoxia to determine the presence, extent, and severity of hypoxia in each individual.

Tumors are inconsistent when it comes to their hypoxic profile, each circumstance should be considered unique and individually identified for their potential to be treated with these bio-reductive drugs. Assuming all tumor environments are equivalent is an over-representation of hypoxia in preclinical models and may contribute to the exaggeration of bioreductive prodrug toxicity ratios in the laboratory.

Another overwhelming challenge to HAPs is the difficulty in delivering these compounds to target cells. Anticancer drugs must penetrate relatively long distances through the extravascular compartment, which is particularly limiting for HAPs, which require these conditions for activation. TH-302 had promising clinical trials but showed substantial drug depletion in hypoxic regions due to diffusion limitations[80]. These

studies revealed that for optimum prodrug efficacy, reduction kinetics need to be balanced to accommodate competing properties of metabolic stability (for tissue penetration) and metabolism to the cytotoxic metabolite (for cytotoxicity in hypoxic cells)[81, 82].

CHAPTER II SITE-SPECIFIC FUNCTIONALIZATION AT C-TERMINI: EXPRESSED PROTEIN LIGATION WITHOUT INTEIN*[1]

2.1 Introduction

The native chemical ligation concept was first developed by Dawson et al. in 1994, in which one protein or peptide with a C-terminal thioester and the other with a N-terminal cysteine selectively undergo thiol-thioester exchange and then S-to-N acyl transfer to form a larger protein or peptide (Figure 13A).[83] Given that a protein with an N-terminal cysteine can be recombinantly produced, the development of the concept made it feasible to synthesize large proteins with a functionalized N-terminus to include either posttranslational or purely chemical modifications. To expand the synthetic scope of native chemical ligation, a related technique termed expressed protein ligation in which a recombinant fusion protein with a C-terminal intein is used to generate a protein thioester was also developed for the synthesis of a protein with a functionalized C-terminus.[3] Another notable related technique is peptide hydrazide ligation that uses nitrous acid or acetyl acetone to convert a chemically stable peptide hydrazide to a peptide acyl azide or a peptide acyl pyrazole respectively and then a peptide thioester for further native chemical or expressed protein ligation (Figure 13A).[84, 85] Evident by the original publication garnering more than 2700 citations to date, the advent of native chemical and expressed protein ligation techniques has revolutionized the protein and

peptide chemistry field. Groundbreaking applications include the synthesis of a large variety of proteins such as histones, kinases, and RAS proteins with posttranslational modifications for driving basic research advances and the production of many proteins or enzymes for therapeutic and biotechnological purposes.[86-93] Although developed extensively, further technological improvement in protein ligation is still necessary. The production of a protein thioester using the intein fusion approach is not guaranteed for many proteins. The stringent requirement for the intein catalysis to generate a protein thioester prevents the processing of many fusion proteins that are expressed insolubly and hard to fold.[94] The C-terminal residue of a targeted protein that is immediate to the intein N-terminus also significantly impacts the protein splicing efficiency, which leads to low splicing efficiency for residues such as proline at this site.[95, 96] The purification of an intein fusion also requires significant caution for avoiding premature hydrolysis.[3, 97] A split intein may be used to prevent premature hydrolysis but adds more procedural complexity.[86, 98] Using a protein ligase for expressed protein ligation resolves some issues related to the intein fusion approach but requires a specific amino acid sequence context at the ligation site. Therefore, a simple method to functionalize a recombinant protein at its C-terminus for expressed protein ligation that requires no enzymatic catalysis, can be broadly applied, and maintains high efficiency in different protein C-terminal sequence contexts is highly desired. In this work, we report such a method and its application in the synthesis of a number of proteins or peptides that can be used in both basic research and therapy.

Among 20 proteinaceous amino acids, cysteine is the most nucleophilic. Its direct activation for generating a peptide thioester was previously explored by Kajihara, Otaka, and their co-workers.[99, 100] To develop a more general, cysteine-based approach that could be applied to proteins for their C-terminal functionalization, we followed a century old industrial chemical process, leather tanning by cyanides. Cyanide salts that reduce disulfide bonds in proteins were used in the early 20th century to treat animal hides and wools.[101] During the process, a cyanide covalently attached to a protein cysteine to form a thiocyanate that underwent reversible intramolecular addition with the cysteine N-amide to generate a 1-acyl-2-iminothiazolidine intermediate. The amide bond in this intermediate was significantly weakened in comparison to a regular protein amide and therefore slowly hydrolyzed to split the protein (Figure 13B).[102]

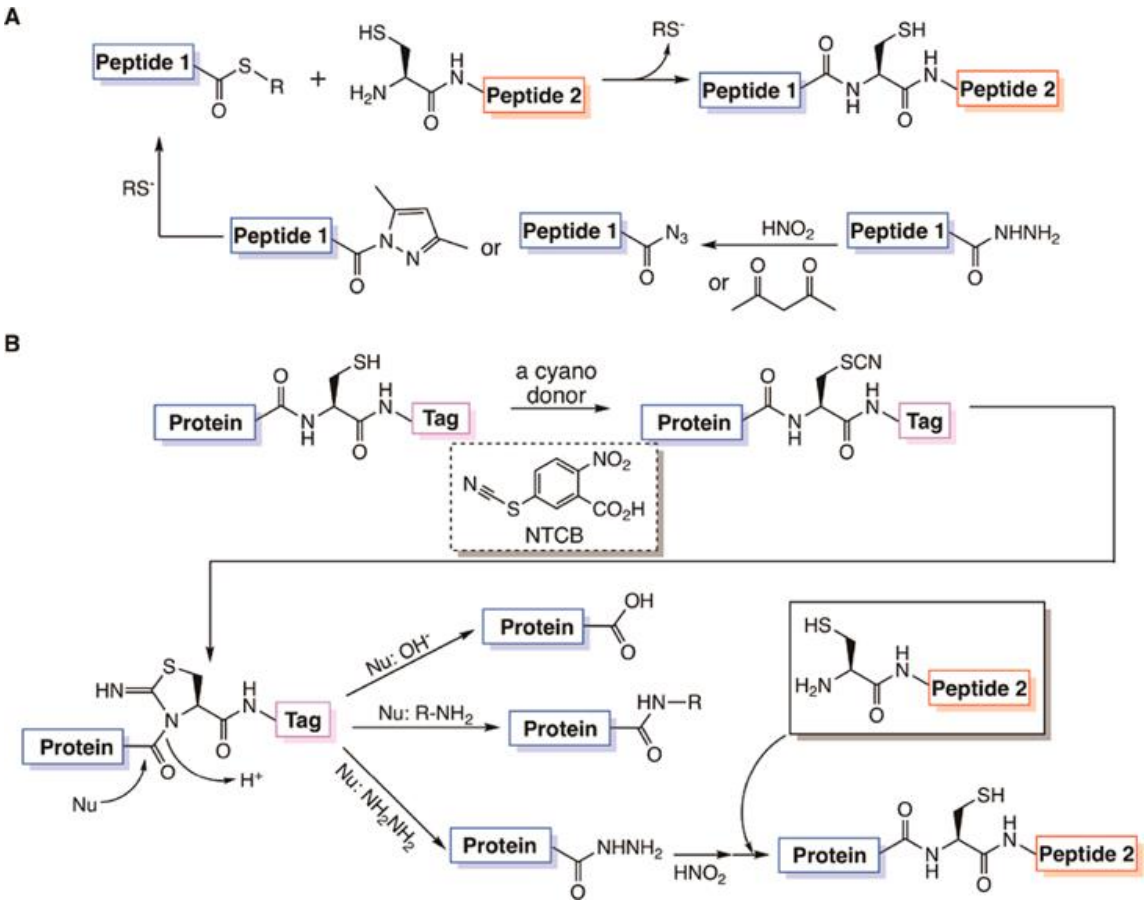


Figure 13 Protein synthesis by ligation techniques. (A) Native chemical ligation and a derivative technique, peptide hydrazide ligation. (B) A proposed ACPL technique based on nucleophilic acyl substitution of an activated cysteine residue in a recombinant protein by a nucleophilic amine. Without a nucleophilic amine, the protein undergoes hydrolysis. When the nucleophile is hydrazine, the afforded protein hydrazide can then undergo peptide hydrazide ligation to form a larger protein.

Early protein chemists exploited this reaction for mapping protein sequences and replaced cyanide salts with other cyanylating reagents such as 2-nitro-5-thiocyanatobenzoic acid (NTCB) that transfers the cyano group directly to a reduced

protein cysteine for avoiding the formation of highly toxic cyanide wastes.[103, 104] According to this reaction mechanism, providing a strongly nucleophilic amine in the reaction will trigger nucleophilic acyl substitution with the 1-acyl-2-iminothiazolidine intermediate to replace 2-iminothiazolidine and potentially circumvent the hydrolysis process. The afforded small molecule amine-ligated product can be used for a further protein ligation process. Since this proposed expressed protein ligation that we termed as activated cysteine-based protein ligation (ACPL) does not involve an enzyme and is purely chemically based, it can be highly controllable, selective, and versatile such as functioning for proteins both soluble and insoluble.

2.2 Results and discussion

2.2.1 Feasibility of ACPL

To demonstrate the feasibility of ACPL, we synthesized Boc-Xxx-Cys-OMe dipeptides in which the Xxx identity varied between seven native amino acids including proline and carried out their reactions with an equivalent amount of NTCB and then ligation with propargylamine (Pa) in dichloromethane (DCM). Our result showed that all dipeptides reacted with Pa to form desired products with varied yields (Table 1). Since the reactions were performed in DCM, amino acids with a hydrophobic and large side chain tended to have high yields, while amino acids with a relatively small and hydrophilic side chain resulted in relatively low yields. To further characterize the effect of the reaction on the chirality of the amino acid at the N-terminal side of cysteine, we synthesized the Boc-L-isoleucine-Pa using ACPL and compared it with both Boc-L-isoleucine-Pa and Boc-DL-

isoleucine-Pa that were synthesized using the standard amidation approach as controls. As shown in the ^{1}H NMR spectra (Figure 14), no diastereomeric CH peak was observed for the ligated product indicating that this reaction did not lead to the racemization of the amino acid before cysteine.

Table 1 Yields of dipeptide ligation with propargylamine

Peptide	Amount (g)	Ligation products	Amount (g)	Yield (%)
Boc-Gly-Cys-OMe	0.6	Boc-Gly-Pa	0.12	56.6
Boc-Ala-Cys-OMe	0.6	Boc-Ala-Pa	0.11	48.6
Boc-Leu-Cys-OMe	2.0	Boc-Leu-Pa	0.7	87.5
Boc-Phe-Cys-OMe	1.8	Boc-Phe-Pa	0.65	88.3
Boc-Trp-Cys-OMe	2.5	Boc-Trp-Pa	0.8	78.4
Boc-Asp-Cys-OMe	0.9	Boc-Asp-Pa	0.05	13.8
Boc-Pro-Cys-OMe	2.0	Boc-Pro-Pa	0.16	25.4

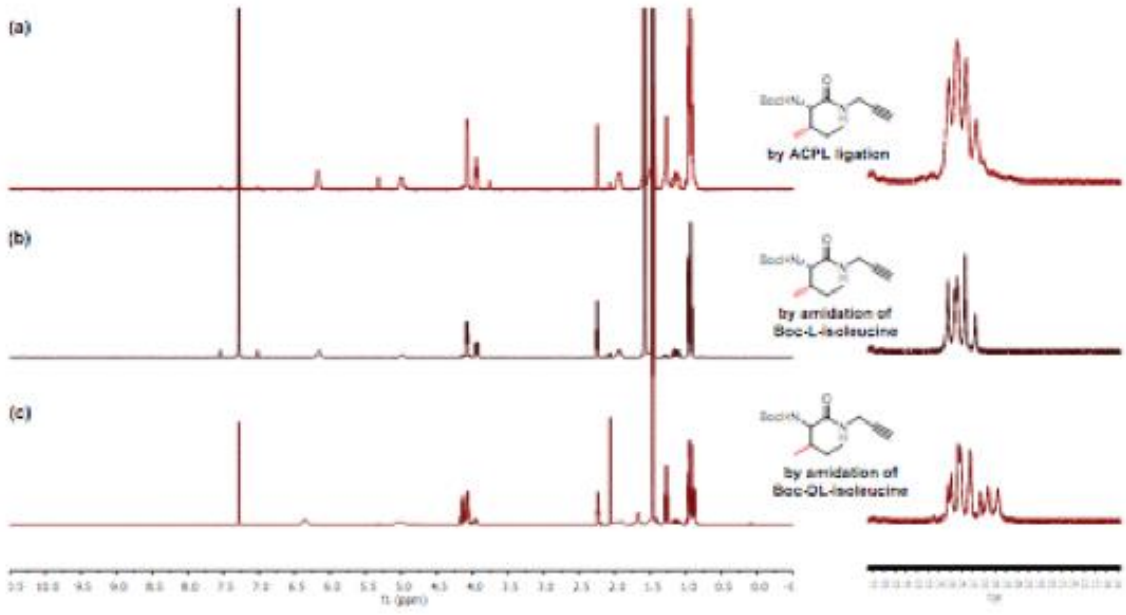


Figure 14 [1] HNMR spectra of tert-butyl (3-methyl-1-oxo-1-(prop-2-yn-1-ylamino)pentan-2-yl) carbamate from ACPL ligation (a), amidation of Boc-L-isoleucine (b), and amidation of Boc-DL-isoleucine (c) ($CDCl_3$, 400 MHz).

2.2.2 Versality of ACPL

Encouraged by our small molecule results, we tested ACPL further with recombinant proteins. Ubiquitin (Ub) is natively devoid of cysteine.[105] We chose it as a model protein for our demonstration. We produced recombinant native Ub and Ub with both a G76C mutation and a C-terminal 6×His tag (Ub-G76C-6H) in E. coli and purified them to homogeneity. We then ligated Ub-G76C-6H with 12 small molecule amines including Pa, allylamine (Aa), hydrazine (Ha), and L- and D-amino acids (Figure 11A) by adding 5 mM NTCB and a 50−1000 mM amine simultaneously to a 2 mg/mL Ub-G76C-6H solution at pH 9 for an overnight incubation at 37 °C. We carefully selected seven L-amino acids for our reactions to represent amino acids in different chemical categories and also different sizes. For all tested compounds including proline that has a secondary amine and two D-amino acids, we obtained ligation products with 50−90% yields that were estimated by the SDS-PAGE analysis of reaction mixtures (Figure 16, Table 2). After using Ni^{2+} charged resins to simply remove unreacted intermediates, we analyzed all 12 ligation products and the two original Ub and Ub-G76C-6H proteins by electrospray ionization mass spectrometry (ESI-MS) analysis. For all analyzed proteins, their deconvoluted ESI-MS spectra displayed clearly observable monoisotopic peaks. Since there is no commercial software for calculating protein monoisotopic peaks, we wrote a Python script to calculate all theoretical monoisotopic masses and their relative

intensities for all proteins and compared them to the determined ESI-MS spectra. Our results showed that determined monoisotopic masses for all proteins agreed very well with their theoretic values in terms of both molecular weight and intensity. Hydrolysis products were either nondetectable or at very low levels. To simplify the comparison, we wrote another Python script to integrate deconvoluted monoisotopic peaks and then calculate the average molecular weights and intensities for all detected protein species in a particular spectrum. The final results are presented in Figures 15B and 15C. For all determined average molecular weights, they matched their theoretical values with a deviation of ±0.3 Da (Table 2). For all 12 ligation products, we detected very few minor peaks in their ESI-MS spectra indicating that all reactions were very selective. One ligation product Ub-G76G is native Ub itself. Its ESI-MS spectrum in Figure 15C matched that of recombinantly expressed native Ub in Figure 15B. So far, our data demonstrated that ACPL works exactly according to what we proposed on a recombinant protein and this reaction is effective for amines that are primary, secondary, Ha, and amino acids with different configurations, characteristics, and sizes. The ligation with Ha was done in both native and denatured conditions. The results from two conditions showed minimal differences. Since conditions used for ACPL so far were at 37 °C with a long incubation time that might be a concern for some proteins, we also carried out reactions under lower temperatures and shorter times. We performed the reaction between Ub-G76C-6H and Pa under three different temperatures (8, 23, and 37 °C) as well as four different reaction times (4, 8, 12, and 16 h). The SDS-PAGE analysis of the final reaction mixtures exhibited no clear differences among all conditions. All

these conditions yielded the desired product as 50%. Therefore, a reaction at 8 °C for 4 h is sufficient for ACPL.

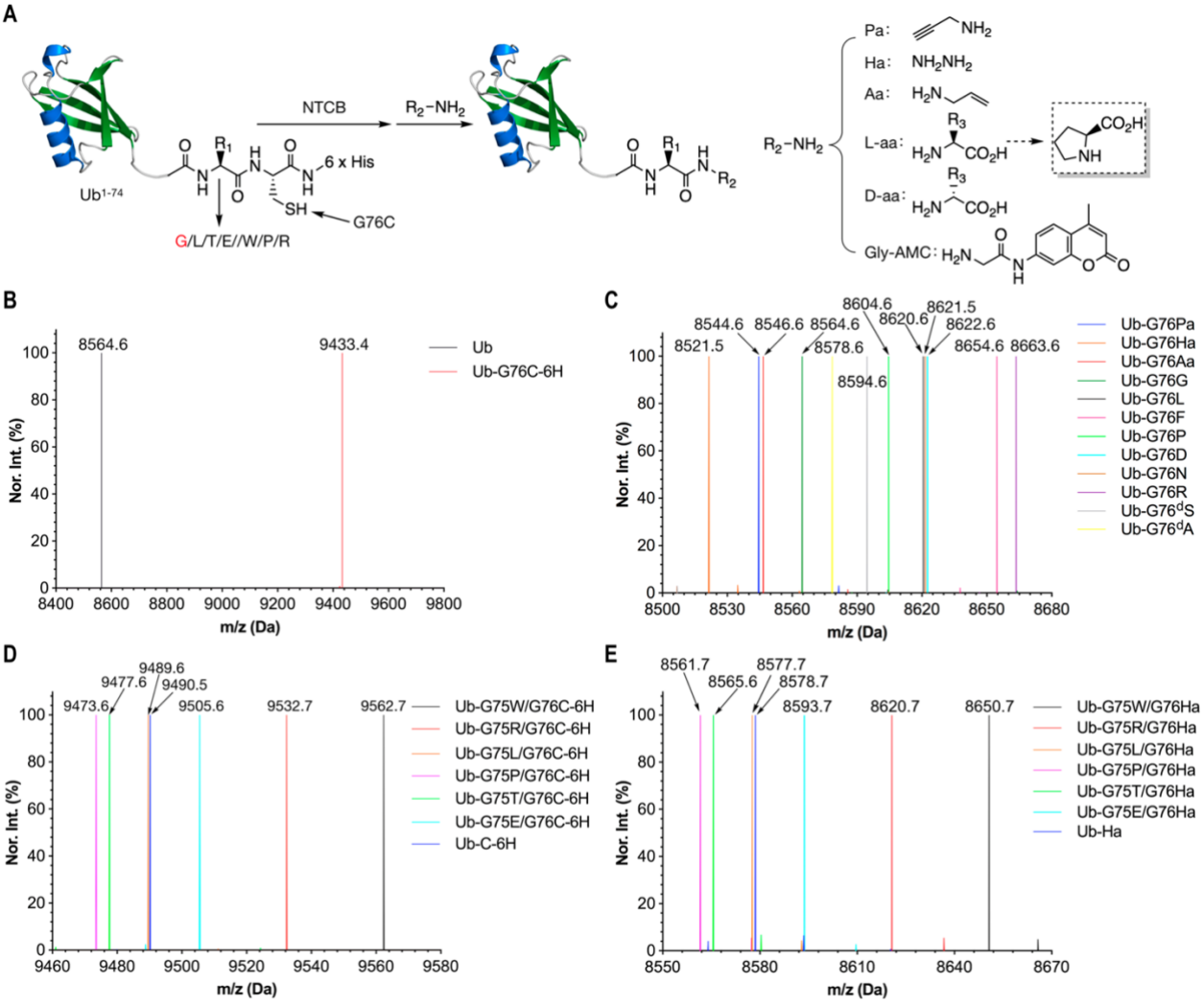
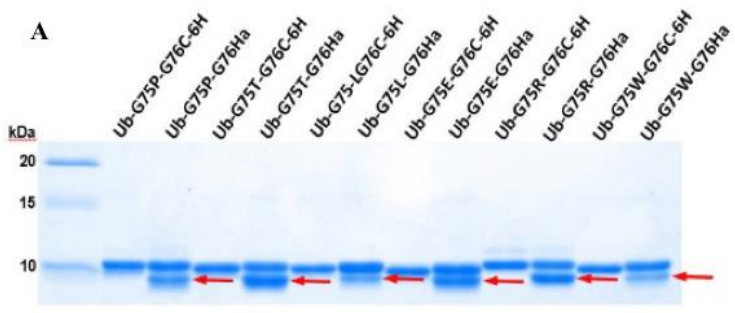


Figure 15 Synthesis of Ub conjugates by ACPL. (A) A schematic diagram to show the activation of recombinant Ub proteins containing a cysteine by NTCB followed by nucleophilic acyl substitution with amines, both primary and secondary, to generate different Ub conjugates. The native Ub has 76 residues and glycine at the 75th and 76th positions. (B) The deconvoluted and integrated ESI-MS spectra of wild type Ub and Ub-

G76C- 6H. 6H represents a 6×His tag. (C) The deconvoluted and integrated ESI-MS spectra of Ub conjugates that were converted from Ub-G76C-6H and had different ligated molecules at the G76 position. Pa, Ha, and Aa are three small molecule amines shown in A. All other ligated molecules are amino acids whose one letter codes are used for labeling. All amino acids are in the L-configuration except two D-amino acids with a footnote d. (D−E) The deconvoluted and integrated ESI-MS spectra of 7 recombinant Ub proteins and products of their reactions with NTCB and Ha. C in Ub-C-6H represents cysteine. All detected molecular weights agreed well with theoretic values in a deviation range of ±0.3 Da.

Table 2 Theoretical and determined average molecular weights of Ub, Ub-G76-6H, and the ligation products of Ub-G76C-6H with 12 amine containing compounds

Protein	Observed Mass (Da)	Theoretical Mass (Da)	Yield (%)
Ub	8564.6	8564.8	N/A
Ub-G76C-6H	9433.4	9433.7	N/A
Ub-G76G	8564.6	8564.8	70
Ub-G76P	8604.6	8604.8	50
Ub-G76D	8622.6	8622.8	70
Ub-G76N	8621.5	8621.8	80
Ub-G76L	8620.6	8620.9	75
Ub-G76R	8663.6	8663.9	85
Ub-G76F	8654.6	8654.9	65
Ub-G76^{d}A	8578.6	8578.8	85
Ub-G76^{d}S	8594.6	8594.8	90
Ub-G76Aa	8546.6	8546.8	60
Ub-G76Pa	8544.6	8544.8	55
Ub-G76Ha	8521.5	8521.7	90
Ub-G76Ha (native)	8521.6	8521.7	85



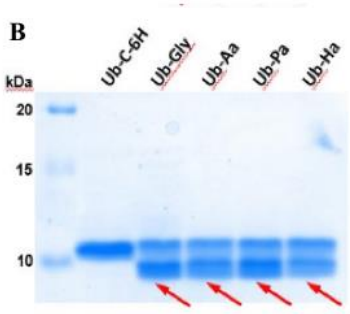


Figure 16 (A) SDS-PAGE of Ub-C-6H reactions with 4 amine containing compounds (B) SDS-PAGE of Ub-G75X/G76C-6H with hydrazine (X denotes P, T, L, E, R or W).

Ubiquitin natively has a G75 residue that has the lowest steric hindrance among all amino acids. In Ub-G76C-6H, the glycine immediately *N*-terminal to G76C might have permitted easy processing of the ligation. Other residues that have different chemical properties and/or are sterically hindered might impede the ligation. To resolve this concern, we mutated G75 in Ub-G76C-6H to six other residues that are large in size, charged, and/or having a secondary amine, recombinantly expressed them, analyzed them with ESI-MS (Figure 15D), and then reacted them in a one-pot fashion with NTCB and Ha. We chose Ha in our demonstration since its ligation products are protein hydrazides that can undergo further peptide hydrazide ligation for making even larger proteins. All reactions progressed exceedingly well and their reaction products displayed average molecular weights matching well to their theoretic values (Figures 15E; Table 3), demonstrating that the residue immediate to the *N*-terminal to the targeted cysteine has little detrimental effect on the ligation process. However, large amino acids such as tryptophan did lead to lower yields in comparison to small amino acids such as threonine (Figure 15A, Table 3). Putting a cysteine residue right after Ub G76 led to similar ligation results with Ha, Aa, Pa and glycine with an at least 50 % conversion rate (Figures 15D, 15E, 16B and Table 3). Ub has a flexible *C*-terminus that may facilitate the ligation. To show that the ligation may work in a more structurally constrained environment, we introduced a cysteine mutation at K48 and K63, two residues in the globular region of Ub and used the two afforded Ub mutants (Table 3) to undergo ACPL

with Ha. ESI-MS of reaction mixtures showed successful formation of two desired protein hydrazides (Table 3) indicating that ACPL works well in a structurally constrained protein region. Ligation both in a structurally constrained protein region and under a denatured condition is something that the traditional intein and ligase-based methods cannot perform well. Collectively our data strongly demonstrates the versatility of the ACPL technique.

Table 3 Theoretical and determined molecular weights of six Ub-G75X/G76C-6H proteins, Ub-C-6H, and their ligation products with hydrazine (data for the ligation products of Ub-C-6H with glycine and for the ligation products of Ub-K48C and Ub-K63C with hydrazine are also included). X denotes one of the six amino acids: W, R, L, P, T and *E*.

Protein	Observed Mass (Da)	Theoretical Mass (Da)	Yield (%)
Ub-G75P-G76C-6H	9473.6	9473.8	N/A
Ub-G75T-G76C-6H	9477.6	9477.7	N/A
Ub-G75L-G76C-6H	9489.6	9489.8	N/A
Ub-G75E-G76C-6H	9505.6	9505.8	N/A
Ub-G75R-G76C-6H	9532.7	9532.8	N/A
Ub-G75W-G76C-6H	9562.7	9562.8	N/A
Ub-G75P-G76Ha	8561.7	8561.8	30
Ub-G75T-G76Ha	8565.6	8565.8	70
Ub-G75L-G76Ha	8577.7	8577.8	20
Ub-G75E-G76Ha	8593.7	8593.8	45
Ub-G75R-G76Ha	8620.7	8620.9	55
Ub-G75W-G76Ha	8650.7	8650.9	20
Ub-C-6H	9490.5	9490.7	N/A
Ub-Ha	8578.7	8578.8	50
Ub-Aa	8603.8	8603.8	60
Ub-Pa	8601.7	8601.8	50
Ub-G	8621.7	8621.8	70
Ub-K48C	8539.5	8539.7	N/A
Ub-K63C	8539.5	8539.7	N/A
Ub_{1-47}-Ha	5255.8	5256.1	
Ub_{1-62}-Ha	7017.8	7017.9	

Use of ACPL To Synthesize Ub and Ub-Like Protein (Ubl) Probes

In eukaryotic cells, Ub and Ubls can be posttranslationally attached to proteins for their functional regulation.[106-108]It has been shown that replacing the C- terminal glycine in Ub and Ubl proteins including SUMO1−3, NEED8, and ISG15 with Pa using either the intein based approach or total synthesis afforded excellent probes to conjugate covalently to deubiquitinases (DUBs) or ubiquitin- like proteases (ULPs) that catalytically remove Ub or Ubls from their conjugated proteins in cells.[109-113] To recapitulate these results and demonstrate the broad application scope of our ACPL technique in the probe synthesis, we recombinantly expressed Ub and a number of Ubl proteins including SUMO1−4, NEDD8, ISG15, GABARAP, GABARAPL2, UFM1, URM1, and MNSFβ (FLAG-Ub/Ubl-GxC-6H: x denotes the terminal glycine position) that all contained a C- terminal Gly-to-Cys mutation and were also fused with a N-terminal FLAG tag and a C-terminal 6×His tag, purified them to homogeneity, and then carried out their reactions with Pa in the presence of NTCB to afford their Pa-conjugated products. ISG15, SUMO1−4, and MNSFβ natively contain a cysteine residue. This cysteine was mutated to alanine or serine to avoid nontargeted reaction at its location. The yields of all 12 reactions varied between 25% and 80% as indicated by their SDS-PAGE analysis. ESI-MS analysis of all 12 products showed their successful and efficient synthesis (Figure 17A). The final results are presented in Figure 17A that displayed very little side products for all 12 Pa-conjugated products. In comparison to both the intein based and total synthesis approach, our method for the synthesis of these Pa conjugates is much simpler and easier to control. To reproduce some literature results, we used our

synthesized Pa-conjugated FLAG-Ub (FLAG-Ub-G76Pa) to react with seven DUBs and observed efficient covalent adduct formation for all tested enzymes in both SDS-PAGE analysis and Western blotting (Figure 17B). We also performed similar tests for seven Pa-conjugated FLAG-Ubl probes and observed their covalent binding to a number of ULPs as shown in Figure 17C. Some ULPs such as SENP1 have only been vaguely confirmed in previous work to deconjugate corresponding Ubls such as SUMO4.[114] All synthesized FLAG- Ub/Ubl conjugates, of which six are synthesized for the first time, are activity-based probes that can be potentially used to profile DUB and ULP proteomes in different tissues or cells. As a demonstration, we incubated the HEK293T cell lysate with FLAG-Ub-G76Pa and then probed the FLAG-Ub- conjugated proteins by an anti-FLAG antibody in Western blotting. The result showed the formation of a number of higher molecular weight species compared to the original FLAG-Ub-G76Pa, indicating conjugation with other proteins in the cell lysate. However, a control reaction using FLAG-Ub-G76C-6H showed no covalent conjugation with any other proteins in the HEK293T cell lysate.

Ub and Ubls conjugated directly to 7-amido-4-methylcoumarin (AMC) at their C-terminus are useful fluorogenic substrates of DUBs and ULPs.[115] To demonstrate the synthesis of Ub/Ubl-AMC conjugates using our ACPL technique, we made Ub-AMC and FLAG-SUMO1−3-AMC by reacting recombinantly produced Ub-G76C-6H and FLAG-SUMO1− 3-GxC-6H proteins with Gly-AMC in the presence of NTCB. The ESI-MS analysis of all four products confirmed their successful formation (Figures 17A) and the following activity assays showed that they served as active substrates for cysteine

proteases UCHL1 and SENP1, respectively (Figure 17D). Overall, our combined data of Ub/ Ubl probe synthesis establish the broad application scope of the ACPL technique and this technique can make Ub/Ubl probes readily available in a manner that can be performed in almost any biology lab for advancing Ub and Ubl biology studies.

