UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE QUÍMICA DE SÃO CARLOS

Medicinal and Biological Chemistry Group - NEQUIMED/IQSC/USP

Fernanda dos Reis Rocho

Investigating the specificity and inhibition of SARS-CoV-2 main protease and its variants

SÃO CARLOS

2023

FERNANDA DOS REIS ROCHO

Investigating the specificity and inhibition of SARS-CoV-2 main protease and its variants

Thesis submitted to the Instituto de Química de São Carlos da Universidade de São Paulo in partial fulfillment of the requirements for the degree of Doctor of Science.

Concentration area: Organic and Biological Chemistry

Supervisor: Prof. Carlos A. Montanari, Ph.D.

SÃO CARLOS

2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico para fins de estudo e pesquisa, desde que citada a fonte.

Assinatura:

Fernanda dos R. Rocho

Data:
13/12/2023

Acknowledgments

First and foremost, I'd like to extend my heartfelt gratitude to my family for their unwavering support and encouragement throughout my journey. Your belief in me has been a constant source of inspiration, and I hope that I am making each one of you immensely proud.

I am deeply indebted to my advisor, Prof. Ph.D. Carlos Alberto Montanari, for an incredible journey of six years working together. Your mentorship, and unwavering support have been instrumental in shaping my academic and professional growth. I am very thankful for the knowledge that both Prof. Ph.D. Montanari and Prof. Ph.D. Andrei Leitão have generously shared with me throughout the years.

I also want to thank Prof. Ph.D. Guy Salvesen for all his "Guydance" and mentorship in his lab in San Diego. I have the most amazing time working and learning with you. Go Wave!! And also, I would like to thank Scott Snipas, Guy's lab manager, for all the kinetics fun and farming. You both turned into incredible friends and I'm very grateful to have the opportunity to work you.

A special acknowledgment to Ph.D. Fabiana Rosini, our laboratory technician, whose assistance throughout my masters and PhD has been fundamental to my research. To my friends Ph.D. Pedro Henrique Jatai Batista, Gustavo Tateyama, MS Priscila Felício, MS Johnatan Mucelini, Ph.D. Vinicius Bonatto, Rafael Janke, Ph.D. Isis Arantes, Ph.D. Felipe Canisares, MS Sabrina del Sarto, MS Isabella Villarinho, and Ph.D. Thiago Kelvin, I extend my heartfelt thanks. Your companionship, shared laughter, and advice have made this journey exceptionally smooth. Above all, I am deeply grateful for your unwavering support and for tolerating my moments of drama and complaints. I cherish each one of you.

I would also like to thank the XFR Taekwondo team, especially my friend and teacher Pedro Colloca, whose support was fundamental during the most difficult time of my PhD. I'm incredible thankful for all the lessons, and specially, for all the kicks!

A special thanks to Ph.D. Marcin Drag who provided me with the M^{pro}'s substrate, and Ph.D. Rich Williams for the inhibitors.

My heartfelt gratitude extends to my dedicated coworkers within the NEQUIMED group, particularly Ph.D. Anwar Shamim, and Ph.D. Felipe Martins, who played pivotal roles in the synthesis of the compounds. Finally, I am immensely thankful to CAPES for providing me with a scholarship, without which this work would not have been possible. I also extend my gratitude to IQSC and Fapesp for their support, and the Fulbright program for my USA scholarship.