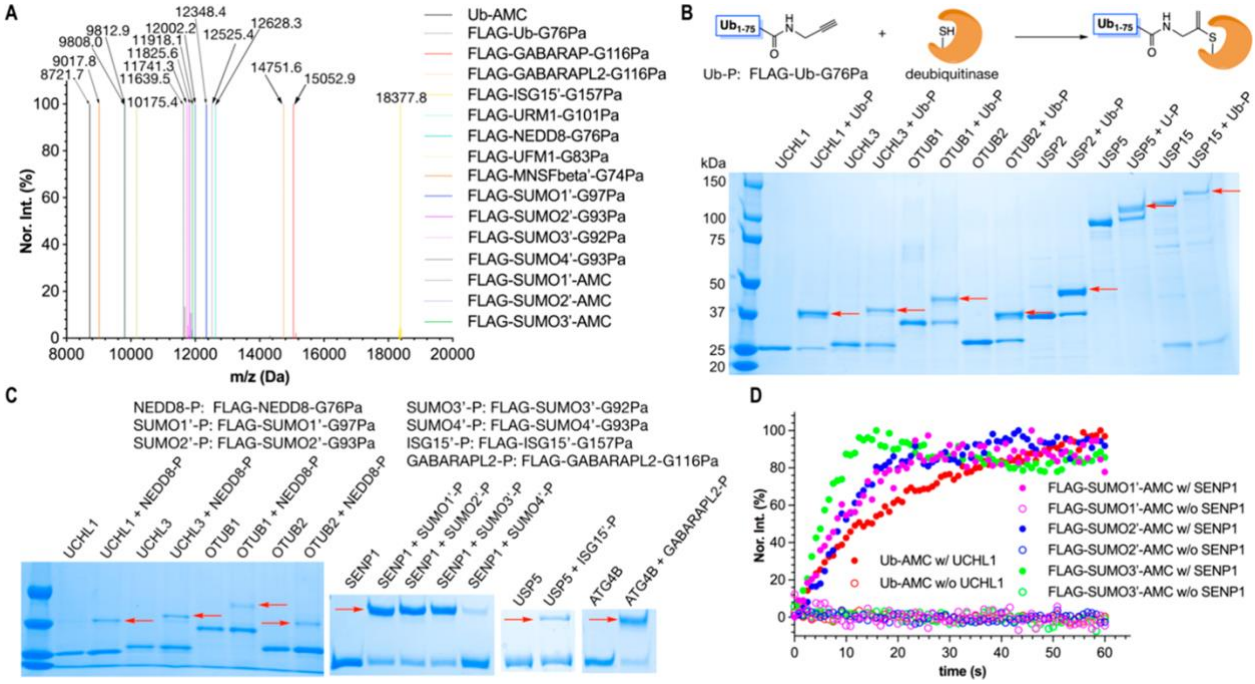


Figure 17 Synthesis of FLAG-Ub/Ubl-Pa and Ub/FLAG-SUMO-1-3-AMC probes by ACPL and their applictions in covalentconjugation or activity assays of DUB/ULPs. (A) The deconvoluted and intergrated ESI-MS of FLAG-UB/UBL-Pa and Ub/FLAG-SUMO01030AMC probes. Ub- AMC was synthesized from Ub-G76C-6H. All other Pa- and AMC-conjugated Ub/Ubls were generated from FLAG-tagged proteins. Ub/Ubls with their C-terminal glycine mutated to cysteine were expressed and purified as a protein fused with a N-terminal FLAG tag and a C-terminal 6×His tag. ISG15, SUMO1−4, and MNSF have a native cysteine residue. This cysteine was mutated to

alanine or serine in all six expressed proteins for avoiding side reactions. The label "′" indicates this mutation. All detected molecular weights agreed well with their theoretic values with a deviation range of 0.5 Da. (B) The formation of covalent adducts between FLAG-Ub-G76Pa and a number of DUBs. Red arrows point to the generated adducts. (C) The formation of covalent adducts, indicated by red arrows, between different FLAG-Ubl-GxPa probes and DUB/ULPs. (D) The DUB/ULP-catalyzed AMC release from Ub-AMC and three FLAG-SUMO-AMC conjugates.

Use of ACPL To Synthesize H2AK129ac.

To further demonstrate its broad application scope, we also carried out ACPL to make nonubiquitin or nonubiquitin-like proteins. In human cells, histone H2A can undergo posttranslational acetylation at its terminal lysine K129.[116] The functional investigation of this acetylation such as how it influences the structure and dynamics of the nucleosome will require the synthesis of the corresponding acetyl-histone, H2AK129ac. We chose to synthesize H2AK129ac to demonstrate that our method can be applied to the synthesis of histones with C- terminal modifications. As a matter of fact, except H3 that natively has a cysteine, all other core histones lack cysteine residues. We first recombinantly produced H2A-K129C-6H, an H2A protein with a K129C mutation and a C-terminal 6×His tag and then ligated it to N^{ε}-acetyl-lysine with the assistance of NTCB. The ESI-MS spectrum of the reaction product showed the formation of H2AK129ac (Figures 18A) with an at least 60% yield. We folded successfully

H2AK129ac into a dimer with H2B and subsequently into a nucleosome (Figure 18B) making it possible to study effects of H2AK129ac on the nucleosome structure and function. In the ESI-MS spectrum of H2AK129ac in Figure 18A, we noticed a minor peak at 13813.6 Da. H2A has two lysine residues, K125 and K127, that are adjacent to K129. Both crystal and cryo-EM structures have indicated all three lysines are located at a flexible C-terminal region of H2A.[117, 118] Potentially either K125 or K127 can undergo intramolecular cyclization with an activated K129C. These intramolecular cyclization products will have a theoretical molecular weight (13813.9 Da) that matches the minor peak in the ESI-MS spectrum. Based on the determined intensities of peaks in the ESI-MS spectrum, this minor peak was lower than 15% of the desired product demonstrating the selectivity of ACPL. By tuning the reaction conditions, this minor peak might be further reduced.

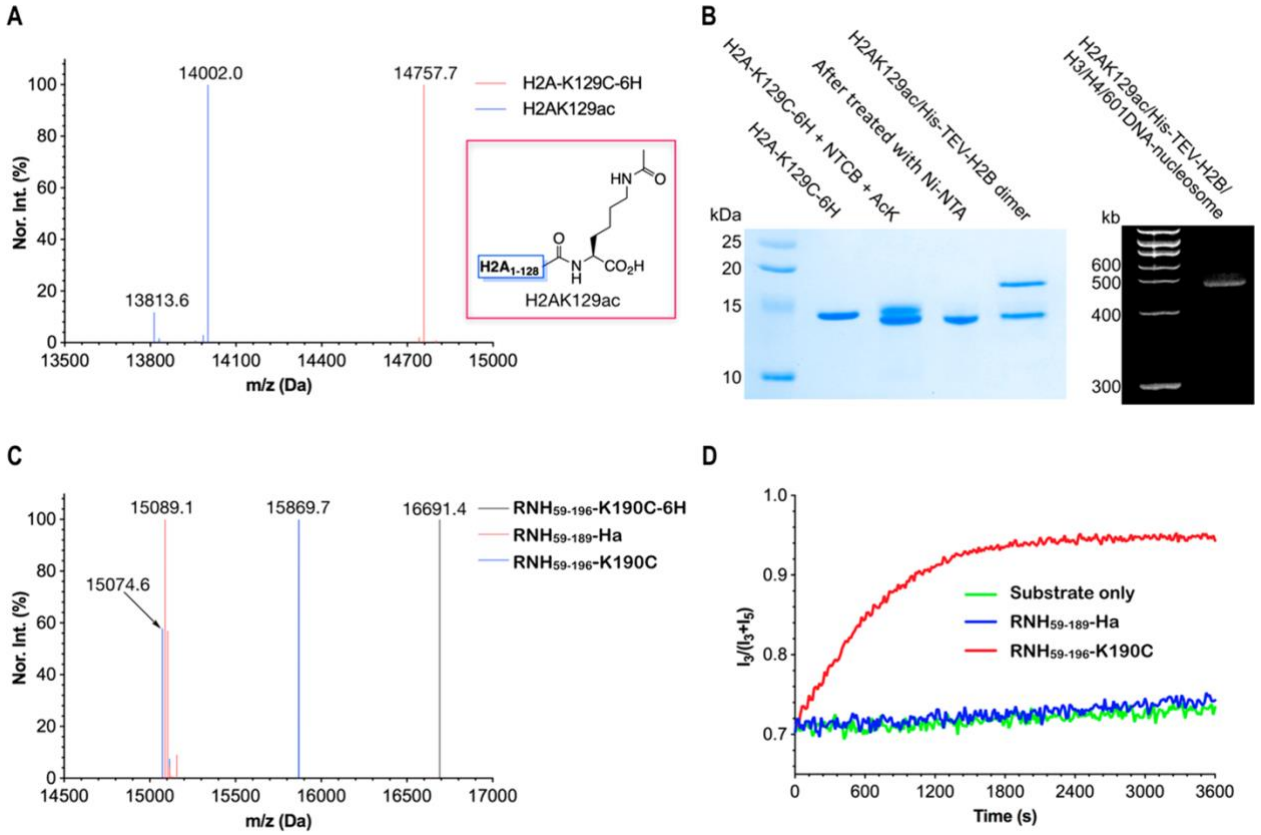


Figure 18 Synthesis of H2AK129ac and RNase H by ACPL (A) The deconvoluted and integrated ESI-MS spectra of H2A-K129C-6H and H2AK129ac. H2A-K129C-6H was recombinantly expressed and then reacted with NTCB and N^{ε}-acetyl-lysine to afford H2AK129ac. (B) The synthesis of H2AK129ac, its isolation, and folding into an H2AK129ac/H2B dimer and then a nucleosome. The purification of H2AK129ac was achieved by extracting the unreacted intermediate using Ni charged resins. (C) The deconvoluted and integrated ESI-MS spectra of $RNH_{59–196}$- K190C-6H, $RNH_{59–189}$-Ha, and $RNH_{59–196}$-K190C. $RNH_{59–196}$-K190C-6H was recombinantly expressed in E. coli. It was reacted with NTCB and Ha to afford $RNH_{59–189}$-Ha that then underwent peptide hydrazide ligation with a 7-mer NH_2-CADYGRK-OH peptide to form a

catalytic active RNH59–196-K190C. (D) The catalytic hydrolysis of an RNA substrate by RNH59–196-K190C. The RNA substrate had a sequence 5′-Cy3-GACACCUGAUUC-Cy5-3′. A DNA fragment 5′-GAATCAGGTGTC-3′ was used to form a double strand with the RNA substrate for binding to RNH59–196-K190C. The hydrolysis led to improved Cy3 (I_3) and decrease Cy5 (I_5) emission.

Use of ACPL in Combination with Peptide Hydrazide Ligation To Synthesize an Active RNase H.

For all ligation reactions that we performed thus far, they involved small molecules with only one amino group for avoiding side product formation. For ligation with larger molecules or peptides that have more than one amino group, one can couple ACPL with peptide hydrazide ligation to resolve nonspecificity issues. To demonstrate this prospect, we recombinantly produced a B. halodurans RNase H region with a C-terminal Cys-6×His tag (RNH59–196-K190C-6H). Its ligation with Ha in the presence of NTCB led to the synthesis of RNH59–189- Ha, a protein hydrazide that proceeded further to undergo peptide hydrazide ligation with a 7-mer peptide, NH_2- CADYGRK-OH to afford a ligated product RNH59–196- K190C.[119] ESI-MS analysis showed the successful synthesis of both RNH59–189-Ha and RNH59–196-K190C (Figure 18C). Similar to what has been found in previous peptide hydrazide ligation reactions, we also detected a minor hydrolysis product at 15074.6 Da.[120] The ligated product RNH59–196-K190C was catalytically active to hydro- lyze an RNA substrate as shown in Figure 18D. In

contrast, RNH59–189-Ha was completely inactive toward this substrate. Our data related to the synthesis of RNase H demonstrated that ACPL can couple to peptide hydrazide ligation for conjugation with large peptides or even protein fragments.

The Use of ACPL To Synthesize Exenatide, an Anti- Diabetic Medication from a Recombinant Precursor.

Exenatide is a 39-mer commercial antidiabetic peptide drug that has a C-terminal amide. The presence of this C-terminal amide makes it difficult to produce exenatide using the recombinant expression technique.[121] This bottleneck can be presumably resolved using our ACPL technique by recombi- nant expression of a precursor protein and then an ACPL reaction with serinamide. To demonstrate this potential, we expressed a 6×His-SUMO-exenatide-S39C-SA-Strep fusion protein that can be largely produced in E. coli followed by the treatment with SUMO protease to obtain the exenatide- S39C-SA-Strep peptide. We then proceeded to carry out its reaction with L-serinamide in the presence of NTCB, purified the final product with a Strep-Tactin column, and analyzed it by ESI-MS. Our results showed that exenatide can be easily procured using this approach

2.3 Conclusions

In summary, we have developed a novel expressed protein ligation technique that uses a cyanylating reagent to directly activate a cysteine in a recombinant protein for ligation with small molecule amines and large peptide or protein fragments when coupling with peptide hydrazide ligation. These small molecule amines are primary or secondary and

include a number of L- and D-amino acids and functional amines such as Ha, Aa, Pa, and Gly-AMC. The technique that is termed ACPL requires no enzymatic catalysis and is controllable, versatile, specific, and very simple to process. Therefore, it can be broadly applied to synthesize a large variety of proteins with unique functionalities for advanced applications in both basic and applied research. One potential industrial application of the technique is to synthesize therapeutic peptides or proteins like exenatide. ACPL requires the activation of cysteine, one of the two lowest occurring amino acids in proteins. Nontargeted cysteines need to be mutated. For proteins with essential cysteines, one solution for using ACPL is to couple it with the noncanonical amino acid mutagenesis technique. Photocaged cysteines have been genetically incorporated into proteins by amber suppression.[122, 123] The incorporation of a photocaged cysteine to essential cysteine sites in a protein followed by ACPL and then de-caging to release protected essential cysteines will allow the processing of proteins with nontargeted cysteines. Overall, our ACPL technique expands to a large extent the synthetic capacity of protein chemistry and will energize the whole field. We anticipate its broad applications in a large variety of research fields and industrial processing of proteins and peptides.

2.4 Experimental Procedures

2.4.1 Synthetic procedures

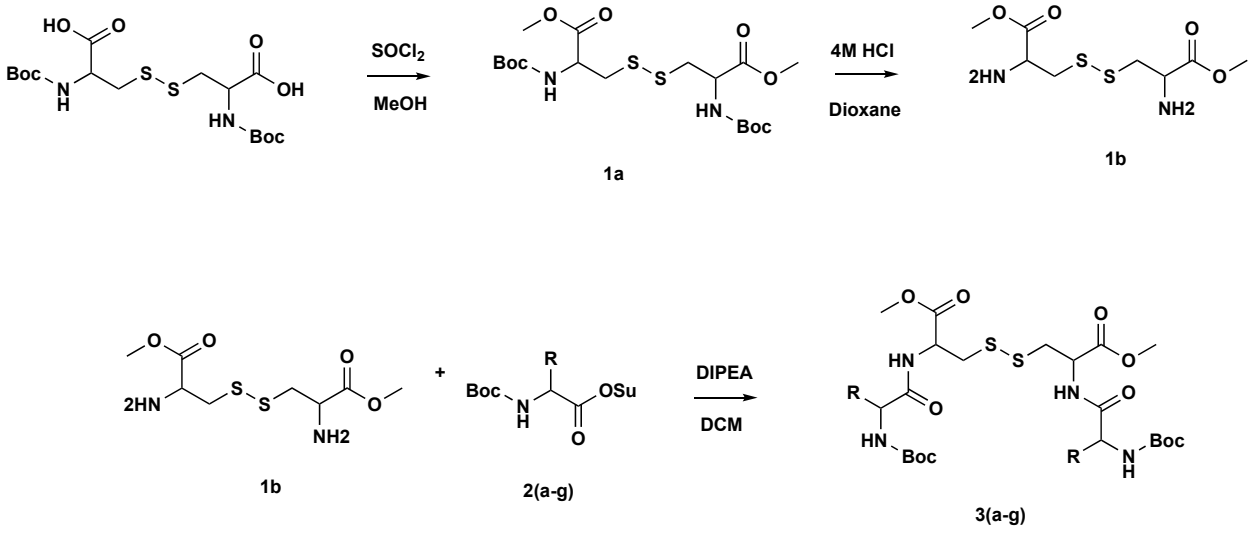


Figure 19 General synthetic route of dipeptides

Boc-Cystine-OMe(1a)

To a suspension of Boc-Cystine-OH (1 eq.) in dry methanol (20 mL per 0.8 g) thionyl chloride (2 eq.) was added slowly at room temperature. Temperature was increased to 80 °C over 15 min. Reaction was stirred for 3 h then left to cool down and stirred overnight at room temperature. Volatiles were evaporated in vacuo and the residue was dissolved in minimal amount of methanol. Crystals of crude product crashed out from solution by addition of diethyl ether. Product was filtered, washed twice with diethyl ether and dried to obtain 1a as white crystals. Final yield was 90%. 1H NMR (CDCl3) δ =1.45 (s, 18 H), 3.16 (d, 4 H, J = 5.0 Hz), 3.76 (s, 6 H), 4.60 (dt, 2 H, J= 6.3, 5.0 Hz), 5.38 (d, 2H, J = 6.3 Hz).

Cystine-OMe(1b)

A solution of 1a (1 eq.) in DCM was cooled to 0 °C. 4 M HCl in Dioxane (1.1 eq.) was added dropwise. Solution was stirred for 2 h and the reaction progress was monitored by TLC. Solvent was removed in vacuo and product was obtained as white solid without further purification. Quantitative yield. 1H NMR (CDCl3) δ = 3.16 (d, 4 H, J = 5.0 Hz), 3.76 (s, 6 H), 4.60 (dt, 2 H, J = 6.3, 5.0 Hz), 5.38 (d, 2H, J = 6.3 Hz) 7.80 (s, 1H).

Gly-Cystine-OMe(3a)

A solution of Boc-Gly-Osu 2a (2 eq.) in dry DCM was added into a mixture of 1b (1 eq.) and DIPEA (2.2 eq.) at 0 °C and stirred for 15-30 min. The solution was then warmed to room temperature and left to stir overnight. The solution was washed with 3 M HCl twice then washed with brine twice. Organic layer was separated, dried over sodium sulfate, and concentrated under reduced pressure. The residue was purified by column chromatography (30% EtOAc/hexane). 1H NMR (CDCl3) δ = 1.42 (s, 18H), 3.16 (d, 4 H, J = 5.0 Hz), 3.76 (s, 6 H), 3.88 (s, 2H) 4.60 (t, 2 H) 7.80 (s, 1H).

Ala-Cystine-OMe (3b)

A solution of Boc-Ala-Osu 2b (2 eq.) in dry DCM was added dropwise into a mixture of 1b (1 eq.) and DIPEA (2.2 eq.) at 0 °C and stirred for 15-30 min. The solution was then warmed to room temperature and left to stir overnight. The solution was

washed with 3 M HCl twice then washed with brine twice. Organic layer was separated, dried over sodium sulfate, and concentrated under reduced pressure. The residue was purified by column chromatography (30% EtOAc/hexane). 1H NMR (CDCl3) δ = 0.98 (d, 6H, J = 6.8 Hz), 1.42 (s, 18H) 3.16 (d, 4 H, J = 5.0 Hz), 3.76 (s, 6 H), 4.60 (t, 2 H, J = 6.3, 5.0 Hz) 7.80 (s, 1H).

Leu-Cystine-OMe (3c)

A solution of Boc-Leu-Osu 2c (2 eq.) in dry DCM was added dropwise into a mixture of 1b (1 eq.) and DIPEA (2.2 eq.) at 0 °C and stirred for 15-30 min. The solution was then warmed to room temperature and left to stir overnight. The solution was washed with 3 M HCl twice then washed with brine twice. Organic layer was separated, dried over sodium sulfate, and concentrated under reduced pressure. The residue was purified by column chromatography (30% EtOAc/hexane). 1H NMR (CDCl3) δ = 0.90 (d, 12H, J = 6.8 Hz), 1.42 (s, 18H), 1.49 (q, 2H), 1.76 (t, 4H), 3.16 (d, 4H, J = 5.0 Hz), 3.76 (s, 6 H), 4.40 (t, 2H), 4.60 (t, 2H, J = 6.3, 5.0 Hz), 7.80 (s, 1H).

Phe-Cystine-OMe (3d)

A solution of Boc-Phe-Osu 2d (2 eq.) in dry DCM was dropped into a mixture of 1b (1 eq.) and DIPEA (2.2 eq.) at 0 °C and stirred for 15-30 min. The solution was then warmed to room temperature and left to stir overnight. The solution was washed with 3 M HCl twice then washed S5with brine twice. Organic layer was separated, dried over sodium sulfate, and concentrated under reduced pressure. The residue was purified by

column chromatography (30% EtOAc/hexane). 1H NMR (CDCl3) δ = 1.42 (s, 18H), 3.16 (d, 2H, J = 5.0 Hz), 3.40 (q, 2H, J = 7.0 Hz) 3.76 (s, 6 H), 4.40 (t, 2H), 4.60 (t, 2H, J = 6.3, 5.0 Hz), 7.14 (m, 10H), 7.80 (s, 1H).

Trp-Cystine-OMe (3e)

A solution of Boc-Trp-Osu 2e (2 eq.) in dry DCM was added dropwise into a mixture of 1b (1 eq.) and DIPEA (2.2 eq.) at 0 °C and stirred for 15-30 min. The solution was then warmed to room temperature and left to stir overnight. The solution was washed with 3 M HCl twice then washed with brine twice. Organic layer was separated, dried over sodium sulfat,e and concentrated under reduced pressure. The residue was purified by column chromatography (30% EtOAc/hexane). 1H NMR (CDCl3) δ = 1.42 (s, 18H), 3.16 (d, 2H, J = 5.0 Hz), 3.40 (q, 2H, J = 7.0 Hz) 3.76 (s, 6 H), 4.40 (t, 2H), 4.60 (t, 2H, J = 6.3, 5.0 Hz), 6.9 (t, 1H), 7.14 (t, 1H), 7.3 (t, 1H), 7.60 (t, 1H), 7.90 (s, 1H).

Asp-Cystine-OMe (3f)

A solution of Boc-Asp-Osu 2f (2 eq.) in dry DCM was added dropwise into a mixture of 1b (1 eq.) and DIPEA (2.2 eq.) at 0 °C and stirred for 15-30 min. The solution was then warmed to room temperature and left to stir overnight. The solution was washed with 3 M HCl twice then washed with brine twice. Organic layer was separated, dried over sodium sulfate, and concentrated under reduced pressure. The residue was purified by column chromatography (30% EtOAc/hexane). 1H NMR (CDCl3) δ = 1.42

(s, 18H), 2.79 (q, 4H, J = 7.0 Hz) 3.16 (d, 2H, J = 5.0 Hz), 3.40 (q, 2H, J= 7.0 Hz) 3.76 (s, 6 H), 4.40 (t, 2H), 4.60 (t, 2H, J = 6.3, 5.0 Hz), 4.90 (t, 1H), 7.90 (s, 1H), 11.20 (s, 1H).

Pro-Cystine-OMe (3g)

A solution of Boc-Asp-Osu 2g (2 eq.) in dry DCM was added into a mixture of 1b (1 eq.) and DIPEA (2.2 eq.) at 0 °C and stirred for 15-30 min. The solution was then warmed to room temperature and left to stir overnight. The solution was washed with 3 M HCl twice then washed with brine twice. Organic layer was separated, dried over sodium sulfate, and concentrated under reduced pressure. The residue was purified by column chromatography (30% EtOAc/hexane). 1H NMR (CDCl3) δ = 1.42 (s, 18H), 2.13 (m. 8H) 3.16 (m, 4H), 3.40 (m, 4H) 3.76 (s, 6 H), 4.26 (br, 2H), 4.78 (br, 2H), 6.90 (br, 1H), 7.90 (s, 1H).

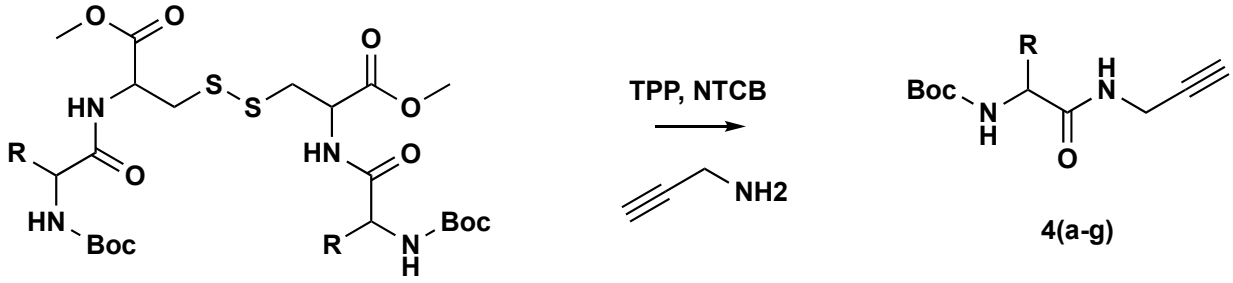


Figure 20 Dipeptide reactions with NTCB and propargylamine

Boc-Gly-Propargylamine (4a)

Boc-Gly-Cys-OMe 3a (1 eq.) and TPP (1.1 eq.) were mixed in DCM at room temperature for 40 min or until reaction completion by monitoring with TLC. After completion NTCB (2 eq.) was added at room temperature and left to stir for 1 h. After the formation of the thiazolidine intermediate, confirmed by TLC, propargylamine was added (1.1 eq.) and left to stir overnight. The reaction was washed in 1 M HCl thrice and brine thrice, the organic layer was separated, dried over sodium sulfate, and concentrated in vacuo. Crude residue was purified by column chromatography with a 0-20% EtOAc/hexane gradient. 1H NMR (CDCl3) δ = 7.29 (s, 1H), 5.41 (s, 1H), 4.08 (dd, 2H, J = 5.4 Hz, 2.7 Hz), 3.84 (d, 2H, J = 5.7 Hz), 2.25 (t, 1H, J = 2.7 Hz), 1.47 (s, 9H).

Boc-Ala-Propargylamine(4b)

Boc-Ala-Cys-OMe 3b (1 eq.) and TPP (1.1 eq.) were mixed in DCM at room temperature for 40 min or until reaction completion by monitoring with TLC. After completion NTCB (2 eq.) was added at room temperature and left to stir for 1 h. After the formation of the thiazolidine intermediate, confirmed by TLC, propargylamine was added (1.1 eq.) and left to stir overnight. The reaction was washed in 1 M HCl thrice and brine thrice, the organic layer was separated, dried over sodium sulfate, and concentrated in vacuo. Crude residue was purified by column chromatography with a 0-20% EtOAc/hexane gradient. 1H NMR (CDCl3) δ= 1.37 (3H, d, J = 6.8 Hz). 1.46 (s, 9H), 2.23 (t, 1H J = 2.4 Hz), 4.05 (m, 2H), 4.16 (m, 1H), 4.92 (s,1H), 6.43 (s,1H).

Boc-Leu-Propargylamine(4c)

S7Boc-Leu-Cys-OMe 3c (1 eq.) and TPP (1.1 eq.) were mixed in DCM at room temperature for 40 min or until reaction completion by monitoring with TLC. After completion NTCB (2 eq.) was added at room temperature and left to stir for 1 h. After the formation of the thiazolidine intermediate, confirmed by TLC, propargylamine was added (1.1 eq.) and left to stir overnight. The reaction was washed in 1 M HCl thrice and brine thrice, the organic layer was separated, dried over sodium sulfate, and concentrated in vacuo. Crude residue was purified by column chromatography with a 0-20% EtOAc/hexane gradient. 1H NMR (300 MHz, CDCl3):δ= 0.90 (br d, 6H), 1.41 (s, 9H), 1.59 (m, 3H) 2.18 (s, 1H), 3.99 (s, 2H), 4.18 (br s, 1H), 5.25 (s, 1H), 7.13 (br s, 1H).

Boc-Phe-Propargylamine(4d)

Boc-Phe-Cys-OMe 3d (1 eq.) and TPP (1.1 eq.) were mixed in DCM at room temperature for 40 min or until reaction completion by monitoring with TLC. After completion NTCB (2 eq.) was added at room temperature and left to stir for 1 h. After the formation of the thiazolidine intermediate, confirmed by TLC, propargylamine was added (1.1 eq.) and left to stir overnight. The reaction was washed in 1 M HCl thrice and brine thrice, the organic layer was separated, dried over sodium sulfate, and concentrated in vacuo. Crude residue was purified by column chromatography with a 0-20% EtOAc/hexane gradient. 1H NMR (CDCl3) δ = 1.37 (s, 9H), 2.18 (s, 1H), 2.98 (m, 1H), 3.06 (m, 1H), 3.95 (s, 2H), 4.47 (m, 1H), 5.49 (d, J = 9.0 Hz, 1H), 7.04 (br s, 1H), 7.17-7.27 (m, 5H).

Boc-Trp-Propargylamine(4e)

Boc-Trp-Cys-OMe 3e (1 eq.) and TPP (1.1 eq.) were mixed in DCM at room temperature for 40 min or until reaction completion by monitoring with TLC. After completion NTCB (2 eq.) was added at room temperature and left to stir for 1 h. After the formation of the thiazolidine intermediate, confirmed by TLC, propargylamine was added (1.1 eq.) and left to stir overnight. The reaction was washed in 1 M HCl thrice and brine thrice, the organic layer was separated, dried over sodium sulfate, and concentrated in vacuo. Crude residue was purified by column chromatography with a 0-20% EtOAc/hexane gradient. 1H NMR (CDCl3) δ =1.41 (s, 9H), 1.86 (m, 2H), 2.19 (m, 2H), 3.28-3.41 (m, 2H), 3.98-4.22 (m, 3H), 6.38 (br s, 1H), 7.23 (br s, 1H).

Boc-Asp-Propargylamine(4f)

Boc-Asp-Cys-OMe 3f (1 eq.) and TPP (1.1 eq.) were mixed in DCM at room temperature for 40 min or until reaction completion by monitoring with TLC. After completion NTCB (2 eq.) was

S8added at room temperature and left to stir for 1 h. After the formation of the thiazolidine intermediate, confirmed by TLC, propargylamine was added (1.1 eq.) and left to stir overnight. The reaction was washed in 1 M HCl thrice and brine thrice, the organic layer was separated, dried over sodium sulfate, and concentrated in vacuo. Crude residue was purified by column chromatography with a 0-20% EtOAc/hexane gradient.

1H NMR (CDCl3) δ = 1.37 (s, 9H), 2.23 (d, 2H, J = 2.4 Hz), 3.06 (m, 1H), 4.05 (m, 2H), 4.92 (t, 1H), 5.7 (br, 1H), 8.1 (br, 1H), 12.30 (br, 1H).

Pro-Propargylamine(4g)

Pro-Cys-OMe 3g (1 eq.) and TPP (1.1 eq.) were mixed in DCM at room temperature for 40 min or until reaction completion by monitoring with TLC. After completion NTCB (2 eq.) was added at room temperature and left to stir for 1 h. After the formation of the thiazolidine intermediate, confirmed by TLC, propargylamine was added (1.1 eq.) and left to stir overnight. The reaction was washed in 1 M HCl thrice and brine thrice, the organic layer was separated, dried over sodium sulfate, and concentrated in vacuo. Crude residue was purified by column chromatography with a 0-20% EtOAc/hexane gradient. 1H NMR (CDCl3) δ =1.41 (s, 9H), 1.86 (m, 2H), 2.19 (m, 2H), 3.28-3.41 (m, 2H), 3.98-4.22 (m, 3H), 6.38 (br s, 1H), 7.23 (b

2.5 Plasmid construction

The gene coding Ub-G76C-6H and Ub-C-6H were inserted into the pETDuet-1- Ub vector[24] to afford the expression vector pETDuet-1-Ub-G76C-6H and pETDuet-1- Ub-C-6H. Constructs for Ub-K6/11/27/29/33/48/63C were obtained from lab cell stocks. Constructs for Ub G75P/T/L/E/R/W mutants were made by site-directed mutagenesis from the pETDuet-1-Ub-G76C-6H vector. The expression vector for FLAG-Ub-G76C-

6H was generated by inserting a FLAG coding DNA at the *N*-terminal side of the Ub-G76C-6H coding sequence in the pETDuet-1-Ub-G76C-6H vector. The DNA fragments for all Ubls that contained both an *N*-terminal FLAG tag and a *C*-terminal 6×His tag were ordered from IDT Inc. and were cloned into the pETDuet-1 vector to afford expression vectors for them. The H2A-K129C-6H expression vector was constructed from the pETDuet-1-6H-TEV-H2A vector. For the RNH$_{59\text{-}196}$-K190C-6H expression vector, the DNA fragment for *B. halodurans* RNase H (residues 59-196 and a K190C mutation) was cloned into the pET28a vector right before the sequence of a C-terminal 6×His tag using the MEGAWHOP technique. The gene fragment for exenatide-40C-GGGSA-Strep was cloned into the pET28a-SUMO vector right after the SUMO cleavage site using the MEGAWHOP technique. The pET28a-6H-SUMO-exenatide-S39C-SA-Strep was constructed by site-directed mutagenesis of the pET28a-6H-SUMO-exenatide-40C-GGGSA-Strep vector. pET28a-SUMO protease was kindly provided by Prof. Pingwei Li.

2.6. Recombinant protein expression and purification

2.6.1. Expression and purification of Ub proteins

The expression of all Ub mutants and FLAG-Ub were performed according to the protocol described before.[24] Briefly, overnight culture of each protein was inoculated into 1 L 2xYT medium and let grow under 37 °C until OD_{600} reached 0.6-1.0. 1 mM IPTG was added to induce the expression at 18 °C overnight. Upon saturation, cells were

collected by centrifugation (6000 rpm, 20 min, 4 °C) and stored at -80 °C if not lysed directly.

For purification of each all seven Ub K to C mutant, the bacterial pellet was resuspended in Ub lysis buffer (50 mM Tris, 1 mM TCEP, pH 7.8) supplied with 0.2 mg/mL lysozyme (Sigma-Aldrich, St. Louis, MO, USA) and lysed by sonication on ice. The total cell lysate was clarified by centrifugation (10000 rpm, 30 min, 4 °C) and supernatant was collected. After that, 6 M HCl solution was gradually added into the

supernatant with constant stirring to adjust pH value to 1.5-2. The white precipitate was then removed by centrifugation (10000 rpm, 30 min, 4 °C) and the pH of the acid purification supernatant was adjusted back to 7.8 using 6 M NaOH and then concentrated using Amicon stirring filtration system with a 5k MWCO membrane (EMD Millipore, Burlington; MA, USA). Subsequently, each protein was desalted into 50 mM ammonium bicarbonate (ABC) buffer by HiPrep 26/10 Desalting column (GE Healthcare, Chicago, IL, USA) using NGCTM chromatography system (Bio-Rad Laboratories, Hercules, CA, USA) and analyzed by 15% SDS-PAGE. The concentration of protein solution was measured by PierceTM 660nm Protein Assay Reagent (Thermo Fisher Scientific, Waltham, MA, USA) and dispensed into 100 nmol aliquots for lyophilization. Eventually, protein pellets were either used for chemical reactions or kept at -80 °C for long-term storage.

For purification of all his-tagged Ub mutants, Ni binding buffer (50 mM NaH_2PO_4, 500 mM NaCl, 5 mM imidazole, 1 mM TCEP, pH 7.8) was used instead of Ub lysis buffer to

resuspend and lyse the cell pellet. The subsequent steps were identical to the purification of Ub-KxC until finishing the acid purification. Acid supernatant was directly passed through high affinity Ni charged resin (Genescript, Piscataway, NJ, USA) without concentration. Resin was then washed by Ni washing buffer (50 mM NaH_2PO_4, 500 mM NaCl, 25 mM imidazole, 1 mM TCEP, pH 7.8) and eluted in 7 mL elution buffer (50 mM NaH_2PO_4, 500 mM NaCl, 300 mM imidazole, 1 mM TCEP, pH 7.8). The following desalting, SDS-PAGE, concentration measurement, lyophilization and storage as described above.

2.6.2. Expression and purification of FLAG-Ubl proteins

An overnight culture of E. coli BL21(DE3) cells harboring a Ubl expression vector was inoculated (1:100 dilution) into LB medium containing 100 μg/mL ampicillin. Cells were let grow at 37 oC until OD_{600} reached 0.6-0.9 and then 1 mM IPTG was added to induce protein expression. Induced cells were let grow in a refrigerated incubating shaker at 18 oC overnight and harvested by centrifugation (4000 rpm, 20 min, 4 oC). The cell pellet was then lysed, purified by Ni^{2+}-NTA resins, and desalted as same as mentioned previously for his-tagged Ub proteins. Yield for all Ubl proteins were around 10-40 mg/L. Eventually, the proteins were aliquoted, lyophilized, and stored at -80 oC.

2.6.3. Expression and purification of H2A-K129-6H

The expression vector pETDuet-1-H2A-K129C-6H was used to transform E. coli BL21(DE3) strain, and a single colony was inoculated into 2×YT medium containing 100 μg/mL ampicillin. After cell grew to OD_{600} about 0.6 at 37 oC, 0.5 mM IPTG was added to induce protein expression. Four hours later, cells were harvested and purified according to a previous protocol.[215]

2.6.4. Expression and purification of RNH59-196-K190C-6H

An overnight culture of E. coli BL21(DE3) cells which contained the pET28a- RNH_{59-196}-K190C-6H vector was inoculated (1:100 dilution) into 2×YT medium containing 50 μg/mL kanamycin. The protein was expressed by inducing with 0.5 mM IPTG when OD_{600} reached 0.6 at 37 oC for 3 h. The harvested cells were lysed in a buffer (20 mM Tris, 200 mM NaCl, pH 8.0) by sonication at 4 oC, the lysate was clarified by centrifugation (10,000 rpm, 30 min, 4 oC). Supernatant was collected and incubated with Ni^{2+}-NTA resins at rt for 45 min before washed with a 20× resin volume of lysis buffer supplemented with 35 mM imidazole. The expressed 6×His-tagged protein was eluted by another buffer (20 mM Tris, 200 mM NaCl, 300 mM imidazole, pH 8.0). Collected elution was concentrated and desalting to 50 mM ABC buffer by HiTrap Desalting columns (GE Healthcare). Proteins were aliquoted and lyophilized for further assays.

2.6.5. Expression and purification of SUMO protease

An overnight culture of E. coli BL21(DE3) cells harboring a pET28a-SUMO

vector was inoculated (1:100 dilution) into LB medium containing 50 μg/mL kanamycin. Cells were let grown at 37 oC until OD_{600} reached 1.3 and 1 mM IPTG was added to induce protein expression. Inducted cells were let grow in a refrigerated incubating shaker at 16 oC overnight and harvested by centrifugation (4000 rpm, 20 min, 4 oC). The cell pellet was resuspended, lysed, purified by Ni^{2+}-NTA resins, and desalted according to the same protocol for Ub proteins except the use of DTT instead of TCEP. The expression yield was about 4.1 mg/L. The finally purified protein was aliquoted into 0.2 mg aliquots and stored at -80 oC for the future usage.

2.6.6. Expression and purification of exenatide

An overnight culture of *E. coli* BL21(DE3) cells harboring the 6H-SUMO- exenatide-S39C-SA-Strep expression vector was inoculated (1:100 dilution) into LB medium containing 50 μg/mL kanamycin. Cells were let grow at 37 oC until OD_{600} reached 0.6-0.9 and then 1 mM IPTG was added to induce protein expression. Induced cells were let grow in a refrigerated incubating shaker at 18 oC overnight and harvested by centrifugation (4000 rpm, 20 min, 4 oC). The sonication with addition of one protease inhibitor cocktail tablet (Sigma-Aldrich), Ni^{2+}-NTA purification, and desalting were performed as same as what were described previously. After desalting, one aliquot of SUMO protease stock solution (0.2 mg) was added into the eluted protein solution along with 1 mM DTT and stirred at 4 oC overnight. On the next day, the digested sample was loaded onto a HisTrap HP column (GE Healthcare) and the flow-through was collected.

The exenatide-S39C-SA-Strep yield was determined as 11 mg/L. The final exenatide-S39C-SA-Strep solution was aliquoted and lyophilized into protein pellets and then stored at -80 oC.

2.7 Chemical reactions on protein

2.7.1. ACPL of Ub and Ubl proteins

A 500 mM TCEP stock and NTCB stock were prepared in water and DMSO, respectively. For reactions with all Ub proteins, solutions of small amine-containing compounds (glycine, proline, aspartic acid, arginine, D-alanine, D-Serine, allylamine, and propargylamine) were prepared in water to a concentration of 1 M along with 6 M guanidinium chloride (GndCl). For asparagine, leucine, and phenylalanine, their concentrations were 500, 150, and 100 mM respectively due to their poor solubility. Hydrazine solution was prepared as 50 mM in 1× PBS buffer along with 6 M GndCl. The pH of all solutions was pre-adjusted to 9 before setting up reactions. 40 mM Gly-AMC solution was prepared in 40% DMSO without pH adjustment. To set up reactions, Ub protein pellets were dissolved by the previously described amine-containing compound solutions and 0.5 mM TCEP and 5 mM NTCB were sequentially introduced into the reaction mixture. For the reaction with Gly-AMC, pH was adjusted to 8 after adding TCEP and NTCB prior to the incubation. The final concentration of proteins was 200 μM (2 mg/mL). All reaction mixtures were incubated at 37 oC for 16 h and followed by desalting using HiTrap Desalting column (GE Healthcare, Chicago, IL, USA). Fractions with UV absorbance was collected and analyzed by 15% SDS-PAGE. Then,

Ni^{2+}-NTA resins were incubated with the desalted products at rt for 30 min. After that, Ni^{2+}-NTA resins were removed by micro syringe filters and flowthroughs were collected for ESI-MS analysis. For reactions between UbK48/63C or Ub-G76C-6H and hydrazine without GndCl, the reaction setup was identical except that no GndCl was provided.

For all Ubl proteins, the reaction setup was identical as aforementioned except in the amine buffer preparation. In this case, no GndCl was added and the amine used was only 1 M propargylamine. The reaction between Gly-AMC and SUMO-1/2/3 was also initiated by mixing all reagents first and then adjusting pH to 8.

2.8 ESI-MS analysis and data processing

Samples for MS were placed in tube inserts inside HPLC sample vials for auto injection. The positive ESI-LC-MS was carried out using Q Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) connected to a liquid chromatography instrument. For the liquid chromatography, 100% water and 100% acetonitrile that both supplied with 0.1% formic acid were used as mobile phase A and B respectively. All protein samples were separated by AccucoreTM 150-C4 analytical HPLC column (150 mm × 2.1 mm, 2.6 μm particle size) (Thermo Fisher Scientific, Waltham, MA, USA). ESI-MS raw data was exported as .txt files using Xcalibur 4.1.31.9 Qual Browser (Thermo Fisher Scientific, Waltham, MA, USA). Deconvolution of each raw file was processed by Bayesian Protein Reconstruction using Analyst^{TM} QS v1.1 software

(Applied Biosystems, Foster City, CA, USA). The deconvolution parameters are set as following: Adduct: hydrogen; Step mass: 0.1 Da; S/N threshold: 20; Minimum intensity (%): 5; Iteration: 20. Start and stop mass were inputted according to the observed mass of each protein sample. Deconvoluted results were then replotted using GraphPad Prism

CHAPTER III DRUG REPURPOSING: BEPRIDIL IS POTENT AGAINST SARS-COV-2 IN VITRO*[2]

3.1 Introduction

The current worldwide impact of the COVID-19 pandemic has been so profound that it is often compared to that of 1918 influenza pandemic.[124, 125] As of October 3rd, 2020, the total global COVID-19 cases had surpassed 34 million, among which more than 1 million had succumbed to death.[126] Important political figures who have been diagnosed COVID-19 positive include the US president Donald Trump Jr. and the UK prime minister Boris Johnson. A modelling study has predicted that this pandemic will continue to affect everyday life and the circumstances may require societies to follow social distancing until 2022.[127] Finding timely treatment options is of tremendous importance to alleviate catastrophic damages of COVID-19. However, the short time window that is required to contain the disease is extremely challenging for a conventional drug discovery process that requires typically many years to finalize a drug and therefore might not achieve its goal before the pandemic ceases. In this January, we did a comparative biochemical analysis between severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2), the virus that has caused COVID-19, and SARS-CoV-1 that led to an epidemic in China in 2003 and proposed that remdesivir was a viable choice for the treatment of COVID-19.[128] We were excited to see that remdesivir was finally approved for emergency use in the

* Reprinted with permission by "DRUG REPURPOSING: BEPRIDIL IS POTENT AGAINST SARS-CoV-2 IN VITRO" by Erol Vatansever, Kai Yang, Kaci C. Kratch, Aleksandra Drelich, Chia-Chuan Cho, Drake M. Mellott, Shiqing Xu, Chien-Te K. Tseng, and Wenshe Ray Liu, 2021, PNAS, © [2021] PNAS

United States and for use in Japan for people with severe symptoms. With only one medicine in stock that provides very limited benefits to COVID-19 patients,[129] the virus

may easily evade it, leading us once again with no medicine to use. Given the rapid spread and the high fatality of COVID-19, finding alternative medicines is imperative. Drug repurposing stands out as an attractive option in the current situation. If an approved drug can be identified to treat COVID-19, it can be quickly proceeded to clinical trials and manufactured at a large scale using its existing GMP lines. Previously, encouraging results were obtained from repurposing small molecule medicines including teicoplanin, ivermectin, itraconazole, and nitazoxanide.[130-133] These antimicrobial agents showed antiviral activity against Ebola, Chikungunya, Enterovirus, Influenza viruses respectively.[134] However, a common drawback of a repurposed drug is its low efficacy level. One way to circumvent this problem is to combine multiple existing medicines to accrue a synergistic effect. To be able to discover such combinations, breaking down the druggable targets of the SARS-CoV-2 to identify drugs that do not cross-act on each other's targets is a promising strategy. For example, a recent study showed that triple combination of interferon b-1b, lopinavir-ritonavir, and ribavirin was safe and superior to lopinavir-ritonavir alone for treating COVID-19 patients.[135]

In our January paper, we recommended four SARS-CoV-2 essential proteins including Spike, RNA-dependent RNA polymerase, the main protease (M^{Pro}), and papain-like protease as drug targets for the development of anti-COVID-19 therapeutics. Among these four proteins, M^{Pro} that was previously called 3C-like protease provides the most facile opportunity for drug repurposing owing to the ease of its biochemical assays. M^{Pro} is a cysteine protease that processes itself and then cleaves a number of nonstructural viral

proteins from two very large polypeptide translates that are made from the viral genomic RNA in the human cell host.[136] Its relatively large active site pocket and a highly nucleophilic, catalytic cysteine residue make it likely to be inhibited by a host of existing and investigational drugs. Previous work has disclosed some existing drugs that inhibit M^{Pro}.[137] However, complete characterization of existing drugs on the inhibition of M^{Pro} is not yet available. Since the release of the first M^{Pro} crystal structure, many computational studies have been carried out to screen existing drugs in their inhibition of M^{Pro} and many potent leads have been proposed.[138-141] However, most of these lead drugs are yet to be confirmed experimentally. To investigate whether some existing drugs can potently inhibit M^{Pro}, we have docked a group of selected FDA/EMA-approved small molecule medicines to the active site of M^{Pro} and selected about 30 hit drugs to characterize their inhibition on M^{Pro} experimentally. Our results revealed that a number of FDA/EMA-approved small molecule medicines have high potency in inhibiting M^{Pro} and bepridil inhibits cytopathogenic effect (CPE) induced by the SARS-CoV-2 virus in Vero E6 and A549/ACE2 cells with low micromolar EC_{50} values. Therefore, the current study encourages further preclinical testing of bepridil in animal models, paving the way to its clinical use against COVID-19.

3.2 Results and discussions

Jin *et al.* released the first crystal structure of M^{Pro} on Feb 5th, 2020.[137, 142] We chose this structure (the pdb entry 6lu7) as the basis for our initial docking study. M^{Pro} has a very large active site that consists of several smaller pockets for the recognition of amino acid residues in its protein substrates. Three pockets that bind the P1, P2, and P4 residues in a protein substrate potentially interact with aromatic and large hydrophobic moieties.[143] Although the P1' residue in a protein substrate is a small residue such as glycine or serine, previous studies based on the same functional enzyme from SARS-CoV-1 showed that an aromatic moiety can occupy the site that originally bind the P1' and P2' residues in a substrate.[144] Based on this analysis of the M^{Pro} structure, we selected 55 FDA/EMA-approved small molecule medicines that have several aromatic or large hydrophobic moieties inter-connected and did a docking analysis of their binding to M^{Pro}. Some of the small molecule medicines used in our docking study were previously reported in other computational studies.[138-141] Autodock was the program we adopted for the docking analysis.[145] The covalent ligand and non-bonded small molecules in the structure of 6lu7 was removed to prepare the protein structure for docking. Four residues His41, Met49, Asnl42, and Glnl89 that have shown conformational variations in the SARS-CoV-1 enzyme were set flexible during the docking process. We carried out a genetic algorithm method with 100 runs to dock each small molecule medicine to the enzyme. We collected the lowest binding energy from the total 100 runs for each small molecule medicine and summarized them in Table 4.

Table 4 Docking results of small molecule medicines (Compounds whose IC_{50} values were tested are asterisked)

Name	$\Delta G_{binding}$ (kcal/mol)	Name	$\Delta G_{binding}$ (kcal/mol)
Rimonabant*	-11.23	Bepridil*	-8.31
Tipranavir*	-10.74	Isoconazole	-8.15
Ebastine*	-10.62	Econazole	-8.14
Saquinavir*	-10.37	Eluxadoline	-8.12
Zopiclone*	-10.10	(R)-Butoconazole	-8.11
Pimozide*	-10.01	(S)-Butoconazole	-8.10
Pirenzepine*	-9.94	Atazanavir	-8.08
Nelfinavir*	-9.67	Cetirizine	-8.01
Doxapram*	-9.55	Efinaconazole	-8.01
Oxiconazole*	-9.18	Amprenavir	-7.99
Indinavir*	-9.13	Hydroxyzine	-7.99
Sertindole*	-9.04	(R)-Tioconazole	-7.98
Metixene*	-9.01	(R)-Carbinoxamine	-7.96
Fexofenadine*	-8.95	Armodafinil	-7.90
Lopinavir*	-8.91	Desipramine	-7.84
Sertaconazole*	-8.87	Ritonavir	-7.74
Reboxetine*	-8.86	Atomoxetine	-7.73
Ketoconazole*	-8.85	Sulconazole	-7.69
Duloxetine*	-8.79	Clotrimazole	-7.67
Isavuconazole*	-8.77	Dipyridamole	-7.67
Lemborexant*	-8.75	Phentolamine	-7.61
Oxyphencyclimine*	-8.74	(S)-Tioconazole	-7.48
Darunavir*	-8.72	Doxylamine	-7.33
Trihexphenidyl*	-8.72	(S)-Carbinoxamine	-7.21
Pimavanserin*	-8.69	Antazoline	-6.86
Clotiapine*	-8.57	Voriconazole	-6.76
Itraconazole*	-8.44	Fluconazole	-6.41
Clemastine*	-8.36		

Among all 55 small molecule drugs that we used in the docking study, 29 showed a binding energy lower than -8.3 kcal/mol. We chose these molecules (Figure 19) to do further experimental characterizations.

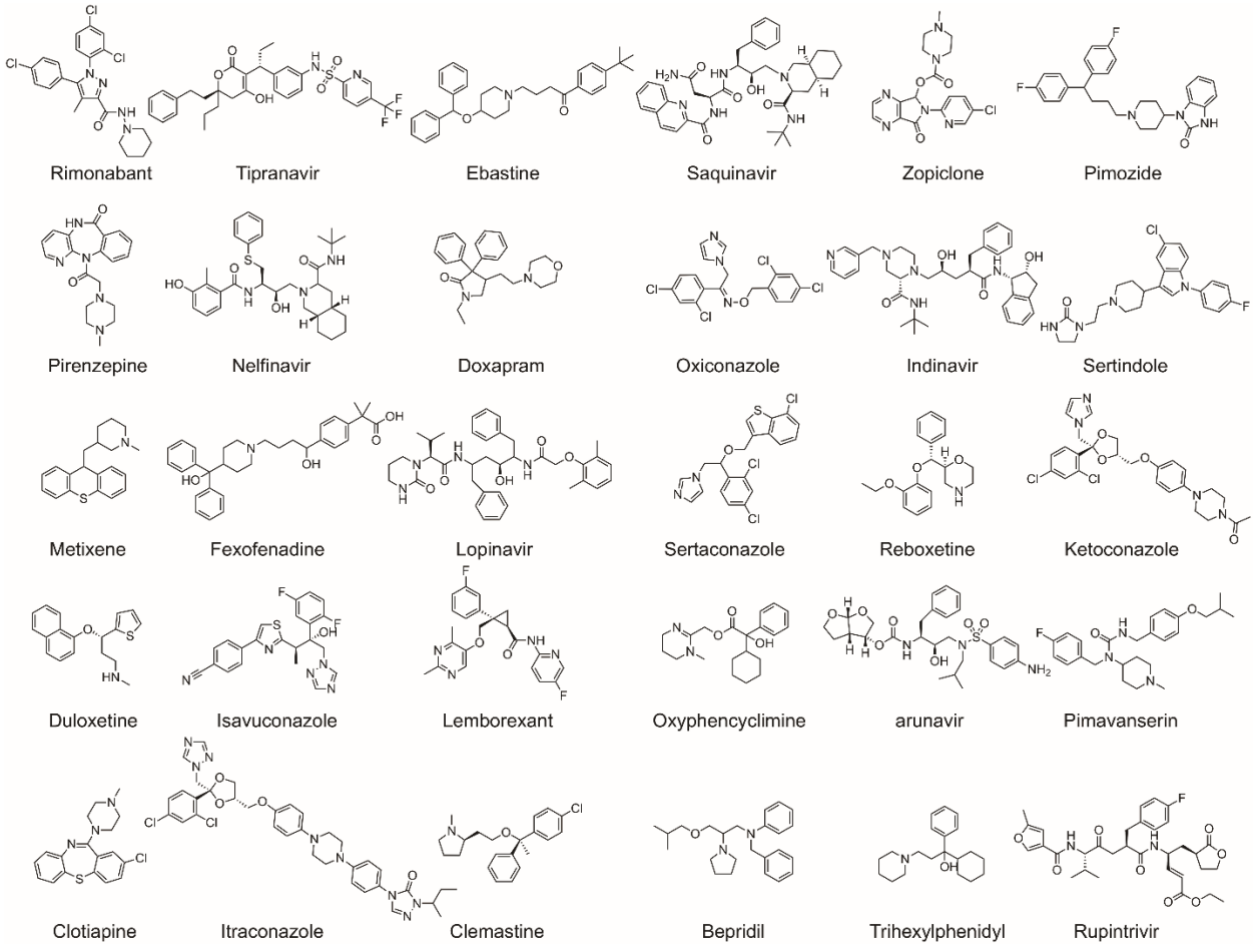


Figure 21 Structure of 29 FDA/EMA- approved medicines and rupintrivir whose IC_{50} values in inhibiting M^{Pro} were determined in this study

To express M^{Pro} for experimental characterizations of 29 selected small molecule medicines, we fused the M^{Pro} gene to a superfolder green fluorescent protein (sfGFP) gene and a 6xHis tag at its 5' and 3' ends respectively in a pBAD-sfGFP plasmid[146] that we used previously in the lab. SfGFP is known to stabilize proteins when it is genetically fused to them.[147] We designed a TEV protease cleavage site between sfGFP and M^{Pro} for the TEV-

catalyzed proteolytic release of M^{Pro} from sfGFP after we expressed and purified the fusion protein. We placed the 6xHis tag right after the M^{Pro} *C*-terminus. The addition of this tag was for straightforward purification with Ni-NTA resins. We expected that the TEV protease cleavage of sfGFP would activate M^{Pro} to cleave the *C*-terminal 6xHis tag so that a finally intact M^{Pro} protein would be obtained. We carried out the expression in *E. coli* TOP10 cells. To our surprise, after expression there was a minimal amount of the fusion protein that we were able to purify. The analysis of the cell lysate showed clearly the cleavage of a substantial amount of M^{Pro} from sfGFP. Since we were not able to enrich the cleaved M^{Pro} using Ni-NTA resins, the *C*-terminal 6xHis tag was apparently cleaved as well. TEV protease is a cysteine protease that cleaves after the Gln residue in the sequence Glu-Asn-Leu-Tyr-Phe-Gln-(Gly/Ser).[148] M^{Pro} is known to cleave the sequence Thr-Val-Leu-Gln-(Gly/Ser).[149, 150] The two cleavage sites share a same P1 residue. It was evident in our expression work that M^{Pro} is able to efficiently cleave the TEV protease cutting site to maturate inside *E. coli* cells. According to a peptide library screening study, it is likely that M^{Pro} has a substrate promiscuity higher than what we have learned from the SARS-CoV-1 enzyme. In this study, activities of SARS-CoV-1 M^{pro} and SARS-CoV-2 M^{pro} enzymes were tested against a combinatorial substrate library. The results showed that both enzymes share a significant similarity in substrate specificity but SARS-CoV-2 M^{pro} tolerates unnatural hydrophobic residues at the P2 position more than SARS-CoV-1 M^{pro}. To purify the cleaved and maturated M^{Pro}, we used ammonium sulfate to precipitate it from the cell lysate and then used the ion exchange and size exclusion chromatography to isolate it to more than 95% purity. We designed and synthesized a fluorogenic

coumarin-based hexapeptide substrate (Sub1) and a FRET-based decapeptide substrate (Sub2) and acquired a commercial FRET-based tetradecapeptide substrate (Sub3) (Fig. 20A). The test of enzyme activities on the three substrates indicated that the enzyme had low activity toward Sub1 (Fig. 20B) and its activity on Sub3 was higher than that on Sub2 (Fig. 20C) under our assay conditions. We subsequently used Sub3 in all the following inhibition analysis. To identify an optimal enzyme concentration for use in our inhibition analysis, we tested activities of different concentrations of M^{Pro} on 10 μM Sub3, the detected catalytic rate of the Sub3 cleavage was not proportional to the enzyme concentration (Fig. 20D). When the enzyme concentration decreased from 50 nM to 10 nM, the Sub3 cleavage rate dropped roughly proportionally to the square of the concentration decrease, characteristics of second-order kinetics. This observation supports previous claims that the enzyme needs to dimerize in order to be active.[137] In all the following assays, 50 nM M^{Pro} and 10 μM Sub3 were used throughout.

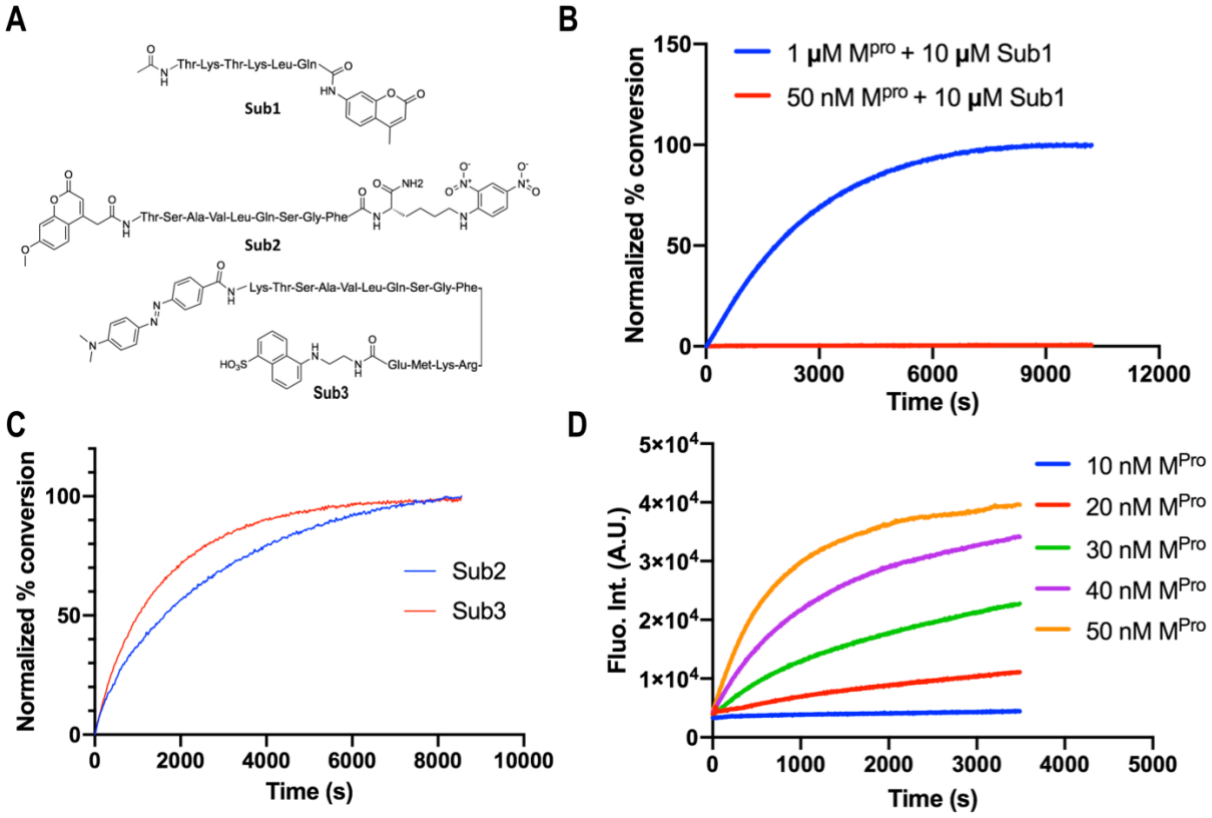


Figure 22 Activity of M^{Pro}. (A) The structure of three substrates. (B) Activity of 50nm M^{Pro} on 10μM Sub1. (C) Activity of 50nM M^{Pro} on 10μM Sub2 and Sub3. The fluorescence signals are normalized for easy comparison. (D) Activity of different concentration of M^{Pro} on 10μM Sub3.