"Quero muito agradecer a mim, porque eu não desisti."
(Larissa de Macedo Machado, 2019)

RESUMO

A protease principal do vírus SARS-CoV-2 (M^{pro}) é um alvo importante no combate à COVID-19, no entanto sua eficácia pode ser prejudicada por mutações virais. A expressão e eficiência catalítica da WT M^{pro} e algumas variantes de preocupação (VOC) foram avaliadas usando SDS-PAGE e cinética enzimática. Foi identificado que diferentes sais cosmotrópicos, especialmente o citrato de sódio, melhora significativamente a atividade catalítica da protease, estabilizando sua forma ativa dimérica. Estudos com as diferentes variantes demonstrou que as mesmas possuem um impacto mínimo no reconhecimento de substratos distintos, sugerindo que diferentes inibidores podem ser eficazes contra todas as variantes. O fato de mutações serem infrequentes na sequência da M^{pro} e não alterarem o sítio ativo da proteína, destaca a sua importância contínua como alvo terapêutico. O Neq1183 (Nirmatrelvir) e inibidores peptídeo miméticos foram testados contra a WT M^{pro} e VOC, e suas respectivas constantes cinéticas foram determinadas. Portanto, por meio da análise de pares moleculares (MMP) e relação estrutura e atividade (SAR), identificou-se que grupos volumosos e aromáticos na posição P2 da molécula influenciam significativamente a afinidade contra a M^{pro}. O Neq1183 também foi estudado usando a técnica de calorimetria de titulação isotérmica (ITC) contra a WT M^{pro}, revelando um perfil termodinâmico favorável de interação. Diferentes warheads irreversíveis foram analisadas, apresentando potencial para o desenvolvimento de sondas baseadas em atividade (ABP). Além disso, considerando o papel da catepsina humana L (hCatL) na COVID-19, a inibição desta CP foi analisada com uma variedade de inibidores covalentes reversíveis, na qual foi identificado que o grupo indol na posição P2 da molécula é importante para atingir alta afinidade e seletividade nesta enzima frente as catepsinas humanas S e B. Logo, este estudo fornece uma compreensão abrangente da especificidade da SARS-CoV-2 M^{pro}, oferecendo insights para o desenvolvimento de inibidores.

Palavras-chave: Cisteíno proteases, SARS CoV-2 M^{pro}, Catepsina L humana

ABSTRACT

The SARS-CoV-2 Main Protease (M^{pro}) is an essential target for combating COVID-19, but viral mutations can hamper its effectiveness. Expression and catalytic efficiency of the wild-type (WT) SARS-CoV-2 M^{pro} and some variants of concern (VOC) are assessed using SDS-PAGE electrophoresis and enzyme kinetics. Kosmotropes, particularly sodium citrate, significantly enhance the catalytic activity of M^{pro} by stabilizing its active dimeric form. The WT M^{pro} and variants study using a peptide small molecule and a protein-based substrate, indicate minimal impact on substrate recognition, suggesting that existing inhibitors remain effective against all variants. The infrequent mutations in M^{pro}'s sequence and do not alter the active site of the protein, highlights its continued importance as a therapeutic target. Neq1183 (Nirmatrelvir) and peptide mimetic inhibitors were tested against both WT and VOC M^{pro} and their respective kinetic constants were determined. Using matched-molecular pair and structure-activity relationship (SAR), it was identified that bulky and aromatic groups at the P2 position significantly influenced affinity. Neq1183 was also investigated using isothermal titration calorimetry (ITC) against WT M^{pro}, revealing a favorable thermodynamic profile. Irreversible warheads were analyzed, presenting the potential for activity-based probe development. Additionally, considering the human cathepsin L (hCatL) role in COVID-19, selective inhibition of this CP was achieved with a range of covalent reversible inhibitors, revealing an essential indole group at the P2 position. Furthermore, this study provides a comprehensive understanding of SARS-CoV-2 M^{pro} specificity, offering insights for inhibitor development.

FIGURE LIST

TABLE LIST

EQUATION LIST

SUMMARY

1 INTRODUCTION

Proteases (EC 3.4), or peptidases, represent a diverse and vital class of enzymes in the biological world. These enzymes play a fundamental role in regulating numerous physiological processes, making them of paramount importance as biological targets for both scientific research and therapeutic intervention. Proteases are responsible for the precise cleavage of peptide bonds within proteins and peptides. The cleavage occurs through a nucleophilic attack performed by amino acid residues at the enzyme's catalytic site [1,2]. This catalytic activity allows them to modulate various cellular functions, including protein turnover, signaling cascades, immune responses, blood clotting, and digestion [3,4]. Due to their essential roles, proteases have emerged as attractive targets for drug development and intervention in various diseases such as cancer, cardiovascular disorders, and Alzheimer's disease, as well as for viral, bacterial, and infectious diseases [4–7].

The spectrum of protease specificity is broad, ranging from nonselective, as seen in the proteasome or lysosomal cathepsins, to high specificity such as caspases, which cleave exclusively after Asp residues [8]. The proteases possess specific subsite pockets that will recognize the corresponding positions of the substrate structure (Fig. 1). Because of this subsite, small molecule inhibitors are usually designed to mimic the substrate structure of the protein. Hence, proteases can execute their substrate cleavage within the polypeptide chain, categorizing them as endopeptidases or, at the termini, classifying them as carboxypeptidases and aminopeptidases [3].

Figure 1: Schechter-Berger nomenclature for protease recognition of substrates. P1, P2, P3, and P4 are the substrate's structure that will interact with the corresponding subsite pocket of the protease.
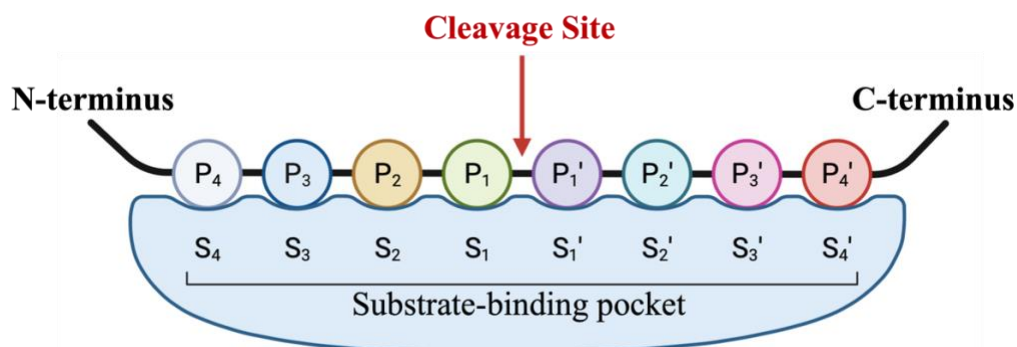
Source: Adapted from Schechter I & Berger A. On the size of the active site in proteases. I. Papain. **Biochem Biophys Res Commun,** v. 27, p. 157–162, 1967

The proteases are broadly classified into several families based on their catalytic mechanisms. The prominent protease families include serine proteases, cysteine proteases,

aspartic proteases, metalloproteases, and threonine proteases. Within each family, one can find multiple protease enzymes, each characterized by unique substrate specificities and functions.

In this work, different cysteine proteases (CPs) were studied, particularly emphasizing the main protease (M^{pro}) of the SARS-CoV-2 virus. CPs are distinguished by a nucleophilic cysteine residue (Cys) responsible for their catalytic activity [2], targeted for potential covalent inhibition. CPs are found in every living organism and are divided into clans, from CA to CD. The most abundant CPs belonged to the papain family (CA clan) due to a high homology shared with the papain (plant protease) structure [10,11]. Herein, SARS-CoV-2 M^{pro}, the causative agent of COVID-19, will be studied regarding its specificity towards different substrates and inhibition with a range of covalent reversible inhibitors. Furthermore, the human cathepsins L, S, and B were also investigated amid selectivity inhibition of these CPs.