We purchased all 29 small molecule medicines from commercial providers without further purification and characterization. Rupintrivir is a previously developed 3C protease inhibitor.[151] It has a key lactone side chain in the P1 residue that has a demonstrated role in tight binding to 3C and 3C-like proteases. Since it has been an investigational antiviral, we purchased it as well with a hope that it could be a potent inhibitor of M^{Pro}. We dissolved

most purchased small molecule medicines in DMSO to make 5 mM stock solutions and proceeded to use these stock solutions to test inhibition on M^{Pro}. Except itraconazole that has low solubility in DMSO, all tested small molecule medicines were diluted to a 1 mM final concentration in the inhibition assay conditions. We maintained 20% DMSO in the final assay condition to prevent small molecule medicines from precipitation. The activity of M^{Pro} in 20% DMSO was 29% lower than that in a regular buffer (Figure 21) but satisfied our assay requirement which is to have a fluorescence signal strong enough to produce straight line for reliable and reproducible calculation of initial slopes.

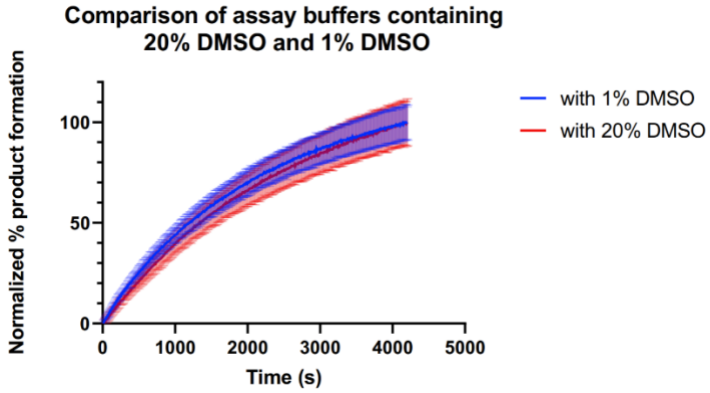


Figure 23 Comparison of assay buffers containing 20% DMSO and 1% DMSO

An M^{Pro} activity assay in the absence of a small molecule medicine was set up as a comparison. Triplicate repeats were carried out for all tested small molecules and the control. The results presented in Figure 22 displayed two easily discernable characteristics. First, about half of the tested compounds showed strong inhibition of M^{Pro} at the 1 mM concentration level (itraconazole was at 0.14 mM due to its low solubility in DMSO), supporting the practical use of a docking method in guiding the drug repurposing

research of COVID-19. Second, several small molecule medicines including fexofenadine, indinavir, pirenzepine, reboxetine, and doxapram clearly activated M^{Pro} (> 15%). This was to the contrary of what the docking program predicted. This observation strongly suggests that frontline clinicians need to exhibit caution in repurposing medicines for COVID-19 patients before they are thoroughly investigated on influencing the SARS-CoV-2 biology. A not-well-understood drug might deteriorate the already devastating symptoms in COVID-19 patients. Although it is not the focus of the current study, the observation that M^{Pro} can be activated by existing drugs needs to be further investigated.

We selected 17 small molecule medicines and rupintrivir that displayed strong inhibition of M^{Pro} to conduct further characterizations of their IC_{50} values in inhibiting M^{Pro} by varying the small molecule concentration from 1 μM to 10 mM. Results collectively presented in Figure 15 identifies that of the 18 tested compounds, 7 had an IC_{50} value below 100 μM. These include pimozide, ebastine, rupintrivir, bepridil, sertaconazole, rimonabant, and oxiconazole.

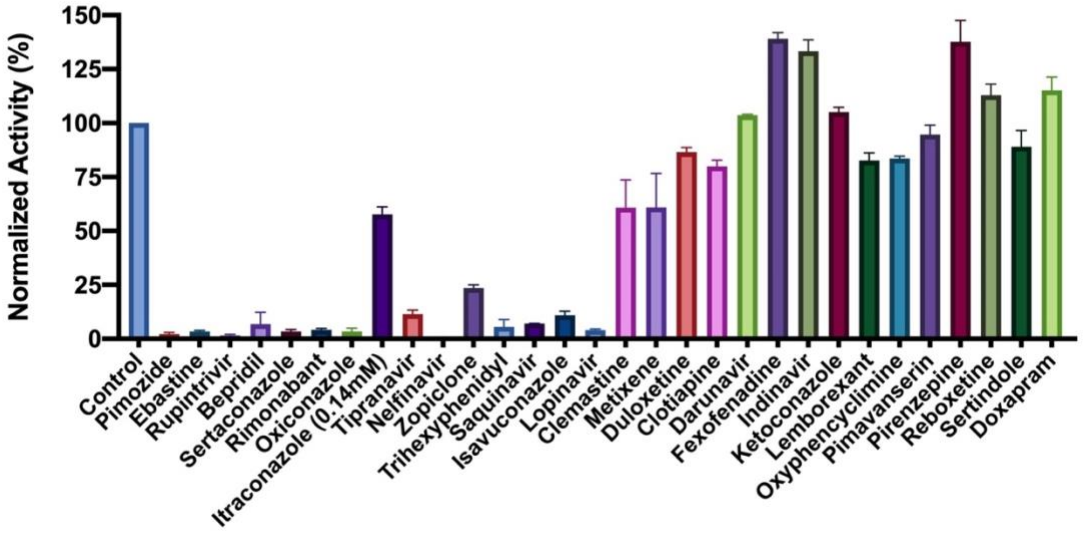


Figure 24 Initial screening of MPro inhibition by 29 FDA/EMA-approved medicines and rupintrivir. 1mM (0.14 mM for Itraconazole due to its low solubility in DMSO) was used for each inhibitor to perform the inhibition assay. Fluorescence intensity was normalized with respect to the control that had no small molecule provided. Triplicate experiments were performed for each compound, and the value was presented as mean ± standard error (SE).

There is no strong correlation between calculated binding energies and determined IC_{50} values. This discrepancy can be explained by limited factors that were involved in the calculation. Pimozide, ebastine, and bepridil were the three most potent FDA/EMA-approved medicines with IC_{50} values as 42 ± 2, 57 ± 12 and 72 ± 12 μM, respectively. Although rupintrivir is a covalent inhibitor that was developed specifically for 3C and 3C-like proteases, its IC_{50} value (68 ± 7 μM) is higher than that of pimozide and ebastine. The

relatively low activity of rupintrivir in inhibiting M^{Pro} might be due to the change of the amide bond between the P2 and P3 residues to an methyleneketone. This conversion served to increase the serum stability of rupintrivir, but has likely eliminated a key hydrogen bonding interaction with M^{Pro}.[137, 142] The repurposing of HIV medicines for the treatment of COVID-19, particularly those targeting HIV1 protease, has been area of much attention. In fact, the cocktail of lopinavir and ritonavir was previously tested in China for the treatment of COVID-19.[152] The IC_{50} value of lopinavir in inhibiting M^{Pro} is about 500 μM, which possibly explains why this anti-HIV viral cocktail demonstrated no significant benefit for treating patients. Nelfinavir was previously shown having high potency in inhibiting the entry of SARS-CoV-2 into mammalian cell hosts. A cell-based study in inhibiting the SARS-CoV-2 entry indicated a 1 μM EC_{50} value.[153, 154] However, our IC_{50} determination against M^{Pro} resulted in a value of 234 ± 5 μM, highlighting that nelfinavir likely inhibits another key SARS-CoV-2 enzyme or protein or interferes with key cellular processes that are required for the SARS-CoV-2 entry into host cells. These possibilities need to be studied further.

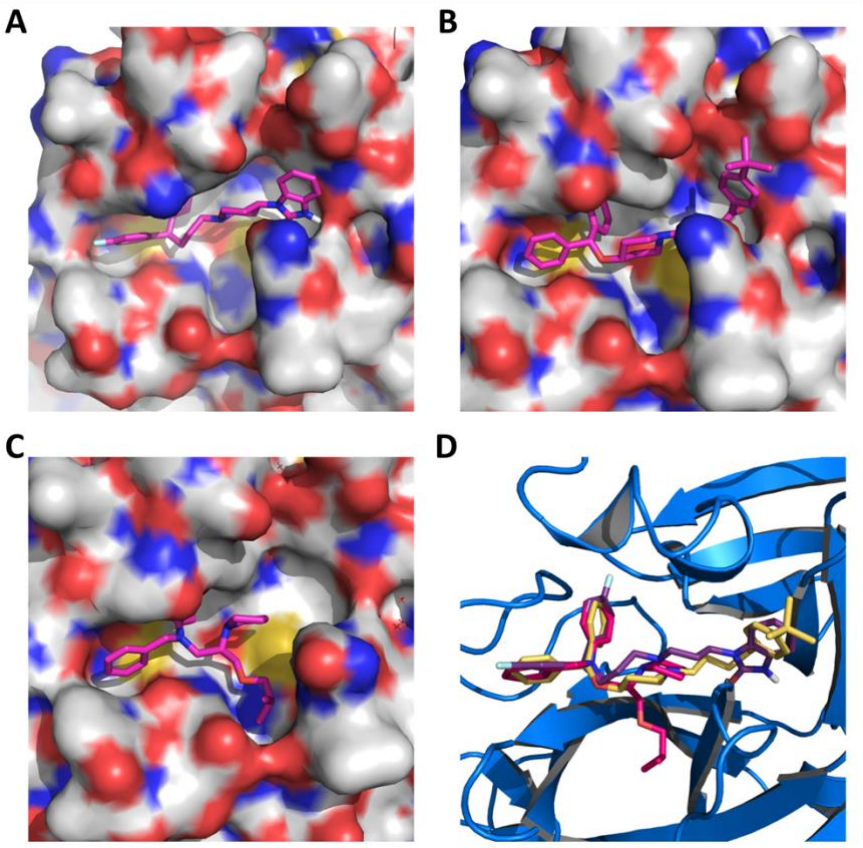


Figure 25 Pimozide (A), ebastine (B) bepridil (C) and their overlay (D) in the active site of M^{Pro}. The protein surface topography in A, B, and C is presented to show the concave active site.

Structurally the two most potent medicines pimozide and ebastine share a same diphenylmethyl moiety. A spatially similar structure moiety N-phenyl-N-benzylamine exists in bepridil. Our docking results suggested a same binding mode for this similar structure moiety in all three drugs (Figure 19). The two aromatic rings occupy the enzyme pockets that associate with the P2 and P4 residues in a substrate.

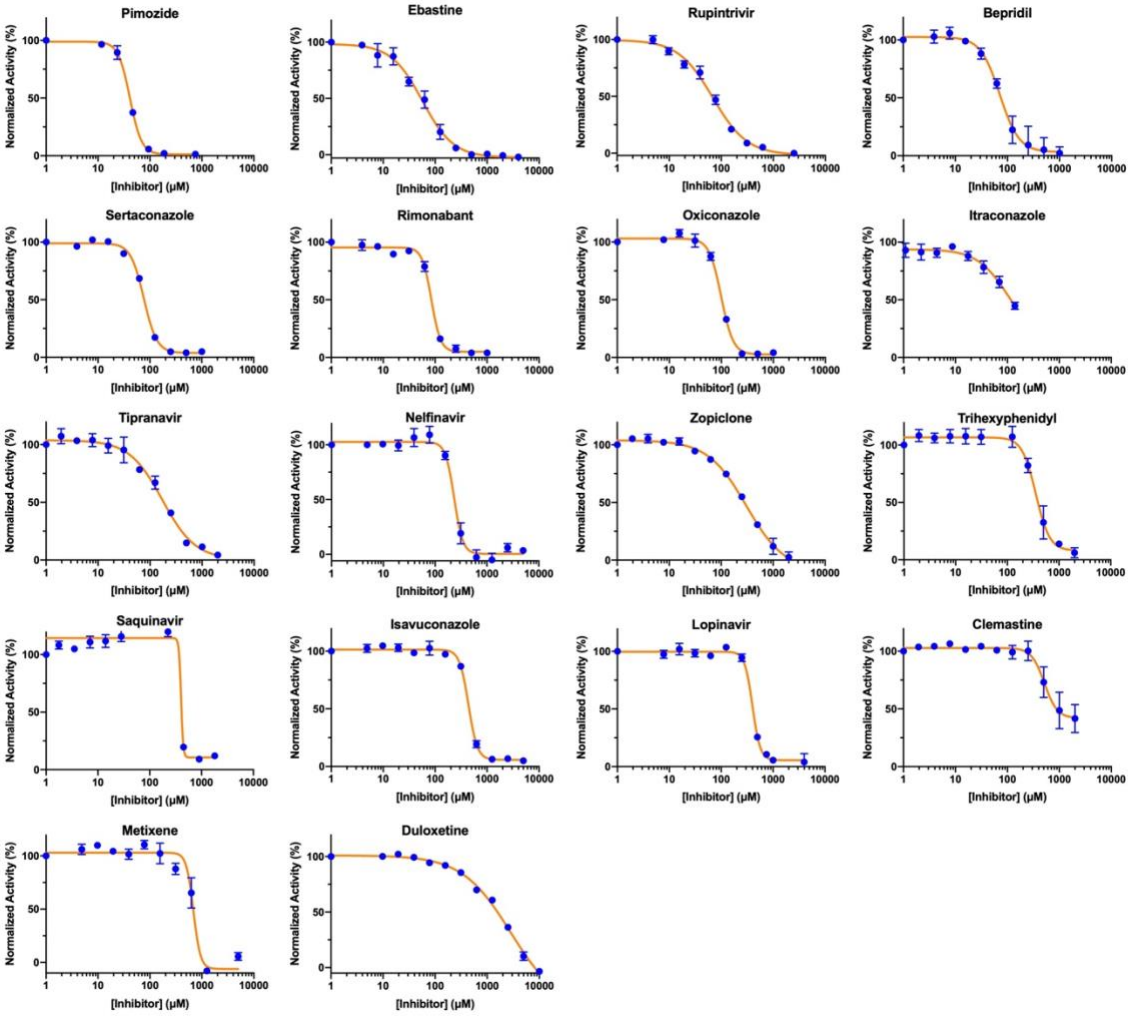


Figure 26 IC50 assays for 18 small molecule medicines on their inhibition of M^{Pro}. Triplicate experiments were performed for each compound and the IC_{50} value was presented as mean ± standard error (SE). GraphPad Prism 8.0 was used to perform data analysis.

This observation is in line with a crystallographic study that showed two aromatic rings with a single methylene linker bound to the active site of the SARS-CoV-1 enzyme.[155] We

believe that the inclusion of the diphenylmethyl moiety in structure-activity relationship studies of M^{Pro}-targeting ligands will likely contribute to the identification of both potent and high cell-permeable M^{Pro} inhibitors. Figure 22 also revealed large variations in Hill coefficients of IC_{50} curves for different small molecule medicines (IC_{50} values and Hill coefficients are summarized in Table 5).

Table 5 IC_{50} and Hill coefficient values of 18 identified inhibitors

Name	IC_{50} (μM)	Hill Slope
Pimozide	42 ± 2	3.1 ± 0.4
Ebastine	57 ± 12	1.5 ± 0.2
Rupintrivir	68 ± 7	1.4 ± 0.2
Bepridil	72 ± 3	2.9 ± 1.0
Sertaconazole	76 ± 2	3.5 ± 0.2
Rimonabant	85 ± 3	5.0 ± 0.4
Oxiconazole	99 ± 6	3.8 ± 0.4
Itraconazole	111 ± 35	1.6 ± 0.2
Tipranavir	180 ± 20	1.4 ± 0.2
Nelfinavir	234 ± 15	5.4 ± 1.0
Zopiclone	349 ± 77	1.2 ± 0.2
Trihexyphenidyl	370 ± 53	8.9 ± 6.4
Saquinavir	411 ± 6	26.8 ± 2.6
Isavuconazole	438 ± 11	5.2 ± 0.7
Lopinavir	486 ± 2	29.9 ± 2.4
Clemastine	497 ± 148	11.2 ± 7.3
Metixene	635 ± 43	8.7 ± 5
Duloxetine	3047 ± 634	0.93 ± 0.07

Duloxetine and zopiclone gave the two highest Hill coefficients with a gradual M^{Pro} activity decrease over an increasing inhibitor concentration. On the contrary, saquinavir and lopinavir yielded lowest Hill coefficients with highly steep IC_{50} curves. There are

three possible explanations for the large discrepancy in Hill coefficients. It could be attributed to different solubility of tested compounds. It is possible that a high DMSO percentage and a relatively high inhibitor concentration created phase transition for some inhibitors.[156] A high Hill coefficient may also be due to different ligand-to-enzyme ratios when tested compounds bind to M^{Pro}. An additionally possible reason is the co-existence of the M^{Pro} dimer and monomer in the assay conditions. A previous report showed a K_d value of the M^{Pro} dimerization as 2.5 μM.[143] In theory, the catalytically active dimer species was present at a very low concentration in our assay conditions, leaving the catalytically inactive monomer species as the major form of M^{Pro}. In this situation, the inhibitors that preferentially bind to the M^{Pro} dimer and the inhibitors that have a higher affinity to the M^{Pro} monomer might yield different Hill coefficients. Although there is no report on the physiological concentration of M^{Pro} in infected cells, it is unlikely that it can reach 1 mM (34 ppm). Even at 1 mM, the majority of M^{Pro} is in its inactive monomeric form. Therefore, we believe that our assay conditions mimic physiological states of M^{Pro}.

The endocytic pathway has been proposed as a key step for the SARS-CoV-2 entry into host cells.[157], making strategies that disrupt this process attractive therapeutic candidates for COVID-19. Based on this concept, hydroxychloroquine that has an ability to raise endosomal pH [158, 159] has been clinically investigated for treating COVID-19 albeit with close to no effects.[160-162] From chemistry point of view, our three lead compounds pimozide, ebastine, and bepridil share a similarity. They are all basic small molecules that can potentially raise endosomal pH.[163] Among the three drugs, bepridil can be very

interesting because it previously provided 100% protection from Ebola virus infections in mice at a dose of 12 mg/kg.[164] Bepridil is a calcium channel blocker with a significant anti-anginal activity. For patients with chronic stable angina, recommended daily dose of bepridil is 200-400 mg.[155] Mice administered with a bepridil dose as high as 300 mg/kg/day did not show alteration in mating behavior and reproductive performance, indicating that bepridil has very low toxicity.[165] Moreover, a previous study showed that bepridil can increase the pH of acidic endosomes.[166] Administration of a high dose of bepridil may have dual functions to slow down the virus replication in host cells by both inhibiting M^{Pro} and raising the pH of endosomes. To demonstrate this prospect, we conducted a live virus-based microneutralization assay to evaluate efficacy of pimozide, ebastine and bepridil in their inhibition of SARS-CoV-2 infection in a kidney epithelial cell line isolated from African green monkey (Vero E6) and human A549 cells stably transduced with human ACE2 viral receptor (A549/ACE2). We tested three medicines in a concentration range from 0.78 to 200 mM. CPE was clearly observable for pimozide and ebastine at all tested concentrations. However, bepridil prevented completely the SARS-CoV-2-induced CPE in Vero E6 and A549/ACE2 cells when its concentration reached 6.25 μM (Table 4) *with* n*o* signs of cytotoxicity observed under the microscopy.

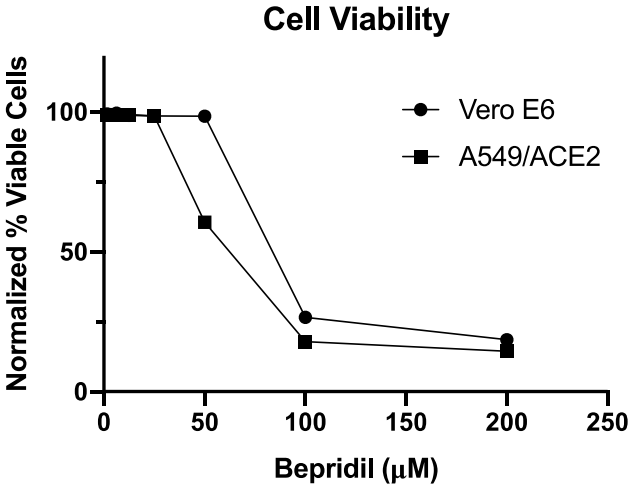


Figure 27 Viability of Vero E6 and A549/ACE2 cells at different concentrations of bepridil.

A separate cell toxicity assay showed that bepridil was not toxic to Vero E6 and A549/ACE2 cells until its concentration reached above 25 and 50 μM, respectively (Figure 25). In order to characterize bepridil EC_{50} values in inhibiting SARS-CoV-2 in the two cell lines, both cells were treated with bepridil at different concentrations and infected with SARS-CoV-2 at an MOI of 0.5. Infected cells were then cultured for three days for Vero E6 and four days for A549/ACE2 cells before assessing the yields of infectious progeny virus. Based on the efficacy to inhibit SARS-CoV-2 infection, EC_{50} values were estimated to be 0.86 and 0.46 μM in Vero E6 and A549/ACE2 cells, respectively (Figure 26).

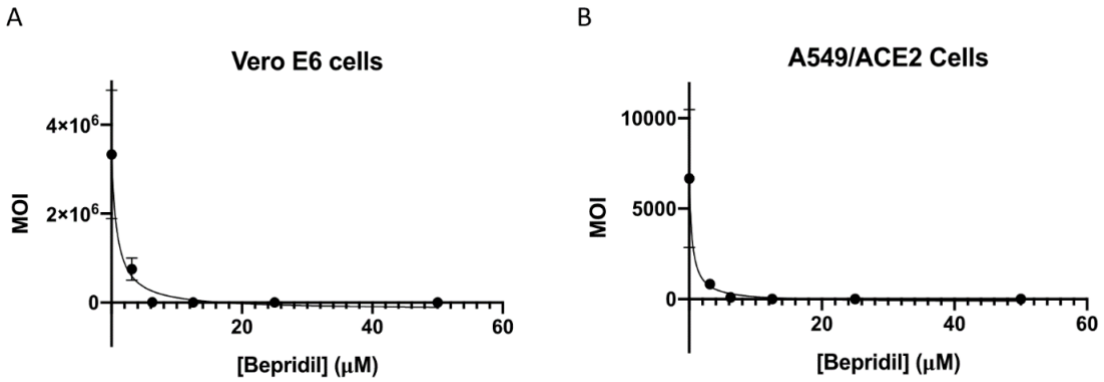


Figure 28 The SARS-CoV-2 inhibition by bepridil in (A0 Vero E6 cells and (B) A549/ACE2 cells. Cells were incubated with different concentrations of bepridil and then infected with 0.5 MOI (multiplicity of infection) of SARS-CoV-2.

These values are similar to Remdesivir`s reported EC_{50} value of 0.77 μM in Vero E6 cells.[167] The strong inhibition of SARS-CoV-2-induced CPE in Vero E6 and A549 cells by bepridil at a concentration much lower than its IC_{50} value for inhibiting M^{Pro} is likely due to the aforementioned dual functions or other cellular effects of bepridil. In patients, bepridil can reach a state C_{max} as 3.72 mM.[168] This concentration is effective in inhibiting SARS-CoV-2 based on our virus microneutralization analysis. Collectively, our results indicate that bepridil is an effective medicine in preventing SARS-CoV-2 from entry and replication in mammalian cell hosts. Therefore, we urge the consideration of clinical tests of bepridil in the treatment of COVID-19.

Bepridil was voluntarily withdrawn from US market in 2004 due to its side effects such as QT prolongation.[169] It has also been implicated that its usage might cause ventricular arrhythmia.[170] Despite those potential drawbacks, bepridil is still marketed in Japan, China, and France. Several factors such as its *in vivo* efficacy against SARS-CoV-2 and cardiovascular effects of COVID-19 need to be assessed before attempting to explore clinical benefits of using it in COVID-19 patients.

3.3 Conclusion

Guided by a computational docking analysis, we experimentally characterized about 30 FDA/EMA-approved drugs on their inhibition of the essential M^{Pro} enzyme of the COVID-19 pathogen SARS-CoV-2. From the study, we identified six FDA/EMA-approved drugs that can potently inhibit M^{Pro} with an IC_{50} value lower than 100 μM. One medicine bepridil exhibited strong inhibition of SARS-CoV-2 from entry and replication inside Vero E6 and A549 cells at a low micromolar concentration. Although the IC_{50} value for Bepridil (72 μM) is relatively high, the better indicative of the potency of an antiviral drug candidate is its potency against live virus. Given that bepridil has been previously demonstrated to show efficacy in Ebola infected mice, we urge a serious consideration of its clinical tests in treating COVID-19. Our current study indicates that there is a large amount of FDA/EMA-approved drug space open for exploration that could hold promise for repurposing existing drugs to target COVID-19. Performing screening studies on different SARS-CoV-2 protein targets are necessary to uncover existing medicines that may be combined for cocktail treatments of COVID-19. More explorations in this direction are imperative.

3.4 Experimental Procedures

Chemicals

We purchased econazole nitrate, duloxetine hydrochloride, doxapram hydrochloride monohydrate, clemastine fumarate salt, sertaconazole nitrate, isavuconazole, rupintrivir, and zopiclone from Sigma-Aldrich, pimavanserin, trihexyphenidyl hydrochloride, reboxetine mesylate, sertindole, bepridil hydrochloride, darunavir, nelfinavir mesylate, indinavir sulfate, lopinavir, tipranavir, saquinavir, pirenzepine hydrochloride, oxiconazole nitrate, pimozide, and rimonabant from Cayman Chemicals, ebastine and itraconazole from Alfa Aesar, metixene hydrochloride hydrate and lemborexant from MedChemExpress, fexofenadine hydrochloride from TCI Chemicals, ketoconazole from Acros Organics, clotiapine from Enzo Life Sciences, and oxyphencyclimine from Boc Sciences. We acquired Sub3 with the sequence as DABCYL-Lys-Thr-Ser-Ala-Val-Leu-Gln-Ser-Gly-Phe-Arg-Lys-Met-Glu-EDANS from Bachem Inc.

3.4.1 Docking

Autodock 4 was used for all docking analysis. For each small molecule, the genetic algorithm-based calculation was carried out for 100 runs with each run having a maximal number of evaluations as 2,500,000.

3.4.2 M^{pro} Expression and Purification

We constructed the plasmid pBAD-sfGFP-M^{pro} from pBAD-sfGFP. The M^{Pro} gene was inserted between DNA sequences that coded sfGFP and 6xHis. The overall sfGFP-M^{Pro}-6xHis fusion gene was under control of a pBAD promoter. The antibiotic selection marker was ampicillin. To express sfGFP-M^{Pro}-6xHis, E. coli TOP10 cells were transformed with pBAD-sfGFP-M^{Pro}. A single colony was picked and grew in 5 mL LB medium with 100 µg/mL ampicillin overnight. In the next day, we inoculated this starting culture into 5 L 2xYT medium with 100 µg/mL ampicillin in 5 separate flasks at 37 °C. When the OD reached to 0.6, we added L-arabinose (working concentration as 0.2%) to each flask to induce protein expression at 37 °C for 4 h. Then, the cells were pelleted at 4000 rpm at 4 °C, washed with cold PBS and stored in -80 °C until purification. To purify the expressed protein, we re-suspended frozen cells in a 125 mL buffer containing Tris pH 7.5, 2.5 mM DTT, and 1.25 mg lysozyme. We sonicated resuspended cells using a Branson 250W sonicator with 1 second on, 4 second off, and a total 5 min 60% power output in two rounds. After sonication, we spun down the cellular debris at 16000 rpm for 30 min at 4 °C. We collected the supernatant and recorded the volume. The whole cell lysate analysis showed almost all of the fusion protein was hydrolyzed to two separate proteins sfGFP and M^{Pro}. We were able to obtain an insignificant amount of M^{Pro} when Ni-NTA resins were used for purification. Therefore, we did ammonium sulfate precipitation of the whole cell lysate method. This was done by the addition of a saturated ammonium sulfate solution at 0 °C. We collected the fraction between 30% and 40% of ammonium sulfate.

We dissolved the collected fraction in buffer A (20 mM Tris, 10 mM NaCl, and 1 mM DTT at pH 8.0) and dialyzed the obtained solution against the same buffer to remove ammonium sulfate. Then, we subjected this solution to anion exchange column chromatography using Q sepharose resins. We eluted proteins from the Q sepharose column by applying a gradient with increasing the concentration of buffer B (20 mM Tris, 1 M NaCl, and 1 mM DTT at pH 8.0). We concentrated the eluted fractions that contained M^{Pro} and subject the concentered solution to size exclusion chromatography using a HiPrep 16/60 Sephacryl S-100 HR column with a mobile phase containing 10 mM sodium phosphate, 10 mM NaCl, 0.5 mM EDTA and 1 mM DTT at pH 7.8. The final yield of the purified enzyme was 1 mg/L with respect to the original expression medium volume. We determined the concentration of the finally purified M^{pro} using the Pierce™ 660nm protein assay and aliquoted 10 µM M^{Pro} in the size exclusion chromatography buffer for storage in -80 °C.

3.4.3 The synthesis of Sub1

We loaded the first amino acid (0.5 mmol, 2 equiv.) manually on chlorotrityl chloride resin (0.52 mmol/g loading) on a 0.25 mmol scale by the addition of DIPEA (3 equiv.). After addition of the first amino acid, automated Fmoc-based solid phases synthesis was performed using a Liberty Blue automated peptide synthesizer. Deprotection of the Fmoc group was carried out with 20% piperidine/DMF. Coupling was done with a Fmoc-protected amino acid (0.75 mmol, 3.0 equiv.) and the coupling reagent HATU (0.9 mmol,

3.6 equiv.) and DIEA in NMP (1 mmol, 4.0 equiv.). The final amino acid was capped by the addition of 25% acetic anhydride (v/v) in DMF and DIEA (0.2mmol, 2.0 equiv.). Coumarin coupling was performed in anhydrous THF using T3P in EtOAc (50% w/v) (3.0 equiv.), DIEA (3 equiv.) and 7-amino-4methyl-coumarin (0.8 equiv.) and mixed for 16 h. The solvent was removed and the peptide was dissolved in DCM and washed with H_2O (4x) followed by HCl (2x) and brine (1x). The organic layer was dried with Na_2SO_4, filtered and concentrated. Global deprotection was then carried out using triisopropylsilane (5%) and trifluoroacetic acid (30%) v/v in DCM and mixed for 2-3 h to result in the crude substrate. The peptide was then purified by semi-preparative HPLC and the fractions containing pure product were pooled, concentrated, and analyzed by LC-MS for purity.