1.1 The SARS-CoV-2 Main Protease (M^{pro})

The M^{pro} (or $3CL^{pro}$)[11] (EC 3.4.22.69) is the major cysteine protease found in the SARS-CoV-2 β-coronavirus, the etiological agent of the coronavirus disease-19 (COVID-19) that caused a global pandemic in march 2020 [5]. Human coronaviruses (HCoVs) represent a group of positively stranded RNA viruses responsible for a significant portion of upper respiratory tract infections in humans. Within the HCoV genome, there are multiple nonstructural proteins that are synthesized as two extensive polyproteins known as pp1a and pp1ab. These polyproteins undergo cleavage mediated by viral proteases, named papain-like protease (PL^{pro}) and M^{pro} [12]. The processing of these polyproteins involves the generation of 16 individual non-structural proteins (nsps). These nsps play a crucial role in forming the RNA genome replication and mRNA transcription complexes, collectively known as the Replication/Transcription Complex [5,13,14].

Within this context, M^{pro}, encoded by nsp5, plays a pivotal role in its own liberation from the polyprotein through an autoproteolytic mechanism [15]. M^{pro} exhibits the capability to cleave at ten distinct sites situated between nsp6 and nsp16, ultimately facilitating the release of mature nsp proteins. Thus, M^{pro} is the best-characterized drug target of the HCoVs due to its role in virus replication and structural conservation [16]. Additionally, it has been reported that M^{pro} plays a crucial role in blocking the recognition of host sensors to deceive the immune system [17].

M^{pro} features a three-domain structure, with the domains I and II forming a chymotrypsin-like fold where the catalytic residues are present in a cleft between the two

domains [13,18]. Domain III comprises a globular antiparallel α-helical cluster, a unique feature required for homodimer formation (Fig. 2) [13,16]. Unlike other cysteine proteases (CPs), M^{pro} does not have the conventional catalytic triad Cys(Ser)-His-Asp(Glu). Instead, M^{pro} has a catalytic dyad formed by Cys145 and His41. Studies suggest a water molecule might assist the catalytic reaction through a strong hydrogen bond with His41 stabilized by the conserved residues His164 and Asp187 [19].

Figure 2: SARS-CoV-2 M^{pro} structure. (A) Crystallographic structure of SARS-CoV-2 M^{pro}, the protomers A and B are shown in red and green, respectively, with the catalytic dyad highlighted in blue. (B) Protomer B colored by spectrum (rainbow) in which the color blue indicates the N-terminus of the protein and red the C-terminus, the different domains are labeled in roman numbers. PDB: 6WTM.
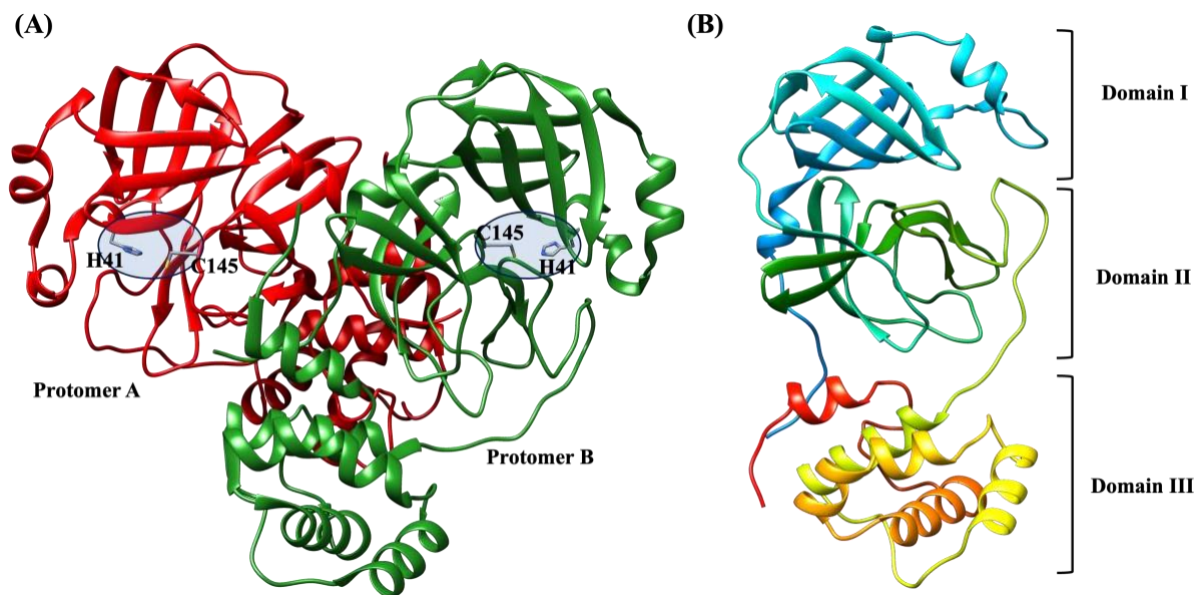
The enzyme exhibits a highly specific cleavage specificity, which preferentially recognizes the Gln residue at P1 (Fig. 3) [5,13,20]. So far, no known human protein exhibits a similar cleavage sequence specificity. Therefore, given its pivotal role in viral replication and the absence of a closely related counterpart in humans, M^{pro} emerges as a highly promising candidate for antiviral drug development. Inhibiting M^{pro} can disrupt the viral life cycle, effectively stopping viral replication and the spread of infection within the host.

Figure 3: Cleavage site specificity of SARS-CoV-2 M^pro. The arrow designates the site of cleavage.
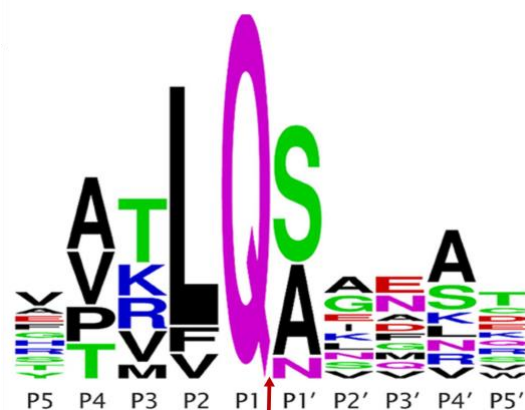
Source: Adapted from Miltner, N., *et al*. Identification of SARS-CoV-2 Main Protease (Mpro) Cleavage Sites Using Two-Dimensional Electrophoresis and In Silico Cleavage Site Prediction. **International Journal of Molecular Science, Washingtons**, v. 24, n. 4, p. 3236, 2023

Researchers worldwide are designing and screening small molecules that can inhibit M^pro's activity [21,22]. In this scenario, researchers from Pfizer, Inc. developed a high affinity inhibitor for M^pro dubbed Nirmatrelvir (Fig. 4) [23]. By the end of 2021, the inhibitor was approved for emergency use to treat COVID-19 with the tradename Paxlovid (Nirmatrelvir + Ritonavir), being the first approved drug to treat COVID-19 (https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-oral-antiviral-treatment-covid-19). Pfizer's compound is a covalent reversible inhibitor that bears a nitrile as a warhead. These findings brought relevance to covalent inhibitors and the nitriles' use as a warhead, with many reviews published on the subject [24–26].

Figure 4: Nirmatrelvir structure with its inhibition constant against SARS-CoV-2 M^pro. The nitrile warhead is highlighted in red.

PF-07321332 (Nirmatrelvir)

$K_i = 3.11$ nM
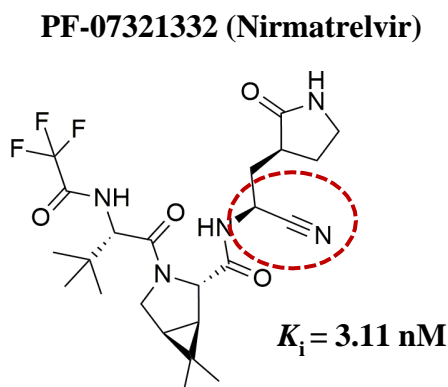
Source: Adapted from Owen, D. R., *et al*., An oral SARS-CoV-2 M^pro inhibitor clinical candidate for the treatment of COVID-19. **Science, Washington**, v. 374, p. 1586–1593, 2021.