3.4.4 The synthesis of Sub2

We performed automated Fmoc-based solid phase synthesis on a Liberty Blue automated peptide synthesizer. Synthesis was conducted on a 0.1 mmol scale with Fmoc Rink amide MBHA resin (0.52 mmol/g loading) and 3 equiv. of protected amino acids. Deprotection of the Fmoc group was carried out with 20% piperidine/DMF. Coupling was done using the desired Fmoc-protected amino acid (0.2 mmol, 2.0 equiv.), coupling reagent Oxyma (0.4 mmol, 4.0 equiv.) and DIC (0.4 mmol, 4.0 equiv.). After the final amino acid had been coupled on, the resin was washed trice with DMF and DCM. Cleavage from the resin was performed using trifluoroacetic acid (95%), triisopropylsilane (2.5%) and water

(2.5%) with agitation for 4 h. The peptide was drained into cold methyl tert-butyl ether where it precipitated out. We centrifuged the precipitate, decanted mother liquor, dissolved the pellet in DMF and then purified the peptide by LCMS.

3.4.5 Screening assay

5 mM stock solutions of medicines were prepared in DMSO. The final screening assay conditions contained 50 nM M^{pro}, 10 µM Sub3, and 1 mM medicine. We diluted enzyme stock and substrate stock solutions using a buffer containing 10 mM sodium phosphate, 10 mM NaCl, and 0.5 mM EDTA at pH 7.6 for reaching desired final concentrations. We ran the assay in triplicates. First, we added 30 µL of a 167 nM M^{pro} solution to each well in a 96-well plate and then provided 20 µL of 5 mM stock solutions of medicines in DMSO. After a brief shaking, we incubated the plate at 37°C for 30 min. Then we added 50 µL of a 20 µM Sub3 solution to initiate the activity analysis. The EDANS fluorescence with excitation at 336 nm and emission at 455 nm from the cleaved substrate was detected. We determined the fluorescence increasing slopes at the initial 5 min and normalized them with respect to the control that had no inhibitor provided.

3.4.6 Inhibition analysis

The final inhibition assay conditions contained 50 nM M^{Pro}, 10 µM Sub3, and a varying concentration of an inhibitor. Similar to screening assay, we diluted enzyme stock and

substrate stock solutions using a buffer containing 10 mM sodium phosphate, 10 mM NaCl, and 0.5 mM EDTA at pH 7.8 for reaching desired final concentrations. We ran the assay in triplicates. For the inhibition analysis, we added 30 µL of a 167 nM M^{Pro} solution to each well in a 96-well plate and then provided 20 µL of inhibitor solutions with varying concentrations in DMSO. After a brief shaking, we incubated the plate at 37°C for 30 min. Then we added 50 µL of a 20 µM Sub3 solution to initiate the activity analysis. We monitored the fluorescence signal and processed the initial slopes in the same way described in screening assay part. We used GraphPad 8.0 to analyze the data and used the [Inhibitor] vs. response - Variable slope (four parameters) fitting to determine the values of both IC_{50} and Hill coefficient.

3.4.7 SARS-CoV-2 inhibition by a cell-based assay

A slightly modified standard live virus-based microneutralization (MN) assay was used as previously described [171-174] to rapidly evaluate the drug efficacy against SARS-CoV-2 infection in Vero E6 and A549 cell culture. Briefly, confluent Vero E6 or A549/ACE2 cells grown in 96-wells microtiter plates were pre-treated with serially 2-folds diluted individual drugs in duplicate over eight concentrations for two hours before infection with ~100 and ~500, respectively, infectious SARS-CoV-2 particles in 100 µL EMEM supplemented with 2% FBS. Vero E6 and A549 cells treated with parallelly diluted dimethyl sulfoxide (DMSO) with or without virus were included as positive and negative controls, respectively. After cultivation at 37 °C for 3 days (Vero E6) or 4 days

(A549/ACE2), individual wells were observed under the microcopy for the status of virus-induced formation of CPE and cytotoxicity. The efficacy of individual drugs was calculated and expressed as the lowest concentration capable of completely preventing virus-induced CPE in 100% of the wells with no signs of cytotoxic effect. All compounds were dissolved in 100% DMSO as 10 mM stock solutions and diluted in culture media The toxicity to the treated cells was assessed by the Cell Cytotoxicity Assay kit (Abcam Cat#ab112118, Cambridge, MA) according to the manufacturer's protocol.

To quantify EC_{50} values, Vero E6 and A549/ACE2 cells grown in 24-well plates were pre-treated with serially 2-folds diluted drug for two hours before infection with SARS-CoV-2 at an MOI (multiplicity of infection) of 0.5 in 200 μL EMEM supplemented with 2% FBS. Cells treated with parallelly diluted DMSO with or without virus were included as positive and negative controls, respectively. After incubation for an hour with viral inoculum, cells were washed 3 times with EMEM, and cultivated with fresh medium for 3 days (Vero E6) or 4 days (A549/ACE2). Supernatants from infected cells were harvested for measuring the infectious virus titers by the tissue culture infective dose (*TCID50*) assay using Vero E6 cells. Briefly, 50 μL supernatants from infected cells were serially diluted (10 fold) in EMEM supplemented with 2% FBS; 100 μL of serially diluted samples were added to VeroE6 cells grown in 96 well plates and cultivated at 37°C for 3 days followed by observation under the microcopy for the status of virus-induced formation of CPE in individual wells. The titers were expressed as log TCID50/mL.

CHAPTER IV DRUG DESIGN: IMPROVING HYPOXIA ACTIVATED PRODRUG TRIGGERS

4.1 Hypoxia activated prodrugs Background

Tumor Hypoxia has been pursued as a drug target for over 30 years, typically by employing bioreductive prodrugs, collectively known as Hypoxia Activated Prodrugs (HAPs). These prodrugs are designed to be selective for low oxygen environments and innocuous in healthy, well oxygenated areas. While there have been many attempts to take advantage of this unique tumor environment, no HAP has been approved for use in humans as a prodrug due to reasons ranging from toxicity, lack of selectivity and poor efficacy. This failure demands a closer look into discovery and development of these compounds. The identification of hypoxia in tumors and its association with treatment resistance and poor prognosis is the initiating premise for the development of these small molecules. Clinical trials have incorporated oxygen needle electrode-based measurements in accessible tumors of the head and neck, cervix, and prostate. Similar assessments have been made using indirect assessments of hypoxia like nitroimidazole-based hypoxia probes, used in PET imaging or immunohistochemistry assays. Based on such assessments, the level of hypoxia in various tumors have displayed an incredible amount of diversity. This heterogenous character of different tumor environments highlights the need to differentiate patients whose hypoxic levels are such that would benefit from these compounds from those who would not.

HAPs by design are inactive prodrugs that require enzymatic activation to generate cytotoxic species with selectivity for hypoxic cells. The ideal HAP should possess the following properties: (1) the ability to penetrate from a blood vessel to hypoxic cells within its lifespan; (2) preferential activation in hypoxic conditions (<10mmHg) and not in normal tissues; (3) the released compound is lethal to hypoxic cells; (4) the released compound displays a 'by-standard effect'.[56]

Currently, there are five different classes of HAPs: nitro (hetero)-cyclic compounds, aromatic N-oxides, aliphatic N-oxides, quinone and transition metal complexes. Nitro-aromatic and Nitro imidazoles have been seen to show the most promise in clinical trials. Examples of recent HAPs in clinical trials are shown in Figure 18.

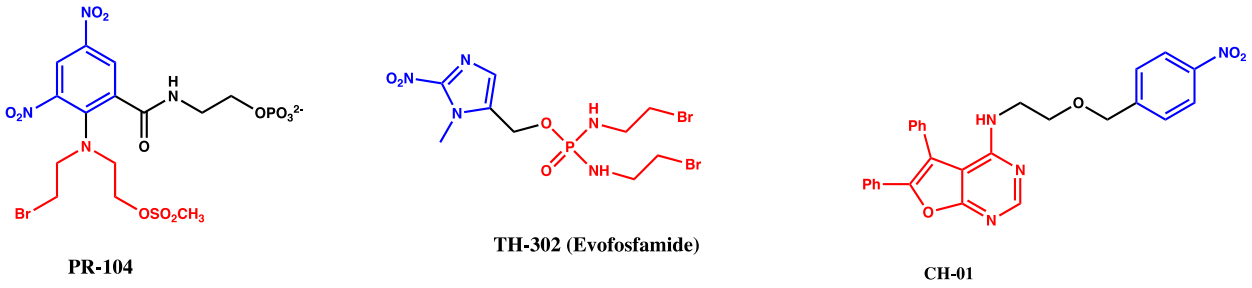


Figure 29 Recent HAPs in clinical trials; nitroaromatic (PR-104), nitro imidazole (TH-302) and nitroaromatic (CH-01).

While they vary in structure, every HAP is made of the same three components: a trigger to catalyze the release of the effector, a linker to deactivate the effector, and an effector (shown in blue black and red, respectively in Figure 28). Of these three components, the

trigger is the most critical. The design of the trigger determines the compound's activation and selectivity. Triggers must be reduced enzymatically- either by a one or two electron mechanism, most commonly carried out by Cytochrome p450 oxidoreductase(POR).[64] Currently, HAP triggers are designed based on reduction potential, focusing on facilitating the first single electron transfer to the nitro trigger. This has been believed to be the rate determining step of the reduction of these compounds.

4.2 Results and Discussion

4.2.1 Trigger Reduction and design

The Nitro functionality on Hypoxia activated prodrugs act as an electronic switch for these therapeutics. This switch is responsible for releasing the effector molecule in the correct environment and at an effective dose. To improve the activity, specificity and overall effectiveness of HAPs, a better understanding of the factors that influence the reactivity of the trigger is crucial. Many current HAPs are designed to specifically employ electron withdrawing (EW) groups around their aromatic ring, like PR-104 shown in figure 28. The desire to include these groups is to emphasize the nitrogen's electron deficiency and increase the compound's overall reduction potential. This is believed to increase the rate of enzymatic reduction and lower the cytotoxicity of the final metabolite.

A high reduction potential is broadly accepted to be the determining factor in the reduction of these compounds but with low relative correlation between rate and reduction potential. This poor understanding of reduction can be attributed to multiple factors. For one, the mechanism of the flavoenzymes that reduce these compounds biologically is poorly understood when it comes to their substrate specificity and mechanism. Being that they're (1) a ubiquitous enzymatic superfamily and that they're (2) role in breaking down xenobiotics makes them incredibly versatile. It is unclear how the enzyme reactivity is influenced by the electron-accepting potency of nitroaromatics and the efficiency of their interaction with the active center, i.e., the structural parameters.[175, 176] Another factor to be noted in regards to the rate of reduction is the reversibility of these nitro triggers. The nitro trigger has the ability to revert back to its' parent compound in the presence of oxygen, this makes the reduction potential; the compounds willingness to accept the initial electron, more influenced by its environment than its structure. PR-104 shown in figure 28, has a desirable reduction potential based on calculations but showed poor results in clinical trials due to its' off target interactions, losing its selectivity for hypoxic cells.

Previously it has been suggested that the steric crowding of nitro groups in polynitrobenzimidazoles increasing their torsion angle may facilitate their reduction by NQO1, by decreasing the distance of hydride transfer.[175] A previous study using large number of nitro compounds with a broad range of reduction potentials has demonstrated that an increase in the electron accepting potency of nitroaromatics increases their enzymatic reactivity[177], although the reactivity is also strongly influenced by the

structural features of the compounds. Oxygen-dependent bioreduction in mammals is thought to occur predominantly via a one-electron mechanism, but the propensity of a particular group to undergo bioreduction does not correlate directly with a compound's electrochemical reduction potential. In addition, the instability of the $ArNO_2^-$ radical in water makes measuring these values difficult. The reduction potentials $E(ArNO_2/ArNO_2^-)$ versus the normal hydrogen electrode (NHE) in water at pH 7 for commonly used nitroaryl bioreductive groups have been estimated[178] and are shown in Figure 29.

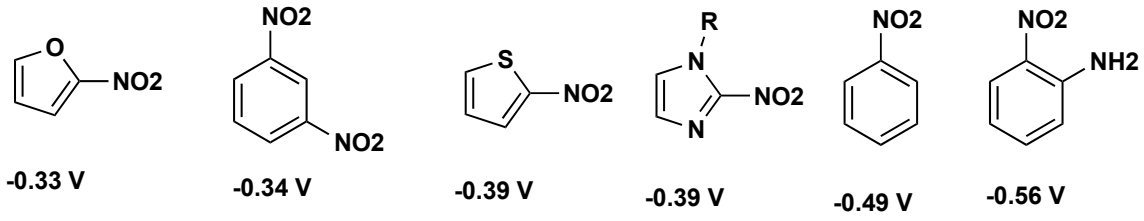
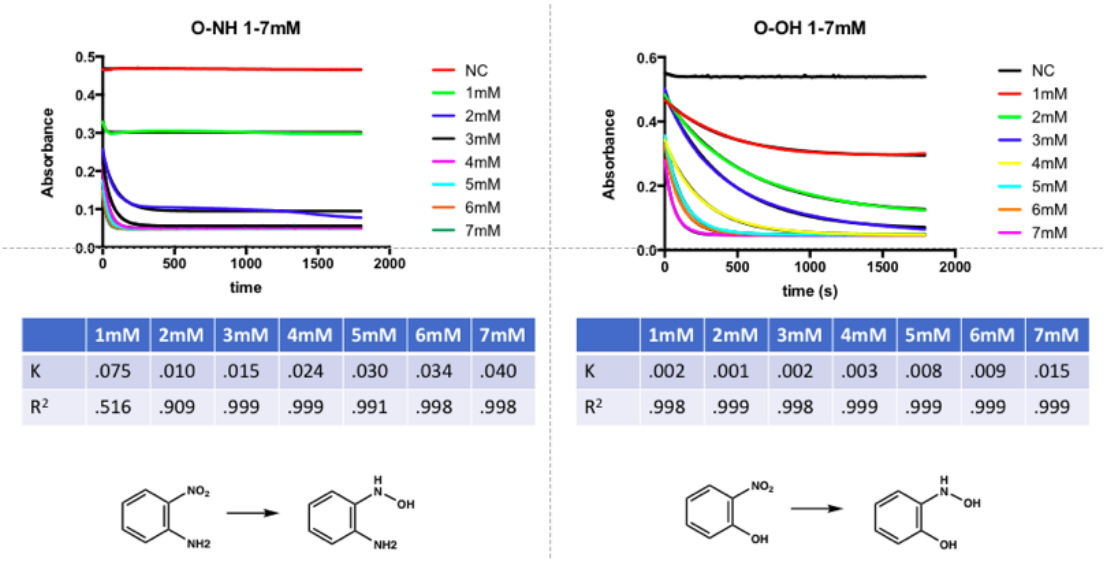


Figure 30 Estimated values of reduction potentials E(arNO2/ArNO2-) versus NHE in water at pH 7 for commonly used nitroaryl bioreductive systems.[178] R= alkyl or hydroxylalkyl.

The window of -400 to -200 mV is predicted to be suitable for the observation of hypoxic selectivity with nitroaromatic compounds, and anything outside of this window may be too negative to allow efficient enzymatic reduction.[179]

For our research, the preferred orientation of the Nitro group was investigated first. Nitro benzyl substituted compounds were chosen over nitroimidazoles due to their well-

established enzymatic reduction and their synthetic accessibility. In the electronic reduction of nitrobenzyl compounds, for the trigger to release an effector (undergo an elimination) the linker can be placed on either the ortho or para position to the Nitro group. Various Nitrobenzene derivatives with meta, ortho and para directors, both activating and de-activating, were reduced chemically and their rates were compared. The starting nitro benzene's absorbance was followed, measuring the loss of absorbance as it was reduced overtime in a plate reader placed inside a glove box to monitor the rate of the reaction. Example of spectra obtained from two selected triggers are shown below.



	1mM	2mM	3mM	4mM	5mM	6mM	7mM
K	.075	.010	.015	.024	.030	.034	.040
R^2	.516	.909	.999	.999	.991	.998	.998

	1mM	2mM	3mM	4mM	5mM	6mM	7mM
K	.002	.001	.002	.003	.008	.009	.015
R^2	.998	.999	.998	.999	.999	.999	.999

Figure 31 Loss of Absorbance measured by plate reader of Nitro aromatic triggers NAP and NAA at varying dithionite concentrations.

From our anaerobic kinetic studies, the substituents around the nitro aromatic ring prove to have a significant effect on the rate of degradation of the compounds. From the data

Shown in figure 31 it can be concluded that stronger electron donating groups increase the rate and that an ortho substituted nitro trigger is preferred over para substituted. The reduction rates of various nitro aromatic compounds at varying concentrations are shown in figure 31.

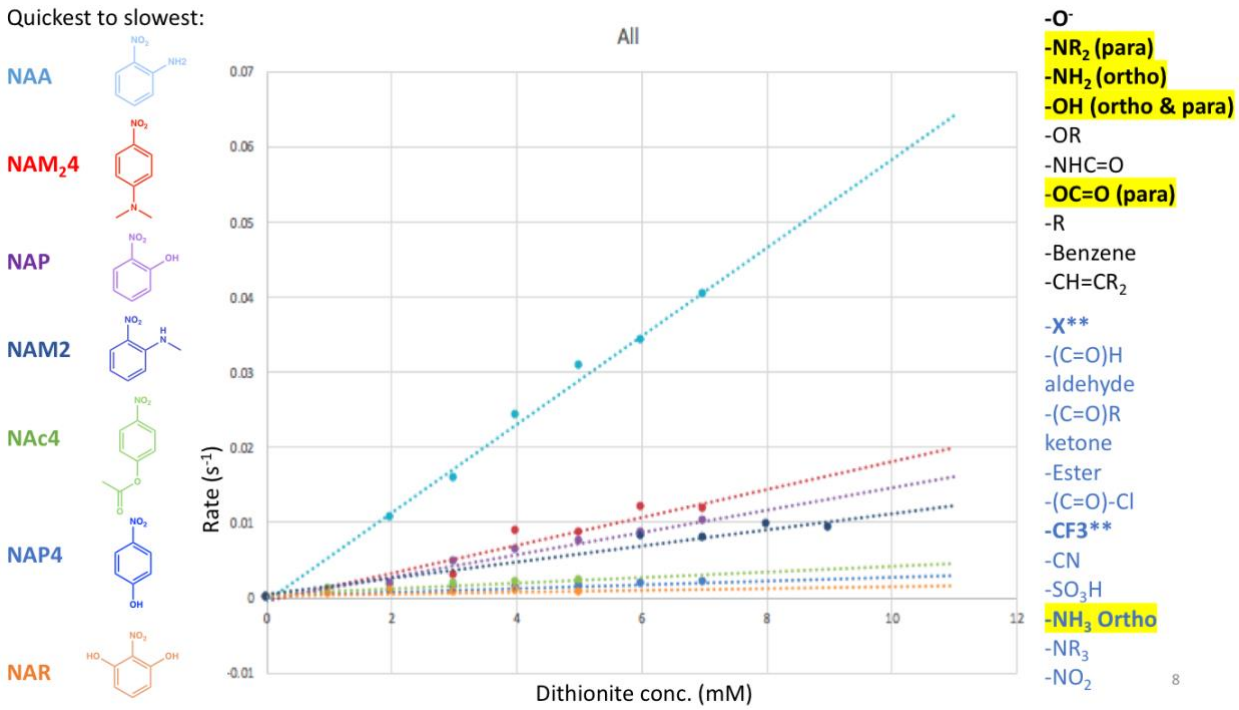


Figure 32 Compounds tested decreasing in rate from top to bottom and color coded to graph, 2-Nitroaniline, (NAA) being the quickest and nitro resorcinol (NAR) being the slowest. Beside the compounds is the graph of the normalized kinetic data. To the right of the graph is a list of electron donating substituents in decreasing electron density from top to bottom (in black) and below them are electron withdrawing substituents (in blue) in increasing electron deficiency from top to bottom

This trend could be contributed to tautomerization and the mesomeric effect, or possibly by proton exchange with the environment or the nitro group itself. Suggesting the largest energy barrier to overcome is not the initial electron transfer but rather exchange of a proton to form the irreversible hydroxylamine. It is relevant to also point out the products of two-electron reduction; nitroso compounds, are much stronger electron acceptors than the parent nitroaromatics, the reduction of nitroso compounds into the corresponding hydroxyl amines being 10^4 times faster than the reduction of nitroaromatics into the nitroso compounds[180]. However, previous studies on two electron reduction of nitroaromatics by *Enterobacter cloacae* nitroreductase have shown that there exists a parallelism between the electron- and hydride-accepting potency of nitroaromatics obtained by means of quantum-mechanical calculations[181] . In general, the log k_{cat}/K_m of nitroaromatic compounds tends to increase with an increase in their reduction potential, however the relationship is poorly expressed ($r^2 = 0.622$) [177] . This uncertainty suggests that the reactivity of nitroaromatics is strongly influenced by other factors.

Based on these initial results, Nitro aromatic phenol (NAP) and Nitro aromatic aniline (NAA) were chosen as the triggers for our kinetic studies, both ortho ED groups for consistency. The fluorescent compounds shown below in figure 32 were synthesized for our reduction studies. Our positive control standard was an unsubstituted nitrobenzene, representative of the current HAP, CH-01.

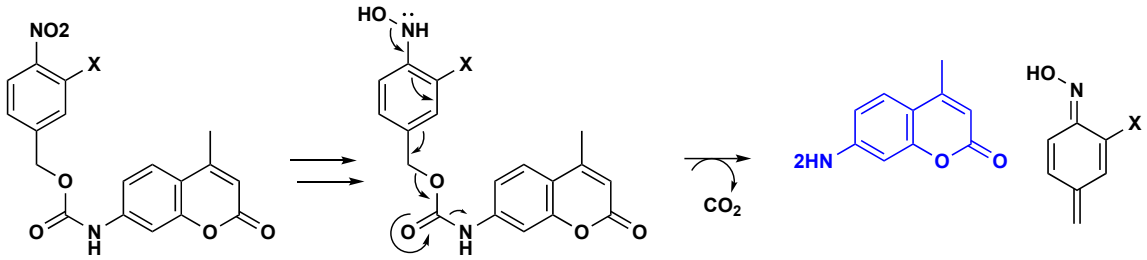


Figure 33 The predicted pathway of the release of the fluorescent coumarin substrate. X= H (CH-01), OH (NAP) or NH_2 (NAA).

4.2.2 Chemical Reduction assays with Dithionite

The fluorescent HAP was subjected to chemical reduction with sodium dithionite. This assay was carried out to examine the effectiveness of our chosen trigger to release a fluorophore upon reduction of the nitro functionality. This demonstrated not only that the nitro group is effectively reduced but also that the carbamate linker is labile towards chemical reduction but stable to various buffers and pH's. To test these conditions the small molecule was treated with dithionite and the rate of reduction was determined by measuring the increase in fluorescence, representing the release of the attached fluorophore. Kinetic assays were carried out in triplicate and plotted using prism.

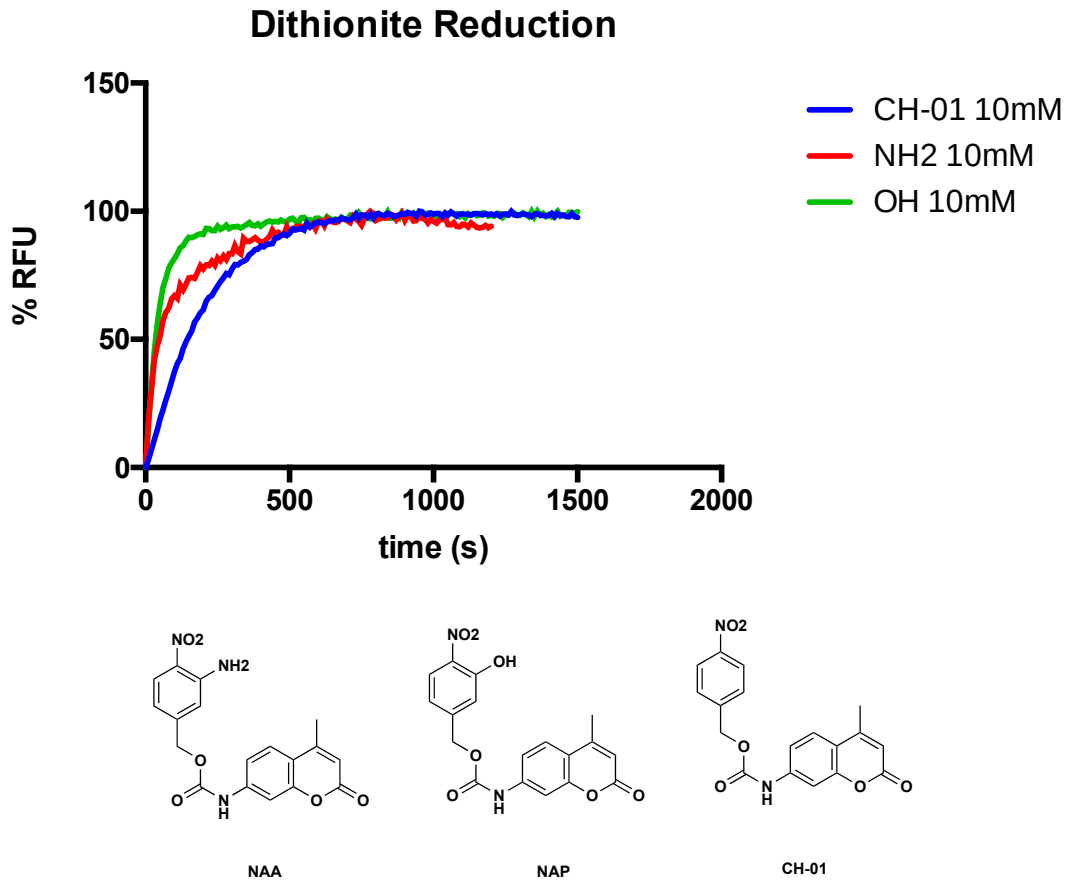


Figure 34 Normalized reductions of NAP, NAA and CH-01 trigger-fluorophores under chemical reduction of sodium dithionite. Rate is determined by release of fluorescent amino coumarin upon nitro reduction.

The results demonstrate that both ortho substituted compounds are reduced at higher rates than the standard unsubstituted nitroaromatic, regardless of their lower reduction potentials. These findings confirmed our initial trigger analysis studies and supported the theory of alternative rate determining steps. One possible cause for these unprecedented trends is the possibility solvent interactions playing significant roles in the rate of

reduction of these nitroaromatics. To better understand the interactions between protonation and reduction rates, deuterium kinetic isotope effects were investigated.

4.2.3 KIE studies

The 1-electron reduction pathway of nitro groups is shown in figure 11. In this scheme there are multiple protonation and deprotonation steps required to transform a nitro to a nitroso and subsequently to its' effective hydroxylamine/aniline. These transfers of protons are facilitated by the solvent in chemical reductions. To examine the effect of solvent on the overall rate, deuterium kinetic isotope assays were performed at pH 7.4. The nitroaromatic compounds were reduced under the same chemical conditions and concentrations while varying the solvent from water to deuterium. The results shown below suggest that the substituents have a strong correlation to the intermolecular protonation rates with the intermediate and the solvent. Each of these showed a normal deuterated isotope effect but interestingly, NAP showed the most significant change in rate with the deuterated solvent system.

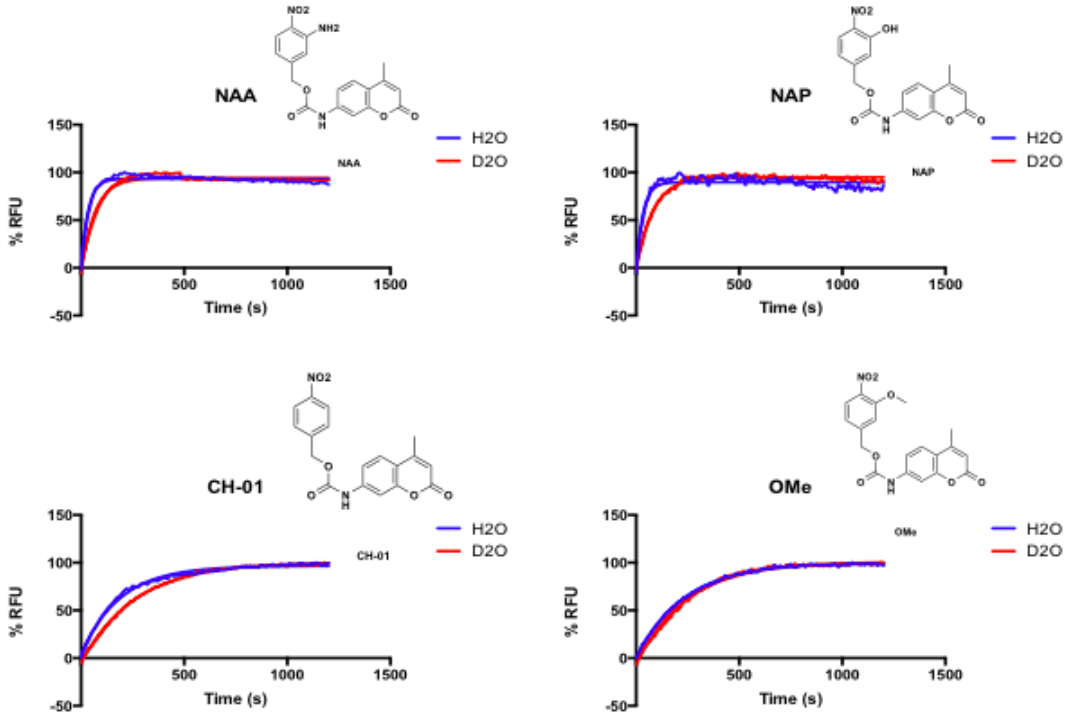


Figure 35 Deuterium kinetic isotope assay. The Rate of release of each compound's fluorophore from chemical reduction by sodium dithionite in water (blue) and in D_2O (red)

From these KIE's it can be determined that the rate determining step is not based on the reduction potential of the nitro group but more closely related to intermolecular protonation's in chemical reductions of these compounds. The KIE of the unsubstituted aromatic is above one but relatively low (KIE= 1.38) but the substituted nitro aromatics, NAP (KIE= 2.48) and NAA (KIE=1.97), show a significant change in KIE while the methoxy substituted compound showed little to no change in rate (KIE=1.049). These results suggest the role of protonation and induction from substituents has a significant

influence on the rate of reduction in these compounds. In previous studies, evidence from compound-specific nitrogen isotope analysis of nitro aromatic compound reduction reactions suggests that the rate limiting step in reduction is either (i) the dehydration of Ar−N(OH)2 to form Ar−NO, or (ii) proton or electron transfer to the nitroaromatic radical anion (Ar−NO2 •−), depending on the reductant and solution chemistry[182]. This could be a result of mesomeric effects in ortho substituted nitro aromatics. By introducing electron donating groups at the ortho position on these aromatic rings, the electron density in the ring is enhanced and the dipole moment of the nitro group is increased, making it more labile to accepting protons from the surrounding solvent. There is also the possibility of stabilization of protonation by Van der Waals forces from the electron rich substituents, facilitating the transfers of protons from the solvent to the nitro group. It is interesting to note that the rate of NAP was most effected by the deuterium solvent, even in comparison to NAA. The aniline substituent is more electron donating in comparison to NAP's hydroxyl group, suggesting the protonation of the acidic phenol plays a role in reduction whereas the aniline would likely be unaffected and maintain its original protons at biological pH's, the constant rate of the methoxy substituted compound also supports this.

4.2.4 Enzymatic Reduction Assays

Nitroreductase Assay

Once the electron rich triggers had been seen to be reduced chemically and stable to aqueous conditions, the two candidate compounds NAP and NAA were tested with

Nitroreductase to determine their ability to be reduced enzymatically. This initial enzymatic reduction was taken to prove that reduction potentials that were believed to be necessary for enzymatic reduction (>~-400mV) were in fact, not necessary at all. The two compounds were not only successfully reduced enzymatically, their rates paralleled their initial trigger reduction findings and their reduction by sodium dithionite. This showed that our compounds were capable of being reduced enzymatically regardless of their low reduction potentials (~ -600—700mV).

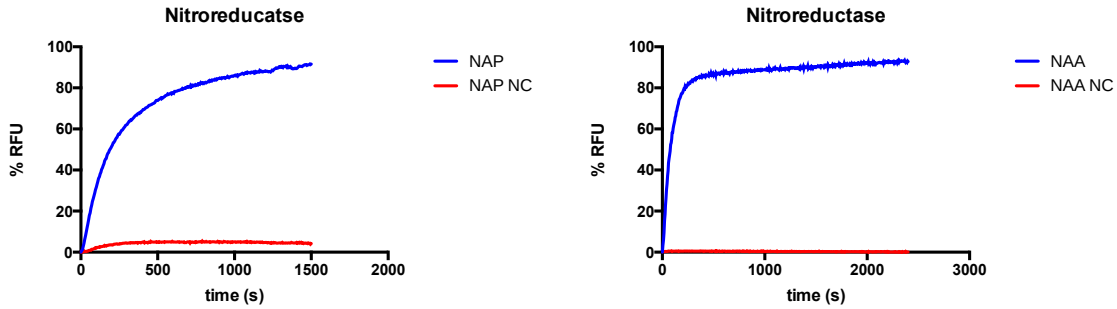


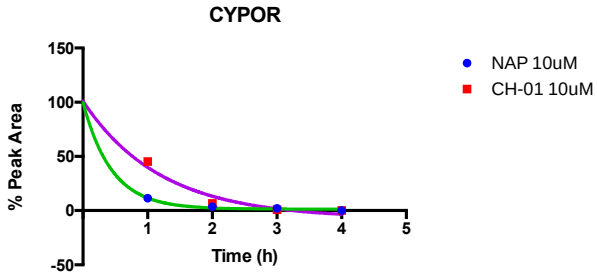
Figure 36 Synthesized fluorescent HAP's NAP and NAA reduction with Nitroreductase

The Rate for NAP was k_{obs}= .003(s^{-1}) and NAA was reduced at a rate of k_{obs}= .012(s^{-1}). These rates were similar to their chemical trigger reductions, NAA being reduced about 3 times more quickly than NAP. These results suggested the initial chemical trigger analysis was representative of enzymatic reduction rates. While these results were encouraging, they were not an accurate description of their reduction in humans. To better exemplify human conditions, we then tested these aromatics with Cytochrome

p450 oxidoreductase, the enzyme most responsible for the reduction of these xenobiotics in mammals.

CYPOR Reduction

Cytochrome P450 has shown to be the main contributing enzyme when breaking down these xenobiotics. From knockout studies, it has been shown that the activity of these compounds is significantly decreased compared to the absence of any other enzyme. While other enzymes have been shown to reduce HAPS, cytochrome p450 oxidoreductase is considered the main target. Cytochrome p450 is a flavoenzyme which is ubiquitous in nature and employs a broad range of functionalities. The enzyme requires NADP and flavin as co-factors to transfer electrons from its heme-site to the compounds being reduced. Because these co-factors have innate fluorescence, our fluorophores' reduction could not be measured by plate reader fluorescence due to overlapping wavelengths. To remedy this problem, degradation of the starting material was monitored by ESI-MS over a 6h period.



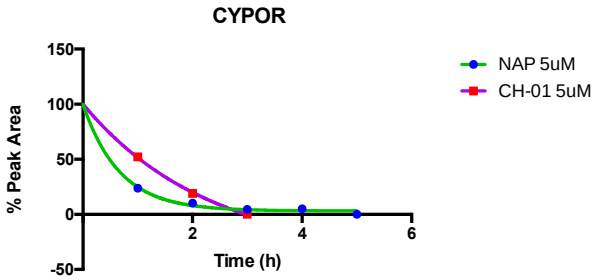


Figure 37 Reduction of fluorescent Hypoxia activated prodrug NAP and CH-01 with Cytochrome p450 oxidoreductase**.**

The compounds were tested at two different concentrations, 10uM and 5uM. The results show that both compounds were effectively reduced until there was no detectable starting material left. While the rates were similarly completed after three hours, the rate of reduction for our NAP compound showed to be reduced much more quickly that the CH-01 analog. The compounds were also tested against the enzyme in normoxic conditions and showed no significant decrease in the starting material, showing that they maintain hypoxic specificity. These results are shown in figure 37.

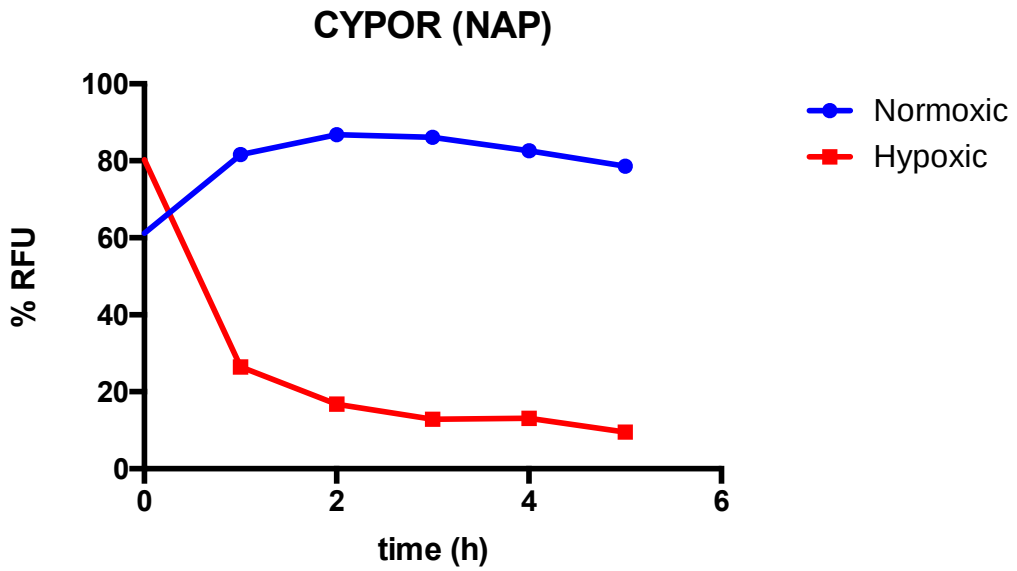


Figure 38 Cytochrome p450 selectivity assay, Normoxic vs. Hypoxic conditions with NAP

This by no means is a completed project, the enzymatic assays were carried out at a n=1 and need to be repeated twice more for each concentration, but these results are still promising and challenge the previous held belief that electron rich nitro compounds were undesirable triggers for the hypoxic enzymatic reduction of prodrugs. Further testing should include cell assays, and further KIE studies with Cytochrome p450. The rates for each compound and the various conditions in which they were run have been put together into table 7.

	NAA	**NAP**	**CH-01**
H_2O **(s^{-1})**	.0295	.0310	.0052
D_2O **(s^{-1})**	.0149	.0125	.0038
Nitro-reductase **(s^{-1})**	.0100	.0037	N/A
CYPOR (5uM) **(h^{-1})**	N/A	1.495	.0408
CYPOR (10uM) **(h^{-1})**	N/A	2.264	.1954

Table 6 Collective Rates of selected compounds from chemical reduction, nitro reductase reduction and CYPOR reduction

4.3 Conclusion

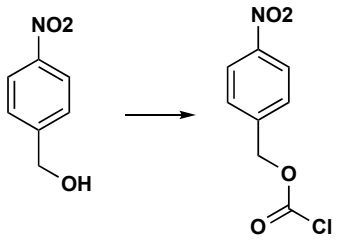
In conclusion, a novel hypoxia activated prodrug trigger has been identified and new insights to the reduction mechanism of these nitro aromatics have been explored. Beginning with elucidating the best trigger by chemical reduction of various substituted nitro compounds we were able to identify the preferred position of the linker to be para to the nitro group. We also were able to see an unprecedented trend in increasing rates and electron rich substituents. From these findings we then synthesized fluorescent analogs as well as the standard CH-01 to compare their rates of reduction. These compounds were reduced chemically in aqueous media as well as enzymatically by both

nitro-reductase and cytochrome p450 oxidoreductase. Our electron rich nitro-aromatic phenol (NAP) showed consistently high reduction rates and was broken down enzymatically effectively and more quickly than its clinical trial counterpart. Our nitro aromatic compounds also displayed deuterium kinetic isotope effects that strongly suggest proton transfer between the environment has a significant effect on their rate of reduction, more so for the electron rich substituents than the unsubstituted CH-01 analog. These results are promising for future endeavors in developing more effective HAPS. While previous studies have solidified effective linkers and effectors, these compounds typically vary and struggle due to their poorly understood triggers. For future progress, enzymatic reactions should be repeated and replicated in deuterium to reassure that proton transfer is in fact rate limiting in the *in vivo* reduction of these compounds. The final assay to be carried out would be the synthesis of NAP conjugated phosphoramide mustard or other inhibiting molecule. This active therapeutic should be tested in appropriate cells lines in both normoxic and hypoxic conditions. Once this data is collected the new prodrug should be considered for further clinical trials.

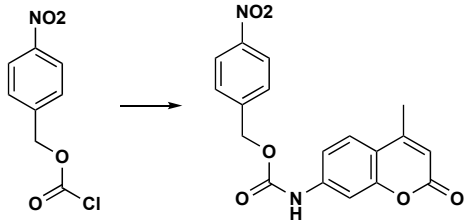
4.4 Experimental

4.4.1 Synthesis of Fluorescent compounds

Nitrobenzyl-AMC



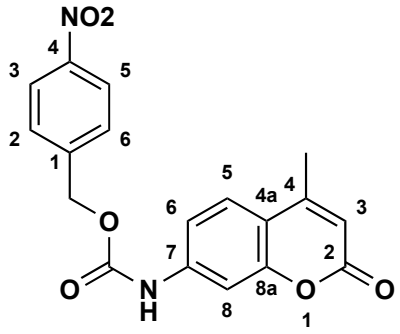
4-nitrobenzyl carbonochloridate: (4-nitrophenyl)methanol (1 eq) was dissolved in THF containing Na_2CO_3 and cooled to 0 C. To the solution was added triphosgene (1 eq) and the reaction was kept stirring for 12h at r.t. The reaction was filtered and the volatiles were subsequently evaporated without heating and the residue dried under vaccuum to give 4-nitrobenzyl carbonochloridate in quantitative conversions. The product was used in the next step without further purification.



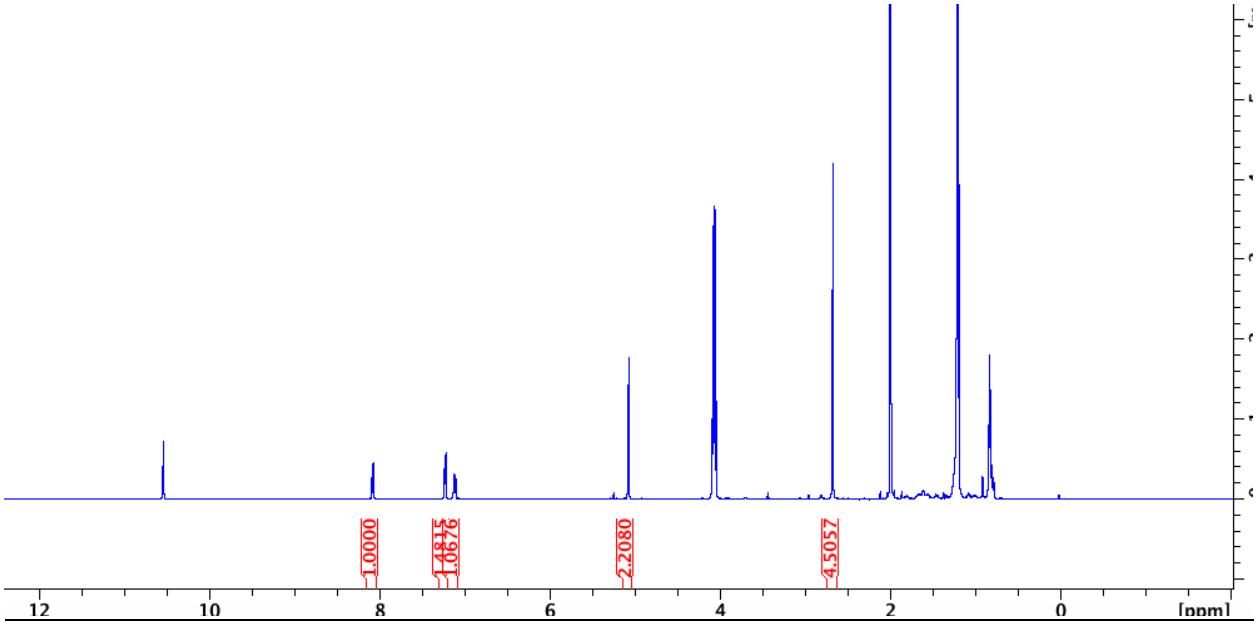
4-nitrobenzyl (4-methyl-2-oxo-2*H*-chromen-7-yl)carbamate: (1.1eq) of amino-methyl coumarin in THF at 0 °C was added (24.0 mmol) of $NaHCO_3$ followed by (1 eq) of 3-hydroxy-4-nitrobenzyl carbonochloridate. After 15, min the reaction mixture was warmed to rt, quenched with H_2O and extracted with 3 x 20 mL of EtOAc. The combined organic layers were dried ($MgSO_4$) and concentrated in vacuo to afford 4-nitrobenzyl (4-methyl-2-oxo-2*H*-chromen-7-yl)carbamate in quantitative yield as an

orange solid that was used without further purification. H NMR (DMSO-d6, 500 MHz) δ 10.42 (s, 1H), 8.28 (d, J = 8.6 Hz, 2H), 7.72−7.70 (m, 3H), 7.55 (d, J = 1.3 Hz, 1H), 7.42 (dd, J = 8.7, 1.3 Hz, 1H), 6.24 (s, 1H), 5.35 (s, 2H), 2.39 (s, 3H).

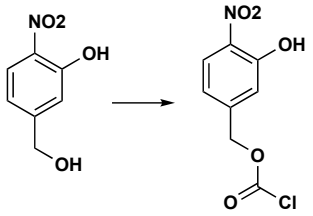
HRMS (ESI-TOF) m/z: [M + H]+ calcd for C18H14N2O6 355.0925



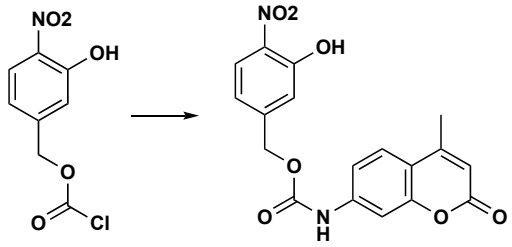
4-nitrobenzyl (4-methyl-2-oxo-2*H*-chromen-7-yl)carbamate



Nitrophenol-AMC

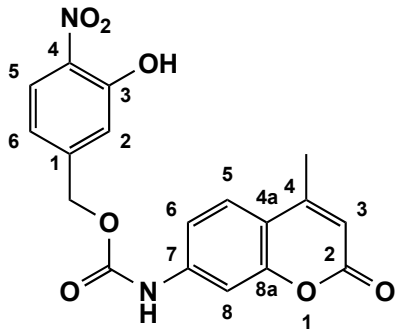


(CH-01) 5-(hydroxymethyl)-2-nitrophenol (1 eq) was dissolved in THF containing Na_2CO_3 and cooled to 0 C. To the solution was added triphosgene (1 eq) and the reaction was kept stirring for 12h at r.t. The reaction was filtered and the volatiles were subsequently evaporated without heating and the residue dried under vaccuum to give 3-hydroxy-4-nitrobenzyl carbonochloridate in quantitative conversions. The product was used in the next step without further purification

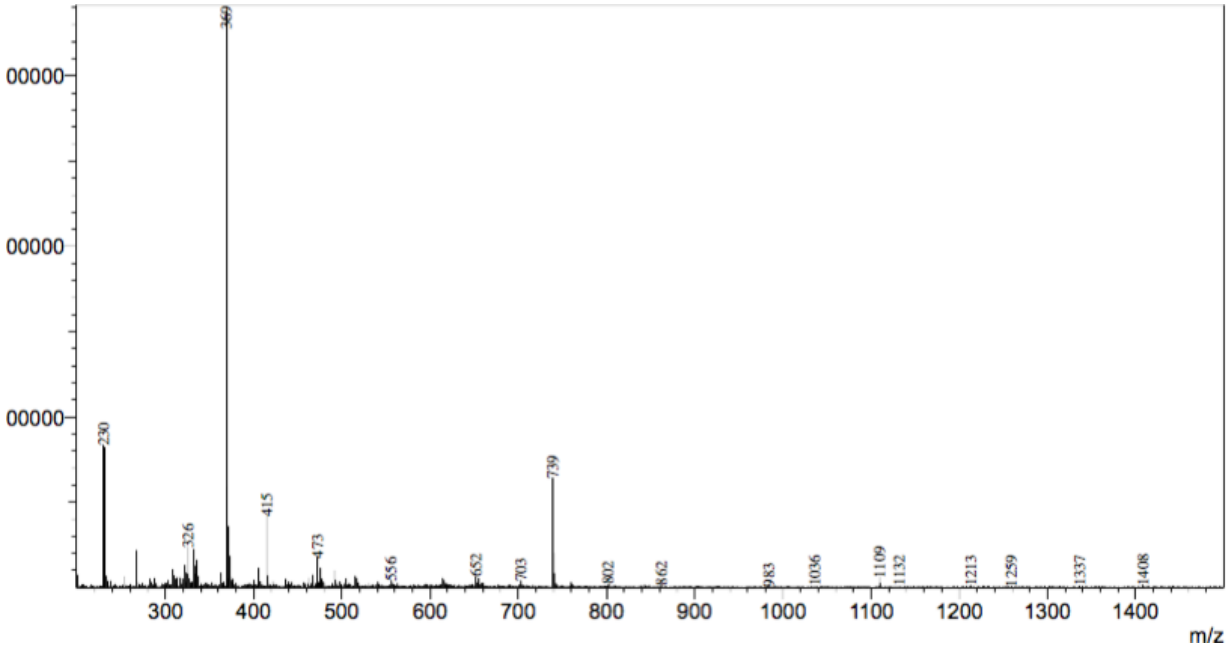


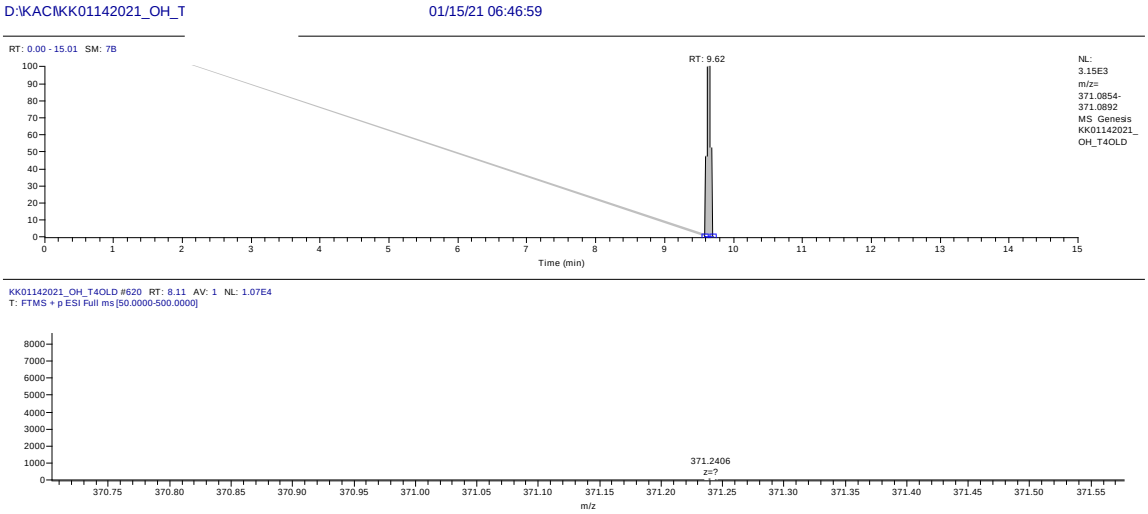
(NAP) 3-hydroxy-4-nitrobenzyl carbonochloridate (1.1eq) of amino-methyl coumarin in THF at 0 °C was added (24.0 mmol) of $NaHCO_3$ followed by (1 eq) of 3-hydroxy-4-nitrobenzyl carbonochloridate. After 15, min the reaction mixture was warmed to rt, quenched with H_2O and extracted with 3 x 20 mL of EtOAc. The combined organic layers were dried ($MgSO_4$) and concentrated in vacuo to afford 3-hydroxy-4-nitrobenzyl (4-methyl-2-oxo-2*H*-chromen-7-yl)carbamate in quantitative yield as an orange solid that was used without further purification. H NMR ($CDCl_3$-d6, 400 MHz) δ 10.8 (s, 1H), 10.42 (s, 1H), 8.28 (d, J = 8.6 Hz, 2H), 7.72−7.70 (m, 3H), 7.55 (d, J = 1.3 Hz, 1H), 7.42

(dd, J = 8.7, 1.3 Hz, 1H), 6.43 (s, 1H), 6.24 (s, 1H), 5.35 (s, 2H), 2.39 (s, 3H). HRMS (ESI-TOF) m/z: [M + H]+ calcd for C18H15N2O7 371.0925

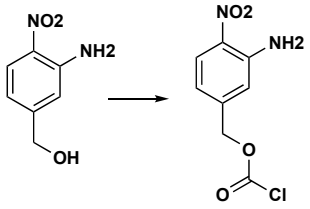


3-hydroxy-4-nitrobenzyl (4-methyl-2-oxo-2*H*-chromen-7-yl)carbamate

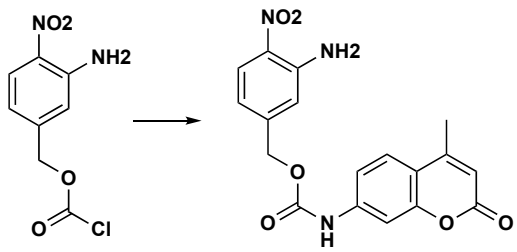




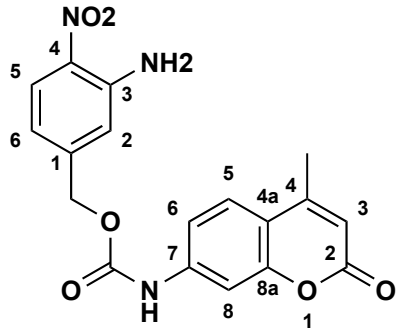
Nitro-Aniline-AMC



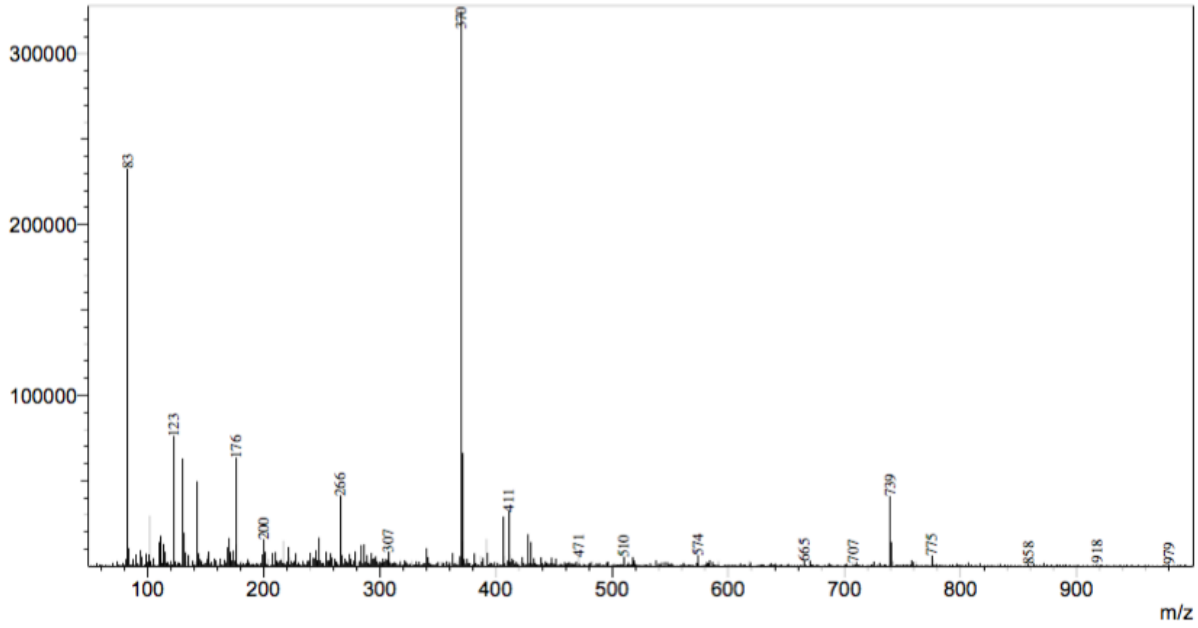
(NAA) 3-amino-4-nitrobenzyl carbonochloridate: (3-amino-4-nitrophenyl)methanol (1 eq) was dissolved in THF containing Na_2CO_3 and cooled to 0 C. To the solution was added triphosgene (1 eq) and the reaction was kept stirring for 12h at r.t. The reaction was filtered and the volatiles were subsequently evaporated without heating and the residue dried under vaccuum to give 4-nitrobenzyl carbonochloridate in quantitative conversions. The product was used in the next step without further purification



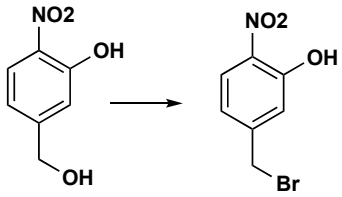
3-amino-4-nitrobenzyl (4-methyl-2-oxo-2*H*-chromen-7-yl)carbamate: (1.1eq) of amino-methyl coumarin in THF at 0 °C was added (24.0 mmol) of $NaHCO_3$ followed by (1 eq) of 3-amino-4-nitrobenzyl carbonochloridate. After 15, min the reaction mixture was warmed to rt, quenched with H_2O and extracted with 3 x 20 mL of EtOAc. The combined organic layers were dried ($MgSO_4$) and concentrated in vacuo to afford 3-amino-4-nitrobenzyl (4-methyl-2-oxo-2*H*-chromen-7-yl)carbamate in quantitative yield as an orange solid that was used without further purification. H NMR ($CDCl_3$-d6, 400 MHz) δ 10.42 (s, 1H), 8.28 (d, J = 8.6 Hz, 2H), 7.72−7.70 (m, 3H), 7.55 (d, J = 1.3 Hz, 1H), 7.42 (dd, J = 8.7, 1.3 Hz, 1H), 6.43 (s, 1H), 6.24 (s, 1H), 5.35 (s, 2H), 2.39 (s, 3H). HRMS (ESI-TOF) m/z: [M + H]+ calcd for C18H15N3O6 370.0925



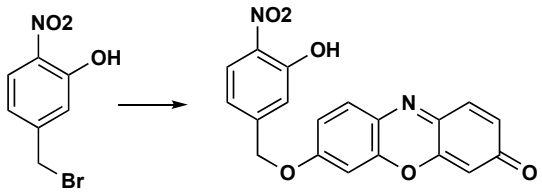
3-amino-4-nitrobenzyl (4-methyl-2-oxo-2*H*-chromen-7-yl)carbamate



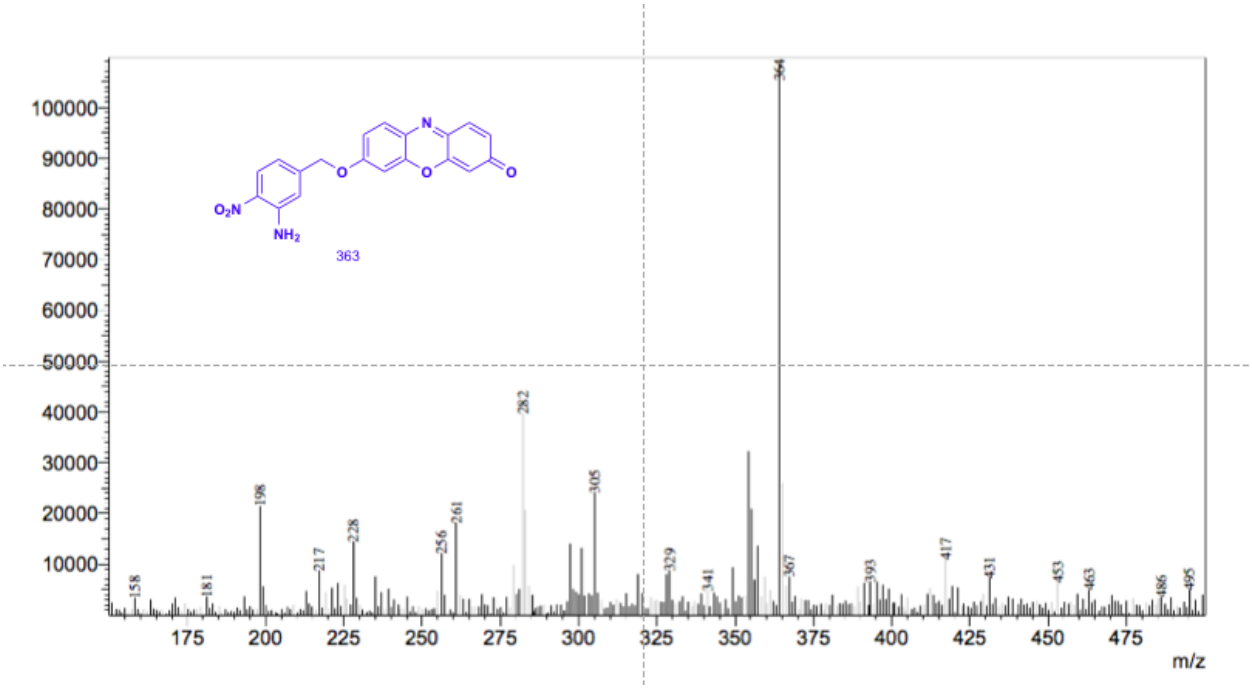
Nitro-Resorufin



To a well stirred solution of the Nitro aromatic alcohol in dry DCM at 0C was added PBr3 (1.1 eq) the reaction was stirred for 15h at 0C under argon. H2O was then added, organic layer separated and aqueous phase extracted twice with DCM. The organic layers were washed with brine and dried over Na2SO4 and filtered to give the product in quantitative yield.