The rapid emergence and widespread dissemination of SARS-CoV-2 variants present a challenge in managing the disease. Thus, some variants of concern (VOCs) were studied, with a specific focus on the mutations identified within the M^pro sequence.

1.1.1 SARS-CoV-2 Variants of Concern

The SARS-CoV-2 VOCs are: Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), Omicron BA.1 and BA.2 [27]. The variants exhibit increased transmissibility, heightened disease severity, and, in some cases, resistance to existing vaccines due to genetic mutations within the virus genome, particularly in the spike protein [27]. While the majority of these mutations are located in the spike protein, some are also present in the M^pro sequence, as observed in the case of the beta and omicron VOCs. For the Beta variant, the mutation K90R was observed for strain B.1.351 and B.1.351.3, and the doubled mutation K90R and A193V for strain B.1.351.2. As for Omicron BA.1 and BA.2, the mutation P132H was observed (gisaid.org). Considering these mutations, in addition to the wild-type M^pro, the M^pro from the VOCs Beta B.1.351 and B.1.351.2 (dubbed beta1 and beta 2), and Omicron BA.1 were also investigated. The M^pro derived from VOCs were investigated to determine whether the mutations impact its catalytic activity or substrate specificity. Furthermore, inhibitors were tested against the WT M^pro and variants.

1.2 Human Cathepsin L

Human cathepsin L (*h*CatL) (EC 3.4.22.B49) is a proteolytic enzyme that belongs to the cathepsin family, which shares a catalytic mechanism and homology with the plant protease papain [28,29]. This family of enzymes comprises 11 members that play essential roles in several physiological processes, such as cell death, proliferation, migration, and extracellular matrix (ECM) remodeling [30,31].

The *h*CatL is an endosomal/lysosomal CP ubiquitously expressed in humans [32,33]. The enzyme is synthesized in the cell as a zymogen (procathepsin L). Subsequently, the activation to its mature form occurs either by an aspartyl cathepsin D or autocatalysis in the lysosome [29,34,35]. The *h*CatL structure consists of two domains, with the catalytic triad residues Cys25, His163, and Asn187 found in a cleft between the two domains (Fig. 5) [36]. Its role as an endopeptidase is crucial in the degradation of endocytosed and intracellular proteins.

Additionally, it plays a role in processing hormones and activating latent forms of other proteolytic enzymes[37].

Figure 5: Crystal structure of human cathepsin L1 showing the catalytic triad residues. The binding subsites region is depicted. PDB: 2XU1
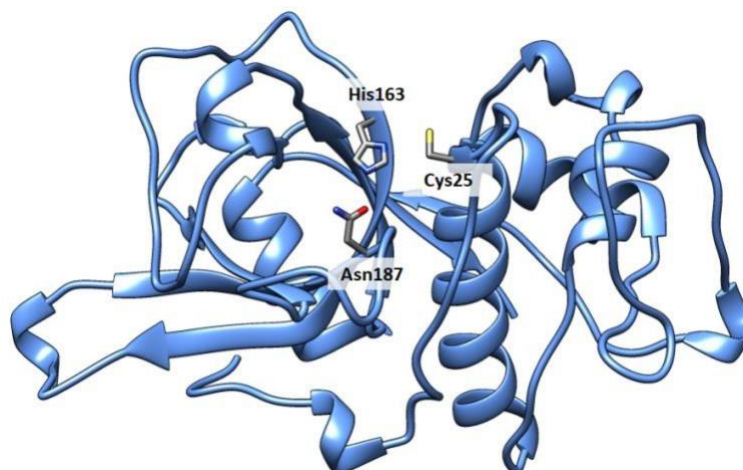
Source: Own author

Recent studies showed that the expression of hCatL is highly dysregulated in several human diseases, like diabetes, abdominal aortic aneurysm, liver fibrosis, different types of cancers, and the COVID-19 disease [38–41]. Due to its involvement in highly invasive forms of cancers, like breast, ovarian, prostate, lung, and colon [32,34,42,43], this enzyme has emerged as an attractive drug target for cancer research.