Resorufin (1eq), benzyl bromide (1.2 eq) TBAB (1eq) K2PO4 (1.5eq) and water were added to the reaction vessel and stirred at room temperature under air. The reaction was monitored to completion by TLC. After completion the reaction was diluted with H2O and extracted with DCM. The organic layer was dried with Na2SO4 and concentrated in vacuo then purified by flash chromatography.



4.4.2 Trigger reduction assays

1M stock solutions were made of each nitro aromatic compound in methanol. 10µl of the nitro aromatic were added to 80µl 1xPBS buffer (7.4 pH). 10µl of Sodium Dithionite (final concentration 100mM) was added last. The Nitro aromatic concentration was

varied from 1mM to 10mM for each compound. A plate reader inside a glove box was used for all measurements. The rate was determined by the loss of absorbance from the starting nitro aromatic. Data were fit to the exponential decay equation -K*Y0 in GrpahPad Prism 9.0 software. The collected data were normalized into a linear plot to show their rates by concentration.

4.4.3 Dithionite reduction assay

1µl of each Nitro compound (A, B, C) (1M stock solution in DMSO) was added to 89µl 1xPBS buffer (7.4 pH). 10µl of Sodium Dithionite (final concentration 100mM) was added last. The solution mixed and the kinetics assay started immediately. The observed rate constant kobs was measured by the fluorescence of the coumarin released. Fluorescence was measured by a fluorometer with excitation 350nM and emission 450nM was used to monitor reaction kinetics.

4.4.4 KIE dithionite reduction assay

1µl of each Nitro compound (A, B, C) (1M stock solution in DMSO) was added to 50µl 1xPBS buffer (7.4 pH) and 49µl of D_2O. 10µl of Sodium Dithionite (final concentration 100mM) was added last. The solution mixed and the kinetics assay started immediately. The observed rate constant kobs was measured by the fluorescence of the coumarin released. Fluorescence was measured by a fluorometer with excitation 350nM and emission 450nM was used to monitor reaction kinetics.

4.4.5 Nitroreductase Assay

Recombinant *E. coli* nitroreductase and dihydronicotinamide adenine dinucleotide (NADPH, reduced form, tetrasodium salt) were purchased from Sigma–Aldrich. 10 μL of a prodrug stock solution (1.7 mM) in ultra-pure water with 13% DMSO were added to 1000 μL of PBS (10 mM, pH 7.4, 37 °C) containing NADPH. The reaction was performed in a glove box and initiated by addition of *E. coli* nitroreductase (25 μL of an initial solution, 1.0 mg/mL in PBS; final concentrations: prodrug: 17 μM; NADPH: 0.9 mM; nitroreductase: 25 μg/mL). Negative controls were prepared in the same manner without the addition *E. coli* nitroreductase. of The solution was mixed and the kinetics assay started immediately. The observed rate constant kobs was measured by the fluorescence of the coumarin released. Fluorescence was measured by a fluorometer with excitation 350nM and emission 450nM was used to monitor reaction kinetics.

4.4.6 General procedure for cytochrome P450 reductase assay

Mouse liver microsomes (20 mg/mL) were purchased from Corning. NADPH was purchased from Sigma-Aldrich. The measurements were made in 10 mM PBS (pH 7.4) according to the following procedure. 1 μL of 5 mM prodrug in were diluted in 1 mL PBS, followed by addition of $MgCl_2$ (5 μmol), NADPH (1 μmol), and 50 μL mouse liver microsomes. The incubation mixtures of test compounds with mouse liver microsomes contained the following at the indicated final concentrations: 10 mM PBS (pH 7.4), 5 μM test compound, 5 mM $MgCl_2$, microsomes (1 mg/mL) with or without 1 mM

NADPH. The mixtures were preincubated at 37 °C. The metabolism reaction was initiated by the addition of 50 μL mouse liver microsomes solution. At 0, 1, 2, 3, 4, and 5 h, 100 μL reaction mixture was quenched by 100 μL cold acetonitrile. The samples were centrifuged at 4 °C for 5 min at 12000 rpm. The supernatant was transferred to a vial for analysis using HPLC.

REFERENCES

1. Lau, J. L.; Dunn, M. K., Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg Med Chem* **2018,** *26* (10), 2700-2707.
2. Tan, Y.; Wu, H. X.; Wei, T. Y.; Li, X. C., Chemical Protein Synthesis: Advances, Challenges, and Outlooks. *J Am Chem Soc* **2020,** *142* (48), 20288-20298.
3. Muir, T. W.; Sondhi, D.; Cole, P. A., Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A* **1998,** *95* (12), 6705-10.
4. Fischer, E. O., E, Synthesis of the derivatives of some dipeptides. *Ber. Deutsch. Chem. Ges* **1903,** *36*.
5. Bergmann, M. Z., L., A general procedure of the peptide synthesis. *Ber. Deutsch. Chem. Ges.* **1932,** *65*.
6. Siffert, R. H. d. V., V., A new synthesis of carnosine, with some observations on the splitting of the benzyl group from carbobenzoxy derivatives and from benzylthioethers. *J. Biol. Chem* **1935,** *108*.
7. Du Vigneaud, V., Ressler, C., Swan, J.M., Roberts, C.W. & Katsoyannis, P.G, The synthesis of oxytocin. *J. Am. Chem. Soc.* **1954,** *76*.
8. Merrifield, B., The chemical synthesis of proteins. *Protein Sci* **1996,** *5* (9), 1947-51.
9. Atherton, E.; Clive, D. L.; Sheppard, R. C., Letter: Polyamide supports for polypeptide synthesis. *J Am Chem Soc* **1975,** *97* (22), 6584-5.
10. Albericio, F. & Giralt, E, Handles and supports. *Houben-Weyl. Methods of Organic Chemistry* **2002,** *E 22a*.
11. Dawson, P. E.; Kent, S. B., Synthesis of native proteins by chemical ligation. *Annu Rev Biochem* **2000,** *69*, 923-60.
12. Klabunde, T.; Sharma, S.; Telenti, A.; Jacobs, W. R., Jr.; Sacchettini, J. C., Crystal structure of GyrA intein from Mycobacterium xenopi reveals structural basis of protein splicing. *Nat Struct Biol* **1998,** *5* (1), 31-6.
13. Xu, M. Q.; Perler, F. B., The mechanism of protein splicing and its modulation by mutation. *EMBO J* **1996,** *15* (19), 5146-53.
14. David, R.; Richter, M. P.; Beck-Sickinger, A. G., Expressed protein ligation. Method and applications. *Eur J Biochem* **2004,** *271* (4), 663-77.
15. Goody, R. S.; Alexandrov, K.; Engelhard, M., Combining chemical and biological techniques to produce modified proteins. *Chembiochem* **2002,** *3* (5), 399-403.
16. Muir, T. W., Semisynthesis of proteins by expressed protein ligation. *Annu Rev Biochem* **2003,** *72*, 249-89.
17. Cole, P. A.; Courtney, A. D.; Shen, K.; Zhang, Z.; Qiao, Y.; Lu, W.; Williams, D. M., Chemical approaches to reversible protein phosphorylation. *Acc Chem Res* **2003,** *36* (6), 444-52.
18. Ottesen, J. J.; Huse, M.; Sekedat, M. D.; Muir, T. W., Semisynthesis of phosphovariants of Smad2 reveals a substrate preference of the activated T beta RI kinase. *Biochemistry* **2004,** *43* (19), 5698-706.

19. Grogan, M. J.; Pratt, M. R.; Marcaurelle, L. A.; Bertozzi, C. R., Homogeneous glycopeptides and glycoproteins for biological investigation. *Annu Rev Biochem* **2002,** *71*, 593-634.
20. Tolbert, T. J.; Wong, C. H., Conjugation of glycopeptide thioesters to expressed protein fragments: semisynthesis of glycosylated interleukin-2. *Methods Mol Biol* **2004,** *283*, 255-66.
21. Warren, J. D.; Miller, J. S.; Keding, S. J.; Danishefsky, S. J., Toward fully synthetic glycoproteins by ultimately convergent routes: a solution to a long-standing problem. *J Am Chem Soc* **2004,** *126* (21), 6576-8.
22. Romanelli, A.; Shekhtman, A.; Cowburn, D.; Muir, T. W., Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proc Natl Acad Sci U S A* **2004,** *101* (17), 6397-402.
23. Cotton, G. J.; Muir, T. W., Generation of a dual-labeled fluorescence biosensor for Crk-II phosphorylation using solid-phase expressed protein ligation. *Chem Biol* **2000,** *7* (4), 253-61.
24. Pellois, J. P.; Hahn, M. E.; Muir, T. W., Simultaneous triggering of protein activity and fluorescence. *J Am Chem Soc* **2004,** *126* (23), 7170-1.
25. Scheibner, K. A.; Zhang, Z.; Cole, P. A., Merging fluorescence resonance energy transfer and expressed protein ligation to analyze protein-protein interactions. *Anal Biochem* **2003,** *317* (2), 226-32.
26. https://coronavirus.jhu.edu/map.html (accessed February, 9).
27. Pislar, A.; Mitrovic, A.; Sabotic, J.; Pecar Fonovic, U.; Perisic Nanut, M.; Jakos, T.; Senjor, E.; Kos, J., The role of cysteine peptidases in coronavirus cell entry and replication: The therapeutic potential of cathepsin inhibitors. *PLoS Pathog* **2020,** *16* (11), e1009013.
28. Morse, J. S.; Lalonde, T.; Xu, S. Q.; Liu, W. R., Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV. *Chembiochem* **2020,** *21* (5), 730-738.
29. Morse, J. S.; Lalonde, T.; Xu, S.; Liu, W., Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV. *ChemRxiv* **2020**.
30. Xiong, M. Y.; Su, H. X.; Zhao, W. F.; Xie, H.; Shao, Q.; Xu, Y. C., What coronavirus 3C-like protease tells us: From structure, substrate selectivity, to inhibitor design. *Med Res Rev* **2021**.
31. Zhang, L. L.; Lin, D. Z.; Sun, X. Y. Y.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* **2020,** *368* (6489), 409-+.
32. Solowiej, J.; Thomson, J. A.; Ryan, K.; Luo, C.; He, M. Y.; Lou, J. H.; Murray, B. W., Steady-state and pre-steady-state kinetic evaluation of severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL(pro) cysteine protease: Development of an ion-pair model for catalysis. *Biochemistry-Us* **2008,** *47* (8), 2617-2630.

33. Lim, L. Z.; Shi, J. H.; Mu, Y. G.; Song, J. X., Dynamically-Driven Enhancement of the Catalytic Machinery of the SARS 3C-Like Protease by the S284-T285-I286/A Mutations on the Extra Domain. *Plos One* **2014,** *9* (7).
34. Chuck, C. P.; Chong, L. T.; Chen, C.; Chow, H. F.; Wan, D. C. C.; Wong, K. B., Profiling of Substrate Specificity of SARS-CoV 3CL(pro). *Plos One* **2010,** *5* (10).
35. Rut, W.; Groborz, K.; Zhang, L. L.; Sun, X. Y. Y.; Zmudzinski, M.; Pawlik, B.; Wang, X. Y.; Jochmans, D.; Neyts, J.; Mlynarski, W.; Hilgenfeld, R.; Drag, M., SARS-CoV-2 M-pro inhibitors and activity-based probes for patient-sample imaging. *Nat Chem Biol* **2021,** *17* (2), 222-+.
36. Xiong, M.; Su, H.; Zhao, W.; Xie, H.; Shao, Q.; Xu, Y., What coronavirus 3C-like protease tells us: From structure, substrate selectivity, to inhibitor design. *Med Res Rev* **2021**.
37. Zhang, L. L.; Lin, D. Z.; Kusov, Y.; Nian, Y.; Ma, Q. J.; Wang, J.; von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J.; de Wilde, A.; Snijder, E. J.; Liu, H.; Hilgenfeld, R., alpha-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J Med Chem* **2020,** *63* (9), 4562-4578.
38. Chuck, C. P.; Chen, C.; Ke, Z. H.; Wan, D. C. C.; Chow, H. F.; Wong, K. B., Design, synthesis and crystallographic analysis of nitrile-based broad-spectrum peptidomimetic inhibitors for coronavirus 3C-like proteases. *Eur J Med Chem* **2013,** *59*, 1-6.
39. Dai, W. H.; Zhang, B.; Jiang, X. M.; Su, H. X.; Li, J. A.; Zhao, Y.; Xie, X.; Jin, Z. M.; Peng, J. J.; Liu, F. J.; Li, C. P.; Li, Y.; Bai, F.; Wang, H. F.; Cheng, X.; Cen, X. B.; Hu, S. L.; Yang, X. N.; Wang, J.; Liu, X.; Xiao, G. F.; Jiang, H. L.; Rao, Z. H.; Zhang, L. K.; Xu, Y. C.; Yang, H. T.; Liu, H., Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **2020,** *368* (6497), 1331-+.
40. Sacco, M. D.; Ma, C. L.; Lagarias, P.; Gao, A.; Townsend, J. A.; Meng, X. Z.; Dube, P.; Zhang, X. J.; Hu, Y. M.; Kitamura, N.; Hurst, B.; Tarbet, B.; Marty, M. T.; Kolocouris, A.; Xiang, Y.; Chen, Y.; Wang, J., Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against M-pro and cathepsin L. *Sci Adv* **2020,** *6* (50).
41. Wu, C. Y.; King, K. Y.; Kuo, C. J.; Fang, J. M.; Wu, Y. T.; Ho, M. Y.; Liao, C. L.; Shie, J. J.; Liang, P. H.; Wong, C. H., Stable benzotriazole esters as mechanism-based inactivators of the severe acute respiratory syndrome 3CL protease. *Chem Biol* **2006,** *13* (3), 261-268.
42. Yang, H. T.; Yang, M. J.; Ding, Y.; Liu, Y. W.; Lou, Z. Y.; Zhou, Z.; Sun, L.; Mo, L. J.; Ye, S.; Pang, H.; Gao, G. F.; Anand, K.; Bartlam, M.; Hilgenfeld, R.; Rao, Z. H., The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *P Natl Acad Sci USA* **2003,** *100* (23), 13190-13195.
43. Lee, C. C.; Kuo, C. J.; Ko, T. P.; Hsu, M. F.; Tsui, Y. C.; Chang, S. C.; Yang, S.; Chen, S. J.; Chen, H. C.; Hsu, M. C.; Shih, S. R.; Liang, P. H.; Wang, A. H.,

Structural basis of inhibition specificities of 3C and 3C-like proteases by zinc-coordinating and peptidomimetic compounds. *J Biol Chem* **2009,** *284* (12), 7646-55.
44. Lee, T. W.; Cherney, M. M.; Huitema, C.; Liu, J.; James, K. E.; Powers, J. C.; Eltis, L. D.; James, M. N. G., Crystal structures of the main peptidase from the SARS coronavirus inhibited by a substrate-like aza-peptide epoxide. *J Mol Biol* **2005,** *353* (5), 1137-1151.
45. Yin, J.; Niu, C.; Cherney, M. M.; Zhang, J.; Huitema, C.; Eltis, L. D.; Vederas, J. C.; James, M. N., A mechanistic view of enzyme inhibition and peptide hydrolysis in the active site of the SARS-CoV 3C-like peptidase. *J Mol Biol* **2007,** *371* (4), 1060-74.
46. Turlington, M.; Chun, A.; Tomar, S.; Eggler, A.; Grum-Tokars, V.; Jacobs, J.; Daniels, J. S.; Dawson, E.; Saldanha, A.; Chase, P.; Baez-Santos, Y. M.; Lindsley, C. W.; Hodder, P.; Mesecar, A. D.; Stauffer, S. R., Discovery of N-(benzo[1,2,3]triazol-1-yl)-N-(benzyl)acetamido)phenyl) carboxamides as severe acute respiratory syndrome coronavirus (SARS-CoV) 3CLpro inhibitors: Identification of ML300 and noncovalent nanomolar inhibitors with an induced-fit binding. *Bioorg Med Chem Lett* **2013,** *23* (22), 6172-6177.
47. Su, H. X.; Yao, S.; Zhao, W. F.; Li, M. J.; Liu, J.; Shang, W. J.; Xie, H.; Ke, C. Q.; Hu, H. C.; Gao, M. N.; Yu, K. Q.; Liu, H.; Shen, J. S.; Tang, W.; Zhang, L. K.; Xiao, G. F.; Ni, L.; Wang, D. W.; Zuo, J. P.; Jiang, H. L.; Bai, F.; Wu, Y.; Ye, Y.; Xu, Y. C., Anti-SARS-CoV-2 activities in vitro of Shuanghuanglian preparations and bioactive ingredients. *Acta Pharmacol Sin* **2020,** *41* (9), 1167-1177.
48. Jin, Z. M.; Du, X. Y.; Xu, Y. C.; Deng, Y. Q.; Liu, M. Q.; Zhao, Y.; Zhang, B.; Li, X. F.; Zhang, L. K.; Peng, C.; Duan, Y. K.; Yu, J.; Wang, L.; Yang, K. L.; Liu, F. J.; Jiang, R. D.; Yang, X. L.; You, T.; Liu, X. C.; Yang, X. N.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W. Q.; Xiao, G. F.; Qin, C. F.; Shi, Z. L.; Jiang, H. L.; Rao, Z. H.; Yang, H. T., Structure of M-pro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020,** *582* (7811), 289-+.
49. Zhou, L.; Liu, Y.; Zhang, W. L.; Wei, P.; Huang, C. K.; Pei, J. F.; Yuan, Y. X.; Lai, L. H., Isatin compounds as noncovalent SARS coronavirus 3C-like protease inhibitors. *J Med Chem* **2006,** *49* (12), 3440-3443.
50. Jacobs, J.; Grum-Tokars, V.; Zhou, Y.; Turlington, M.; Saldanha, S. A.; Chase, P.; Eggler, A.; Dawson, E. S.; Baez-Santos, Y. M.; Tomar, S.; Mielech, A. M.; Baker, S. C.; Lindsley, C. W.; Hodder, P.; Mesecar, A.; Stauffer, S. R., Discovery, Synthesis, And Structure-Based Optimization of a Series of N-(tert-Butyl)-2-(N-arylamido)-2-(pyridin-3-yl) Acetamides (ML188) as Potent Noncovalent Small Molecule Inhibitors of the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) 3CL Protease. *J Med Chem* **2013,** *56* (2), 534-546.
51. Eastman, R. T.; Roth, J. S.; Brimacombe, K. R.; Simeonov, A.; Shen, M.; Patnaik, S.; Hall, M. D., Remdesivir: A Review of Its Discovery and Development Leading to Emergency Use Authorization for Treatment of COVID-19. *Acs Central Sci* **2020,** *6* (5), 672-683.
52. Li, Z.; Li, X.; Huang, Y. Y.; Wu, Y. X.; Liu, R. D.; Zhou, L. L.; Lin, Y. X.; Wu, D. Y.; Zhang, L.; Liu, H.; Xu, X. M.; Yu, K. Q.; Zhang, Y. X.; Cui, J.; Zhan, C.

G.; Wang, X.; Luo, H. B., Identify potent SARS-CoV-2 main protease inhibitors via accelerated free energy perturbation-based virtual screening of existing drugs. *P Natl Acad Sci USA* **2020,** *117* (44), 27381-27387.
53. Vuong, W.; Khan, M. B.; Fischer, C.; Arutyunova, E.; Lamer, T.; Shields, J.; Saffran, H. A.; Mckay, R. T.; van Belkum, M. J.; Joyce, M. A.; Young, H. S.; Tyrrell, D. L.; Vederas, J. C.; Lemieux, M. J., Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication (vol 11, 4282, 2020). *Nature Communications* **2020,** *11* (1).
54. Kaptein, S. J. F.; Jacobs, S.; Langendries, L.; Seldeslachts, L.; ter Horst, S.; Liesenborghs, L.; Hens, B.; Vergote, V.; Heylen, E.; Barthelemy, K.; Maas, E.; De Keyzer, C.; Bervoets, L.; Rymenants, J.; Van Buyten, T.; Zhang, X.; Abdelnabi, R.; Pang, J.; Williams, R.; Thibaut, H. J.; Dallmeier, K.; Boudewijns, R.; Wouters, J.; Augustijns, P.; Verougstraete, N.; Cawthorne, C.; Breuer, J.; Solas, C.; Weynand, B.; Annaert, P.; Spriet, I.; Velde, G. V.; Neyts, J.; Rocha-Pereira, J.; Delang, L., Favipiravir at high doses has potent antiviral activity in SARS-CoV-2-infected hamsters, whereas hydroxychloroquine lacks activity. *P Natl Acad Sci USA* **2020,** *117* (43), 26955-26965.
55. Maciorowski, D.; El Idrissi, S. Z.; Gupta, Y.; Medernach, B. J.; Burns, M. B.; Becker, D. P.; Durvasula, R.; Kempaiah, P., A Review of the Preclinical and Clinical Efficacy of Remdesivir, Hydroxychloroquine, and Lopinavir-Ritonavir Treatments against COVID-19. *Slas Discov* **2020,** *25* (10), 1108-1122.
56. Phillips, R. M., Targeting the hypoxic fraction of tumours using hypoxia-activated prodrugs. *Cancer Chemother Pharmacol* **2016,** *77* (3), 441-57.
57. Koch, C. J.; Jenkins, W. T.; Jenkins, K. W.; Yang, X. Y.; Shuman, A. L.; Pickup, S.; Riehl, C. R.; Paudyal, R.; Poptani, H.; Evans, S. M., Mechanisms of blood flow and hypoxia production in rat 9L-epigastric tumors. *Tumor Microenviron Ther* **2013,** *1*, 1-13.
58. Graves, E. E.; Giaccia, A. J., Imaging tumoral hypoxia: oxygen concentrations and beyond. *Oncology (Williston Park)* **2007,** *21* (3), 368-76; discussion 377-8, 384.
59. Nepali, K.; Lee, H. Y.; Liou, J. P., Nitro-Group-Containing Drugs. *J Med Chem* **2019,** *62* (6), 2851-2893.
60. M., Q., The oxygen effect in radiation inactivation of DNA and enzymes. *Int J Radiat Biol Relat Stud Phys Chem Med.* **1986,** *50* (4), 573-94.
61. Hockel, M.; Vaupel, P., Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. *J Natl Cancer Inst* **2001,** *93* (4), 266-76.
62. Ward, J. F., DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog Nucleic Acid Res Mol Biol* **1988,** *35*, 95-125.
63. Patterson, A. V.; Ferry, D. M.; Edmunds, S. J.; Gu, Y.; Singleton, R. S.; Patel, K.; Pullen, S. M.; Hicks, K. O.; Syddall, S. P.; Atwell, G. J.; Yang, S.; Denny, W. A.; Wilson, W. R., Mechanism of action and preclinical antitumor activity of the novel hypoxia-activated DNA cross-linking agent PR-104. *Clin Cancer Res* **2007,** *13* (13), 3922-32.

64. Quintiliani, M., The oxygen effect in radiation inactivation of DNA and enzymes. *Int J Radiat Biol Relat Stud Phys Chem Med* **1986,** *50* (4), 573-94.
65. McKeown, S. R., Defining normoxia, physoxia and hypoxia in tumours-implications for treatment response. *Br J Radiol* **2014,** *87* (1035), 20130676.
66. Brown, J. M., The hypoxic cell: a target for selective cancer therapy--eighteenth Bruce F. Cain Memorial Award lecture. *Cancer Res* **1999,** *59* (23), 5863-70.
67. Foehrenbacher, A.; Patel, K.; Abbattista, M. R.; Guise, C. P.; Secomb, T. W.; Wilson, W. R.; Hicks, K. O., The Role of Bystander Effects in the Antitumor Activity of the Hypoxia-Activated Prodrug PR-104. *Front Oncol* **2013,** *3*, 263.
68. Mason, R. P.; Holtzman, J. L., The role of catalytic superoxide formation in the O2 inhibition of nitroreductase. *Biochem Biophys Res Commun* **1975,** *67* (4), 1267-74.
69. Kappen, L. S.; Lee, T. R.; Yang, C. C.; Goldberg, I. H., Oxygen transfer from the nitro group of a nitroaromatic radiosensitizer to a DNA sugar damage product. *Biochemistry* **1989,** *28* (11), 4540-2.
70. Dische, S., Chemical sensitizers for hypoxic cells: a decade of experience in clinical radiotherapy. *Radiother Oncol* **1985,** *3* (2), 97-115.
71. Hong, C. R.; Dickson, B. D.; Jaiswal, J. K.; Pruijn, F. B.; Hunter, F. W.; Hay, M. P.; Hicks, K. O.; Wilson, W. R., Cellular pharmacology of evofosfamide (TH-302): A critical re-evaluation of its bystander effects. *Biochem Pharmacol* **2018,** *156*, 265-280.
72. Overgaard, J.; Eriksen, J. G.; Nordsmark, M.; Alsner, J.; Horsman, M. R.; Danish, H.; Neck Cancer Study, G., Plasma osteopontin, hypoxia, and response to the hypoxia sensitiser nimorazole in radiotherapy of head and neck cancer: results from the DAHANCA 5 randomised double-blind placebo-controlled trial. *Lancet Oncol* **2005,** *6* (10), 757-64.
73. Toustrup, K.; Sorensen, B. S.; Metwally, M. A.; Tramm, T.; Mortensen, L. S.; Overgaard, J.; Alsner, J., Validation of a 15-gene hypoxia classifier in head and neck cancer for prospective use in clinical trials. *Acta Oncol* **2016,** *55* (9-10), 1091-1098.
74. Jain, R. K., Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy. *Science* **2005,** *307* (5706), 58-62.
75. Nordsmark, M.; Bentzen, S. M.; Rudat, V.; Brizel, D.; Lartigau, E.; Stadler, P.; Becker, A.; Adam, M.; Molls, M.; Dunst, J.; Terris, D. J.; Overgaard, J., Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy. An international multi-center study. *Radiother Oncol* **2005,** *77* (1), 18-24.
76. Pennacchietti, S.; Michieli, P.; Galluzzo, M.; Mazzone, M.; Giordano, S.; Comoglio, P. M., Hypoxia promotes invasive growth by transcriptional activation of the met protooncogene. *Cancer Cell* **2003,** *3* (4), 347-61.
77. Chang, Q.; Jurisica, I.; Do, T.; Hedley, D. W., Hypoxia predicts aggressive growth and spontaneous metastasis formation from orthotopically grown primary xenografts of human pancreatic cancer. *Cancer Res* **2011,** *71* (8), 3110-20.
78. Ozawa, S.; Sugiyama, Y.; Mitsuhashi, J.; Inaba, M., Kinetic analysis of cell killing effect induced by cytosine arabinoside and cisplatin in relation to cell cycle phase specificity in human colon cancer and Chinese hamster cells. *Cancer Res* **1989,** *49* (14), 3823-8.

79. Huxham, L. A.; Kyle, A. H.; Baker, J. H.; Nykilchuk, L. K.; Minchinton, A. I., Microregional effects of gemcitabine in HCT-116 xenografts. *Cancer Res* **2004,** *64* (18), 6537-41.
80. Durand, R. E.; Olive, P. L., Evaluation of bioreductive drugs in multicell spheroids. *Int J Radiat Oncol Biol Phys* **1992,** *22* (4), 689-92.
81. Hay, M. P.; Pchalek, K.; Pruijn, F. B.; Hicks, K. O.; Siim, B. G.; Anderson, R. F.; Shinde, S. S.; Phillips, V.; Denny, W. A.; Wilson, W. R., Hypoxia-selective 3-alkyl 1,2,4-benzotriazine 1,4-dioxides: the influence of hydrogen bond donors on extravascular transport and antitumor activity. *J Med Chem* **2007,** *50* (26), 6654-64.
82. Hicks, K. O.; Pruijn, F. B.; Sturman, J. R.; Denny, W. A.; Wilson, W. R., Multicellular resistance to tirapazamine is due to restricted extravascular transport: a pharmacokinetic/pharmacodynamic study in HT29 multicellular layer cultures. *Cancer Res* **2003,** *63* (18), 5970-7.
83. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B., Synthesis of proteins by native chemical ligation. *Science* **1994,** *266* (5186), 776-9.
84. Fang, G. M.; Li, Y. M.; Shen, F.; Huang, Y. C.; Li, J. B.; Lin, Y.; Cui, H. K.; Liu, L., Protein chemical synthesis by ligation of peptide hydrazides. *Angew Chem Int Ed Engl* **2011,** *50* (33), 7645-9.
85. Flood, D. T.; Hintzen, J. C. J.; Bird, M. J.; Cistrone, P. A.; Chen, J. S.; Dawson, P. E., Leveraging the Knorr Pyrazole Synthesis for the Facile Generation of Thioester Surrogates for use in Native Chemical Ligation. *Angew Chem Int Ed Engl* **2018,** *57* (36), 11634-11639.
86. Muralidharan, V.; Muir, T. W., Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat Methods* **2006,** *3* (6), 429-38.
87. Shogren-Knaak, M.; Ishii, H.; Sun, J. M.; Pazin, M. J.; Davie, J. R.; Peterson, C. L., Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **2006,** *311* (5762), 844-7.
88. McGinty, R. K.; Kim, J.; Chatterjee, C.; Roeder, R. G.; Muir, T. W., Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **2008,** *453* (7196), 812-6.
89. Zhang, X.; Zhao, Y.; Wang, C.; Ju, H.; Liu, W.; Zhang, X.; Miao, S.; Wang, L.; Sun, Q.; Song, W., Rhomboid domain-containing protein 1 promotes breast cancer progression by regulating the p-Akt and CDK2 levels. *Cell Commun Signal* **2018,** *16* (1), 65.
90. Schwarzer, D.; Cole, P. A., Protein semisynthesis and expressed protein ligation: chasing a protein's tail. *Curr Opin Chem Biol* **2005,** *9* (6), 561-9.
91. Ma, P.; Yuille, H. M.; Blessie, V.; Gohring, N.; Igloi, Z.; Nishiguchi, K.; Nakayama, J.; Henderson, P. J.; Phillips-Jones, M. K., Expression, purification and activities of the entire family of intact membrane sensor kinases from Enterococcus faecalis. *Mol Membr Biol* **2008,** *25* (6-7), 449-73.
92. Dann, G. P.; Liszczak, G. P.; Bagert, J. D.; Muller, M. M.; Nguyen, U. T. T.; Wojcik, F.; Brown, Z. Z.; Bos, J.; Panchenko, T.; Pihl, R.; Pollock, S. B.; Diehl, K. L.; Allis, C. D.; Muir, T. W., ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature* **2017,** *548* (7669), 607-611.