Currently, hCatL has been gaining prominence due to the SARS-CoV-2 pandemic. It was identified that the protein facilitates the virus's entry into the cell by activating the virus Spike glycoprotein (s) [44,45]. The SARS-CoV-2 virus has two mechanisms for entering human cells: one involves the attachment of the S protein to the host receptor angiotensin-converting enzyme 2 (ACE2), and the second through endocytosis [44,46,47]. Primarily, the S protein activation is facilitated by membrane-bound serine protease TMPRSS2, although human cathepsin L can also activate it [46,48]. When an endosomal pathway occurs, the activation of the S protein happens in acidic pH, a condition in which TMPRSS2 is not catalytically active, only hCatL[48].

Also, hCatL is related to the progression and severity of COVID-19 due to protein overexpression in human cell lines during infection [49]. Recent studies point out the use of hCatL inhibitors as possible antiviral agents. Through cell-based assays, it was identified that these inhibitors could successively block the virus infection[48,50]. Despite the importance of hCatL as a drug target, there are no approved drugs or clinical trials regarding hCatL inhibition. This is

probably due to the enormous difficulty of achieving selectivity between proteins of the cathepsin family [37].

Therefore, the inhibition of *h*CatL by small molecules can be explored as an appropriate strategy for developing novel anticancer and antiviral agents. Thus, this study also focuses on determining inhibition constants for three human cathepsins, seeking to find high affinity and selective inhibitors for *h*CatL towards the human cathepsins S and B.

1.2.1 Human Cathepsin S and B

The cathepsins share a high degree of homology, which makes the design of selective inhibitors very challenging [29,37]. In order to have a high-affinity inhibitor for *h*CatL, the inhibitor must not have an affinity for human cathepsins S and B. The reason is that human cathepsin B has no role in the SARS-CoV-2 entry into the cell, and human cathepsin S is essential for the immunological system [50,51].

Human cathepsin B (hCatB) is a lysosomal enzyme, such as hCatL, fundamentally degrades intracellular proteins. It is considered a crucial tumor promotion factor and, due to this role, is an essential biological target for treating different types of cancer [32,52–54]. *h*CatB has a pH-dependent activity, which regulates its ability to act as endo or carboxypeptidase [55,56]. Due to this structural feature, selectivity between *h*CatL and *h*CatB is more easily achieved.

The human cathepsin S (*h*CatS) is a lysosomal enzyme associated with the regulation of the immune response [57]. It shares a 57% sequence identity with *h*CatL, which makes achieving selectivity between these two enzymes challenging. However, the CPs differ in their preference for neutral to basic pHs [58]. *h*CatS is a target for therapeutic interventions in various diseases, including but not limited to multiple sclerosis [59], psoriasis [60], and obesity [61].

1.3 Enzyme Kinetics

Enzyme-catalyzed reactions can be investigated in numerous ways to explore various aspects of catalysis, with kinetic analysis being the most prevalent approach [62]. Enzyme kinetics studies the reaction rates catalyzed by enzymes, which is crucial to elucidate enzyme mechanisms. In an enzyme-catalyzed reaction, the enzyme (E) binds to the substrate (S), forming an enzyme-substrate complex (ES), with subsequent product formation (P) and the enzyme restitution, best shown in the reaction scheme:
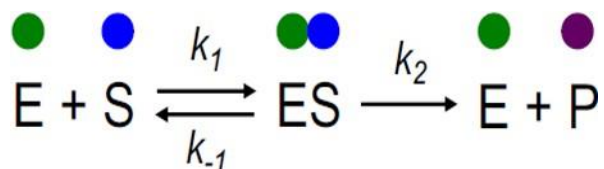
Source: Adapted from Copeland, R. A. *Enzymes: a practical introduction to structure, mechanism, and data analysis*. (Wiley, 2000)

The constants k_1, k_{-1}, and k_2 are related to the formation of the ES complex, ES complex dissociation, and product formation, respectively. This scheme predicts that the reaction rate will be proportionate to the concentration of the ES complex, such that $v = k_2[ES]$ [62].