93. Rak, A.; Pylypenko, O.; Durek, T.; Watzke, A.; Kushnir, S.; Brunsveld, L.; Waldmann, H.; Goody, R. S.; Alexandrov, K., Structure of Rab GDP-dissociation inhibitor in complex with prenylated YPT1 GTPase. *Science* **2003,** *302* (5645), 646-50.
94. Stevens, A. J.; Sekar, G.; Shah, N. H.; Mostafavi, A. Z.; Cowburn, D.; Muir, T. W., A promiscuous split intein with expanded protein engineering applications. *Proc Natl Acad Sci U S A* **2017,** *114* (32), 8538-8543.
95. Amitai, G.; Callahan, B. P.; Stanger, M. J.; Belfort, G.; Belfort, M., Modulation of intein activity by its neighboring extein substrates. *Proc Natl Acad Sci U S A* **2009,** *106* (27), 11005-10.
96. Oeemig, J. S.; Zhou, D.; Kajander, T.; Wlodawer, A.; Iwai, H., NMR and crystal structures of the Pyrococcus horikoshii RadA intein guide a strategy for engineering a highly efficient and promiscuous intein. *J Mol Biol* **2012,** *421* (1), 85-99.
97. Evans, T. C., Jr.; Benner, J.; Xu, M. Q., The in vitro ligation of bacterially expressed proteins using an intein from Methanobacterium thermoautotrophicum. *J Biol Chem* **1999,** *274* (7), 3923-6.
98. Vila-Perello, M.; Liu, Z.; Shah, N. H.; Willis, J. A.; Idoyaga, J.; Muir, T. W., Streamlined expressed protein ligation using split inteins. *J Am Chem Soc* **2013,** *135* (1), 286-92.
99. Okamoto, R.; Morooka, K.; Kajihara, Y., A synthetic approach to a peptide alpha-thioester from an unprotected peptide through cleavage and activation of a specific peptide bond by N-acetylguanidine. *Angew Chem Int Ed Engl* **2012,** *51* (1), 191-6.
100. Miyajima, R.; Tsuda, Y.; Inokuma, T.; Shigenaga, A.; Imanishi, M.; Futaki, S.; Otaka, A., Preparation of peptide thioesters from naturally occurring sequences using reaction sequence consisting of regioselective S-cyanylation and hydrazinolysis. *Biopolymers* **2016,** *106* (4), 531-46.
101. Friedman, M.; Williams, L. D.; Masri, M. S., Reductive alkylation of proteins with aromatic aldehydes and sodium cyanoborohydride. *Int J Pept Protein Res* **1974,** *6* (3), 183-5.
102. Middlebrook, W. R., Identification of the end-amino groups of wool by means of their 2:4-dinitrophenyl derivatives. *Nature* **1949,** *164* (4168), 501.
103. Degani, Y.; Neumann, H.; Patchornik, A., Selective cyanylation of sulfhydryl groups. *J Am Chem Soc* **1970,** *92* (23), 6969-71.
104. Lees, A.; Nelson, B. L.; Mond, J. J., Activation of soluble polysaccharides with 1-cyano-4-dimethylaminopyridinium tetrafluoroborate for use in protein-polysaccharide conjugate vaccines and immunological reagents. *Vaccine* **1996,** *14* (3), 190-8.
105. Komander, D.; Rape, M., The ubiquitin code. *Annu Rev Biochem* **2012,** *81*, 203-29.
106. Swatek, K. N.; Komander, D., Ubiquitin modifications. *Cell Res* **2016,** *26* (4), 399-422.
107. Cappadocia, L.; Lima, C. D., Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chem Rev* **2018,** *118* (3), 889-918.
108. van der Veen, A. G.; Ploegh, H. L., Ubiquitin-like proteins. *Annu Rev Biochem* **2012,** *81*, 323-57.

109. Ekkebus, R.; van Kasteren, S. I.; Kulathu, Y.; Scholten, A.; Berlin, I.; Geurink, P. P.; de Jong, A.; Goerdayal, S.; Neefjes, J.; Heck, A. J.; Komander, D.; Ovaa, H., On terminal alkynes that can react with active-site cysteine nucleophiles in proteases. *J Am Chem Soc* **2013,** *135* (8), 2867-70.
110. Pruneda, J. N.; Durkin, C. H.; Geurink, P. P.; Ovaa, H.; Santhanam, B.; Holden, D. W.; Komander, D., The Molecular Basis for Ubiquitin and Ubiquitin-like Specificities in Bacterial Effector Proteases. *Mol Cell* **2016,** *63* (2), 261-276.
111. Sommer, S.; Weikart, N. D.; Linne, U.; Mootz, H. D., Covalent inhibition of SUMO and ubiquitin-specific cysteine proteases by an in situ thiol-alkyne addition. *Bioorg Med Chem* **2013,** *21* (9), 2511-7.
112. Paudel, P.; Zhang, Q.; Leung, C.; Greenberg, H. C.; Guo, Y.; Chern, Y. H.; Dong, A.; Li, Y.; Vedadi, M.; Zhuang, Z.; Tong, Y., Crystal structure and activity-based labeling reveal the mechanisms for linkage-specific substrate recognition by deubiquitinase USP9X. *Proc Natl Acad Sci U S A* **2019,** *116* (15), 7288-7297.
113. Basters, A.; Knobeloch, K. P.; Fritz, G., How USP18 deals with ISG15-modified proteins: structural basis for the specificity of the protease. *FEBS J* **2018,** *285* (6), 1024-1029.
114. Catic, A.; Fiebiger, E.; Korbel, G. A.; Blom, D.; Galardy, P. J.; Ploegh, H. L., Screen for ISG15-crossreactive deubiquitinases. *PLoS One* **2007,** *2* (7), e679.
115. Dang, L. C.; Melandri, F. D.; Stein, R. L., Kinetic and mechanistic studies on the hydrolysis of ubiquitin C-terminal 7-amido-4-methylcoumarin by deubiquitinating enzymes. *Biochemistry* **1998,** *37* (7), 1868-79.
116. Basu, A.; Rose, K. L.; Zhang, J.; Beavis, R. C.; Ueberheide, B.; Garcia, B. A.; Chait, B.; Zhao, Y.; Hunt, D. F.; Segal, E.; Allis, C. D.; Hake, S. B., Proteome-wide prediction of acetylation substrates. *Proc Natl Acad Sci U S A* **2009,** *106* (33), 13785-90.
117. Matsumoto, S.; Cavadini, S.; Bunker, R. D.; Grand, R. S.; Potenza, A.; Rabl, J.; Yamamoto, J.; Schenk, A. D.; Schubeler, D.; Iwai, S.; Sugasawa, K.; Kurumizaka, H.; Thoma, N. H., DNA damage detection in nucleosomes involves DNA register shifting. *Nature* **2019,** *571* (7763), 79-84.
118. Frouws, T. D.; Barth, P. D.; Richmond, T. J., Site-Specific Disulfide Crosslinked Nucleosomes with Enhanced Stability. *J Mol Biol* **2018,** *430* (1), 45-57.
119. Nowotny, M.; Gaidamakov, S. A.; Crouch, R. J.; Yang, W., Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* **2005,** *121* (7), 1005-16.
120. Zheng, J. S.; Tang, S.; Qi, Y. K.; Wang, Z. P.; Liu, L., Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. *Nat Protoc* **2013,** *8* (12), 2483-95.
121. Raufman, J. P., Bioactive peptides from lizard venoms. *Regul Pept* **1996,** *61* (1), 1-18.
122. Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G., A genetically encoded photocaged amino acid. *J Am Chem Soc* **2004,** *126* (44), 14306-7.
123. Nguyen, D. P.; Mahesh, M.; Elsasser, S. J.; Hancock, S. M.; Uttamapinant, C.; Chin, J. W., Genetic encoding of photocaged cysteine allows photoactivation of TEV protease in live mammalian cells. *J Am Chem Soc* **2014,** *136* (6), 2240-3.

124. Gates, B., Responding to Covid-19 - A Once-in-a-Century Pandemic? *N Engl J Med* **2020,** *382* (18), 1677-1679.
125. Morens, D. M.; Daszak, P.; Taubenberger, J. K., Escaping Pandora's Box - Another Novel Coronavirus. *N Engl J Med* **2020,** *382* (14), 1293-1295.
126. World Health Organization WHO Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/ (accessed 2020, May 13).
127. Kissler, S. M.; Tedijanto, C.; Goldstein, E.; Grad, Y. H.; Lipsitch, M., Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **2020,** *368* (6493), 860-868.
128. Morse, J. S.; Lalonde, T.; Xu, S.; Liu, W. R., Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV. *Chembiochem : a European journal of chemical biology* **2020,** *21* (5), 730-738.
129. Beigel, J. H.; Tomashek, K. M.; Dodd, L. E.; Mehta, A. K.; Zingman, B. S.; Kalil, A. C.; Hohmann, E.; Chu, H. Y.; Luetkemeyer, A.; Kline, S.; Lopez de Castilla, D.; Finberg, R. W.; Dierberg, K.; Tapson, V.; Hsieh, L.; Patterson, T. F.; Paredes, R.; Sweeney, D. A.; Short, W. R.; Touloumi, G.; Lye, D. C.; Ohmagari, N.; Oh, M. D.; Ruiz-Palacios, G. M.; Benfield, T.; Fatkenheuer, G.; Kortepeter, M. G.; Atmar, R. L.; Creech, C. B.; Lundgren, J.; Babiker, A. G.; Pett, S.; Neaton, J. D.; Burgess, T. H.; Bonnett, T.; Green, M.; Makowski, M.; Osinusi, A.; Nayak, S.; Lane, H. C.; Members, A.-S. G., Remdesivir for the Treatment of Covid-19 - Preliminary Report. *N Engl J Med* **2020**.
130. Zhou, N.; Pan, T.; Zhang, J.; Li, Q.; Zhang, X.; Bai, C.; Huang, F.; Peng, T.; Zhang, J.; Liu, C.; Tao, L.; Zhang, H., Glycopeptide Antibiotics Potently Inhibit Cathepsin L in the Late Endosome/Lysosome and Block the Entry of Ebola Virus, Middle East Respiratory Syndrome Coronavirus (MERS-CoV), and Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV). *J Biol Chem* **2016,** *291* (17), 9218-32.
131. Strating, J. R.; van der Linden, L.; Albulescu, L.; Bigay, J.; Arita, M.; Delang, L.; Leyssen, P.; van der Schaar, H. M.; Lanke, K. H.; Thibaut, H. J.; Ulferts, R.; Drin, G.; Schlinck, N.; Wubbolts, R. W.; Sever, N.; Head, S. A.; Liu, J. O.; Beachy, P. A.; De Matteis, M. A.; Shair, M. D.; Olkkonen, V. M.; Neyts, J.; van Kuppeveld, F. J., Itraconazole inhibits enterovirus replication by targeting the oxysterol-binding protein. *Cell Rep* **2015,** *10* (4), 600-15.
132. Mastrangelo, E.; Pezzullo, M.; De Burghgraeve, T.; Kaptein, S.; Pastorino, B.; Dallmeier, K.; de Lamballerie, X.; Neyts, J.; Hanson, A. M.; Frick, D. N.; Bolognesi, M.; Milani, M., Ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity: new prospects for an old drug. *J Antimicrob Chemother* **2012,** *67* (8), 1884-94.
133. Rossignol, J. F., Nitazoxanide: a first-in-class broad-spectrum antiviral agent. *Antiviral Res* **2014,** *110*, 94-103.
134. Mercorelli, B.; Palu, G.; Loregian, A., Drug Repurposing for Viral Infectious Diseases: How Far Are We? *Trends Microbiol* **2018,** *26* (10), 865-876.
135. Hung, I. F.; Lung, K. C.; Tso, E. Y.; Liu, R.; Chung, T. W.; Chu, M. Y.; Ng, Y. Y.; Lo, J.; Chan, J.; Tam, A. R.; Shum, H. P.; Chan, V.; Wu, A. K.; Sin, K. M.;

Leung, W. S.; Law, W. L.; Lung, D. C.; Sin, S.; Yeung, P.; Yip, C. C.; Zhang, R. R.; Fung, A. Y.; Yan, E. Y.; Leung, K. H.; Ip, J. D.; Chu, A. W.; Chan, W. M.; Ng, A. C.; Lee, R.; Fung, K.; Yeung, A.; Wu, T. C.; Chan, J. W.; Yan, W. W.; Chan, W. M.; Chan, J. F.; Lie, A. K.; Tsang, O. T.; Cheng, V. C.; Que, T. L.; Lau, C. S.; Chan, K. H.; To, K. K.; Yuen, K. Y., Triple combination of interferon beta-1b, lopinavir-ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label, randomised, phase 2 trial. *Lancet* **2020,** *395* (10238), 1695-1704.
136. Baez-Santos, Y. M.; St John, S. E.; Mesecar, A. D., The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antivir Res* **2015,** *115*, 21-38.
137. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020,** *582* (7811), 289-293.
138. Nguyen, D. D.; Gao, K.; Chen, J.; Wang, R.; Wei, G. W., Potentially highly potent drugs for 2019-nCoV. *bioRxiv* **2020**, 2020.02.05.936013.
139. Xu, Z.; Peng, C.; Shi, Y.; Zhu, Z.; Mu, K.; Wang, X.; Zhu, W., Nelfinavir was predicted to be a potential inhibitor of 2019-nCov main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation. *bioRxiv* **2020**, 2020.01.27.921627.
140. Wu, C.; Liu, Y.; Yang, Y.; Zhang, P.; Zhong, W.; Wang, Y.; Wang, Q.; Xu, Y.; Li, M.; Li, X.; Zheng, M.; Chen, L.; Li, H., Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm Sin B* **2020,** *10* (5), 766-788.
141. Sadek, H.; Ahmed, M.; Wang, P.; Farag, A., Identification of FDA Approved Drugs Targeting COVID-19 Virus by Structure-Based Drug Repositioning. *ChemRxiv* **2020,** *10.26434/chemrxiv.12003930.v3*.
142. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020,** *10.1038/s41586-020-2223-y*.
143. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* **2020,** *368* (6489), 409-412.
144. Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S. H., An Overview of Severe Acute Respiratory Syndrome-Coronavirus (SARS-CoV) 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy. *J Med Chem* **2016,** *59* (14), 6595-628.
145. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking

with selective receptor flexibility. *Journal of computational chemistry* **2009,** *30* (16), 2785-91.
146. Wang, Y. S.; Fang, X.; Chen, H. Y.; Wu, B.; Wang, Z. U.; Hilty, C.; Liu, W. R., Genetic incorporation of twelve meta-substituted phenylalanine derivatives using a single pyrrolysyl-tRNA synthetase mutant. *ACS chemical biology* **2013,** *8* (2), 405-15.
147. Pedelacq, J. D.; Cabantous, S.; Tran, T.; Terwilliger, T. C.; Waldo, G. S., Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* **2006,** *24* (1), 79-88.
148. Phan, J.; Zdanov, A.; Evdokimov, A. G.; Tropea, J. E.; Peters, H. K., 3rd; Kapust, R. B.; Li, M.; Wlodawer, A.; Waugh, D. S., Structural basis for the substrate specificity of tobacco etch virus protease. *J Biol Chem* **2002,** *277* (52), 50564-72.
149. Rut, W.; Groborz, K.; Zhang, L.; Sun, X.; Zmudzinski, M.; Hilgenfeld, R.; Drag, M., Substrate specificity profiling of SARS-CoV-2 M^{pro} protease provides basis for anti-COVID-19 drug design. *bioRxiv* **2020**, 2020.03.07.981928.
150. Rut, W.; Groborz, K.; Zhang, L.; Sun, X.; Zmudzinski, M.; Pawlik, B.; Młynarski, W.; Hilgenfeld, R.; Drag, M., Substrate specificity profiling of SARS-CoV-2 M^{pro} protease provides basis for anti-COVID-19 drug design. *bioRxiv* **2020**, 2020.03.07.981928.
151. Dragovich, P. S.; Prins, T. J.; Zhou, R.; Webber, S. E.; Marakovits, J. T.; Fuhrman, S. A.; Patick, A. K.; Matthews, D. A.; Lee, C. A.; Ford, C. E.; Burke, B. J.; Rejto, P. A.; Hendrickson, T. F.; Tuntland, T.; Brown, E. L.; Meador, J. W., 3rd; Ferre, R. A.; Harr, J. E.; Kosa, M. B.; Worland, S. T., Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 4. Incorporation of P1 lactam moieties as L-glutamine replacements. *Journal of medicinal chemistry* **1999,** *42* (7), 1213-24.
152. Cao, B.; Wang, Y.; Wen, D.; Liu, W.; Wang, J.; Fan, G.; Ruan, L.; Song, B.; Cai, Y.; Wei, M.; Li, X.; Xia, J.; Chen, N.; Xiang, J.; Yu, T.; Bai, T.; Xie, X.; Zhang, L.; Li, C.; Yuan, Y.; Chen, H.; Li, H.; Huang, H.; Tu, S.; Gong, F.; Liu, Y.; Wei, Y.; Dong, C.; Zhou, F.; Gu, X.; Xu, J.; Liu, Z.; Zhang, Y.; Li, H.; Shang, L.; Wang, K.; Li, K.; Zhou, X.; Dong, X.; Qu, Z.; Lu, S.; Hu, X.; Ruan, S.; Luo, S.; Wu, J.; Peng, L.; Cheng, F.; Pan, L.; Zou, J.; Jia, C.; Wang, J.; Liu, X.; Wang, S.; Wu, X.; Ge, Q.; He, J.; Zhan, H.; Qiu, F.; Guo, L.; Huang, C.; Jaki, T.; Hayden, F. G.; Horby, P. W.; Zhang, D.; Wang, C., A Trial of Lopinavir-Ritonavir in Adults Hospitalized with Severe Covid-19. *N Engl J Med* **2020,** *382* (19), 1787-1799.
153. Musarrat, F.; Chouljenko, V.; Dahal, A.; Nabi, R.; Chouljenko, T.; Jois, S. D.; Kousoulas, K. G., The anti-HIV drug nelfinavir mesylate (Viracept) is a potent inhibitor of cell fusion caused by the SARSCoV-2 spike (S) glycoprotein warranting further evaluation as an antiviral against COVID-19 infections. *J Med Virol* **2020,** *10.1002/jmv.25985*.
154. Yamamoto, N.; Matsuyama, S.; Hoshino, T.; Yamamoto, N., Nelfinavir inhibits replication of severe acute respiratory syndrome coronavirus 2 in vitro. *BioRxiv* **2020,** *10.1101/2020.04.06.026476*.

155. Shapiro, W.; Dibianco, R.; Thadani, U., Comparative Efficacy of 200, 300 and 400 Mg of Bepridil for Chronic Stable Angina-Pectoris. *Am J Cardiol* **1985,** *55* (7), C36-C42.
156. Shoichet, B. K., Interpreting steep dose-response curves in early inhibitor discovery. *J Med Chem* **2006,** *49* (25), 7274-7.
157. Yang, N.; Shen, H. M., Targeting the Endocytic Pathway and Autophagy Process as a Novel Therapeutic Strategy in COVID-19. *Int J Biol Sci* **2020,** *16* (10), 1724-1731.
158. Savarino, A.; Boelaert, J. R.; Cassone, A.; Majori, G.; Cauda, R., Effects of chloroquine on viral infections: an old drug against today's diseases? *Lancet Infect Dis* **2003,** *3* (11), 722-7.
159. Geleris, J.; Sun, Y.; Platt, J.; Zucker, J.; Baldwin, M.; Hripcsak, G.; Labella, A.; Manson, D. K.; Kubin, C.; Barr, R. G.; Sobieszczyk, M. E.; Schluger, N. W., Observational Study of Hydroxychloroquine in Hospitalized Patients with Covid-19. *N Engl J Med* **2020,** *382* (25), 2411-2418.
160. Chen, Z.; Hu, J.; Zhang, Z.; Jiang, S.; Han, S.; Yan, D.; Zhuang, R.; Hu, B.; Zhang, Z., Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial. *medRxiv* **2020**, 2020.03.22.20040758.
161. CHEN Jun, L. D., LIU Li, LIU Ping, XU Qingnian, XIA Lu, LING Yun, HUANG Dan, SONG Shuli, ZHANG Dandan, QIAN Zhiping, LI Tao, SHEN Yinzhong, LU Hongzhou, A pilot study of hydroxychloroquine in treatment of patients with moderate COVID-19. *J Zhejiang Univ (Med Sci)* **2020,** *49* (2), 215-219.
162. Cavalcanti, A. B.; Zampieri, F. G.; Rosa, R. G.; Azevedo, L. C. P.; Veiga, V. C.; Avezum, A.; Damiani, L. P.; Marcadenti, A.; Kawano-Dourado, L.; Lisboa, T.; Junqueira, D. L. M.; de Barros, E. S. P. G. M.; Tramujas, L.; Abreu-Silva, E. O.; Laranjeira, L. N.; Soares, A. T.; Echenique, L. S.; Pereira, A. J.; Freitas, F. G. R.; Gebara, O. C. E.; Dantas, V. C. S.; Furtado, R. H. M.; Milan, E. P.; Golin, N. A.; Cardoso, F. F.; Maia, I. S.; Hoffmann Filho, C. R.; Kormann, A. P. M.; Amazonas, R. B.; Bocchi de Oliveira, M. F.; Serpa-Neto, A.; Falavigna, M.; Lopes, R. D.; Machado, F. R.; Berwanger, O.; Coalition Covid-19 Brazil, I. I., Hydroxychloroquine with or without Azithromycin in Mild-to-Moderate Covid-19. *N Engl J Med* **2020**.
163. Ohkuma, S.; Poole, B., Cytoplasmic vacuolation of mouse peritoneal macrophages and the uptake into lysosomes of weakly basic substances. *The Journal of cell biology* **1981,** *90* (3), 656-64.
164. DeWald, L. E.; Dyall, J.; Sword, J. M.; Torzewski, L.; Zhou, H.; Postnikova, E.; Kollins, E.; Alexander, I.; Gross, R.; Cong, Y.; Gerhardt, D. M.; Johnson, R. F.; Olinger, G. G., Jr.; Holbrook, M. R.; Hensley, L. E.; Jahrling, P. B., The Calcium Channel Blocker Bepridil Demonstrates Efficacy in the Murine Model of Marburg Virus Disease. *J Infect Dis* **2018,** *218* (suppl_5), S588-S591.
165. Wu, W. N.; Pritchard, J. F.; Ng, K. T.; Hills, J. F.; Uetz, J. A.; Yorgey, K. A.; McKown, L. A.; O'Neill, P. J., Disposition of bepridil in laboratory animals and man. *Xenobiotica* **1992,** *22* (2), 153-69.
166. Mitterreiter, S.; Page, R. M.; Kamp, F.; Hopson, J.; Winkler, E.; Ha, H. R.; Hamid, R.; Herms, J.; Mayer, T. U.; Nelson, D. J.; Steiner, H.; Stahl, T.; Zeitschel, U.; Rossner, S.; Haass, C.; Lichtenthaler, S. F., Bepridil and amiodarone

simultaneously target the Alzheimer's disease beta- and gamma-secretase via distinct mechanisms. *J Neurosci* **2010,** *30* (26), 8974-83.
167. Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G., Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res* **2020,** *30* (3), 269-271.
168. Hollingshead, L. M.; Faulds, D.; Fitton, A., Bepridil. A review of its pharmacological properties and therapeutic use in stable angina pectoris. *Drugs* **1992,** *44* (5), 835-57.
169. Somberg, J.; Torres, V.; Flowers, D.; Miura, D.; Butler, B.; Gottlieb, S., Prolongation of QT interval and antiarrhythmic action of bepridil. *Am Heart J* **1985,** *109* (1), 19-27.
170. Behr, E. R.; Roden, D., Drug-induced arrhythmia: pharmacogenomic prescribing? *Eur Heart J* **2013,** *34* (2), 89-95.
171. Agrawal, A. S.; Ying, T.; Tao, X.; Garron, T.; Algaissi, A.; Wang, Y.; Wang, L.; Peng, B. H.; Jiang, S.; Dimitrov, D. S.; Tseng, C. T., Passive Transfer of A Germline-like Neutralizing Human Monoclonal Antibody Protects Transgenic Mice Against Lethal Middle East Respiratory Syndrome Coronavirus Infection. *Scientific reports* **2016,** *6*, 31629.
172. Du, L.; Kou, Z.; Ma, C.; Tao, X.; Wang, L.; Zhao, G.; Chen, Y.; Yu, F.; Tseng, C. T.; Zhou, Y.; Jiang, S., A truncated receptor-binding domain of MERS-CoV spike protein potently inhibits MERS-CoV infection and induces strong neutralizing antibody responses: implication for developing therapeutics and vaccines. *PloS one* **2013,** *8* (12), e81587.
173. Du, L.; Zhao, G.; Yang, Y.; Qiu, H.; Wang, L.; Kou, Z.; Tao, X.; Yu, H.; Sun, S.; Tseng, C. T.; Jiang, S.; Li, F.; Zhou, Y., A conformation-dependent neutralizing monoclonal antibody specifically targeting receptor-binding domain in Middle East respiratory syndrome coronavirus spike protein. *J Virol* **2014,** *88* (12), 7045-53.
174. Li, W.; Schafer, A.; Kulkarni, S. S.; Liu, X.; Martinez, D. R.; Chen, C.; Sun, Z.; Leist, S. R.; Drelich, A.; Zhang, L.; Ura, M. L.; Berezuk, A.; Chittori, S.; Leopold, K.; Mannar, D.; Srivastava, S. S.; Zhu, X.; Peterson, E. C.; Tseng, C. T.; Mellors, J. W.; Falzarano, D.; Subramaniam, S.; Baric, R. S.; Dimitrov, D. S., High Potency of a Bivalent Human VH Domain in SARS-CoV-2 Animal Models. *Cell* **2020,** *183* (2), 429-441 e16.
175. Sarlauskas, J.; Dickancaite, E.; Nemeikaite, A.; Anusevicius, Z.; Nivinskas, H.; Segura-Aguilar, J.; Cenas, N., Nitrobenzimidazoles as substrates for DT-diaphorase and redox cycling compounds: their enzymatic reactions and cytotoxicity. *Arch Biochem Biophys* **1997,** *346* (2), 219-29.
176. Knox, R. J.; Chen, S., Quinone reductase-mediated nitro-reduction: clinical applications. *Methods Enzymol* **2004,** *382*, 194-221.
177. Nivinskas, H.; Staskeviciene, S.; Sarlauskas, J.; Koder, R. L.; Miller, A. F.; Cenas, N., Two-electron reduction of quinones by Enterobacter cloacae NAD(P)H:nitroreductase: quantitative structure-activity relationships. *Arch Biochem Biophys* **2002,** *403* (2), 249-58.

178. Wardman, P., Some reactions and properties of nitro radical-anions important in biology and medicine. *Environ Health Perspect* **1985,** *64*, 309-20.
179. Tercel, M.; Yang, S.; Atwell, G. J.; Smith, E.; Gu, Y.; Anderson, R. F.; Denny, W. A.; Wilson, W. R.; Pruijn, F. B., Hypoxic selectivity and solubility--investigating the properties of A-ring substituted nitro seco-1,2,9,9a-tetrahydrocyclopropa[c]benz[e]indol-4-ones (nitroCBIs) as hypoxia-activated prodrugs for antitumor therapy. *Bioorg Med Chem* **2010,** *18* (14), 4997-5006.
180. Race, P. R.; Lovering, A. L.; Green, R. M.; Ossor, A.; White, S. A.; Searle, P. F.; Wrighton, C. J.; Hyde, E. I., Structural and mechanistic studies of Escherichia coli nitroreductase with the antibiotic nitrofurazone. Reversed binding orientations in different redox states of the enzyme. *J Biol Chem* **2005,** *280* (14), 13256-64.
181. Nivinskas, H.; Koder, R. L.; Anusevicius, Z.; Sarlauskas, J.; Miller, A. F.; Cenas, N., Quantitative structure-activity relationships in two-electron reduction of nitroaromatic compounds by Enterobacter cloacae NAD(P)H:nitroreductase. *Arch Biochem Biophys* **2001,** *385* (1), 170-8.
182. Salter-Blanc, A. J.; Bylaska, E. J.; Johnston, H. J.; Tratnyek, P. G., Predicting reduction rates of energetic nitroaromatic compounds using calculated one-electron reduction potentials. *Environ Sci Technol* **2015,** *49* (6), 3778-86.

APPENDIX A: LIST OF ABBREVIATIONS

Allylamine	Aa
Activated Cysteine-Directed Protein Ligation	ACPL
7-amino-4-methylcoumarin	AMC
angiotensin-converting enzyme 2	ACE2
acetonitrile	ACN
(tert-Butoxycarbonyl)	Boc
carboxybenzyl	CBZ
cyano	CN
Cytochrome p450 oxidoreductase	CYPOR
dalton	Da
N,N′-Dicyclohexylcarbodiimide	DCC
dichloromethane	DCM
N,N-Dimethylformamide	DMF
dimethyl sulfoxide	DMSO
1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide	EDCI

Ethylenediaminetetraacetic acid	EDTA
electrospray ionization-mass spectrometry	ESI-MS
ethyl acetate	EtOAc
Fluorenylmethyloxycarbonyl	Fmoc
Hydrazine	Ha
hydrogen chloride	HCl
high pressure liquid chromatography	HPLC
half maximal inhibitory concentration	IC50
Hypoxia activated prodrug	HAP
iso-propanol	*i*-PrOH
kilodalton	kDa
Liquid chromatography–mass spectrometry	LC-MS
methanol	MeOH
main protease inhibitor	MPI
main protease	M^{pro}
messenger RNA	mRNA
Sodium hydride	NaH

Sodium bicarbonate	$NaHCO_3$
Non-canonical amino acdis	ncAA
Nuclear magnetic resonance	NMR
2-nitro-5-thiocyanobenzoic acid	NTCB
Propargylamine	Pa
Polyacrylamide gel electrophoresis	PAGE
protein data bank	PDB
round-bottom	RB
ribonucleic acid	RNA
sodium dodecyl sulphate–polyacrylamide gel electrophoresis	SDS-PAGE
Small Ubiquitin-like Modifier	SUMO
Solid Phase Peptide Synthesis	SPPS
tert-butyl	tBu
trifluoroacetic acid	TFA
tetrahydrofuran	THF
thin layer chromatography	TLC

Tris(2-carboxyethyl)phosphine	TCEP
tris(hydroxymethyl)aminomethane	Tris
transfer RNA	tRNA
Ubiquitin	Ub