The rate of an enzyme-catalyzed reaction follows pseudo-first-order kinetics. Initially, reaction velocity depends on the substrate concentration following a first-order reaction. As the substrate saturates the enzyme, the rate becomes independent of substrate concentration, displaying zero-order kinetics. During this phase, the reaction reaches a steady-state, meaning that any further increase in substrate concentration does not alter the reaction velocity [63]. For the steady-state to be achieved, the substrate concentration has to be in excess in the reaction ([S] >> [E]). In the steady-state, the constant k_2 can be denominated k_{cat}, the catalytic constant (turnover number), which defines the maximal velocity at which an enzymatic reaction can proceed with a certain substrate [62]. It is the most consistent mean to compare rates between different enzymatic reactions [62].

In 1913, Leonor Michaelis and Maud Menten developed an equation, based on previous studies by Victor Henri [64], to explain the enzyme kinetics:

Equation 1: Michaelis-Menten equation for an enzyme-catalyzed reaction.

$$v_0 = \frac{V_{max}[S]}{K_M + [S]}$$

The kinetic constant K_M, also known as the Michaelis constant, represent the substrate concentration at which the reaction velocity reaches half of the maximal velocity (V_{max}) obtained under saturating substrate conditions (steady-state). The K_M can also be used as an affinity measure when $k_2 >>> k_{-1}$, where the dissociation of the ES complex to form the product is faster than the reverse reaction (k_{-1}). In these conditions, the lower the K_M value, the greater the enzyme's affinity for the substrate [65]. K_M and $V_{max's}$ are determined through graphical analysis by measuring initial velocities at different substrate concentrations.

1.3.1 Irreversible Inhibitors

The activity of enzymes can be blocked by inhibitors, which are molecules usually similar to their substrates, preventing the enzyme from performing its function. Inhibitors can bind to the enzyme's active site or other locations in the enzyme structure, such as allosteric sites. Inhibitors can be categorized into two classes: reversible and irreversible. The kinetics of irreversible inhibitors differ from reversible inhibitors primarily due to their distinct mechanisms of action. Irreversible inhibitors are compounds that are active site-directed. This can chemically modify the active site, leading to a catalytically inactive enzyme and prolonged or permanent inhibition [66].

Since enzymes and the final complex formed with irreversible inhibitors can never reach equilibrium, the most suitable approach for assessing the potency of such inhibitors is through kinetic methods. IC_{50} values obtained for irreversible inhibitors lack significance. Therefore, to assess the inhibitor's potency, the reaction's k_{obs} must be determined first (Eq. 2) [66]. The observed rate constant for inhibition (k_{obs}) is a pseudo-first-order rate constant derived from the product formation, which can be determined by fitting the kinetic data to the following equation:

Equation 2: Determination of k_{obs} for irreversible inhibitors. P is the product formation, v_i is the initial reaction rate, and t is time.

$$[P] = \frac{v_i}{k_{obs}} [1 - e^{-k_{obs}t}]$$

A plot of k_{obs} versus [I] can be done by determining kobs for each inhibitor concentration. The relationship between k_{obs} and [I] reveals the mechanism of inactivation exhibited by the compound. A linear plot indicates direct inactivation, while a hyperbolic plot suggests a two-step mechanism involving a reversible step before covalent bond formation. The k_{obs} versus [I] plot yields the kinetic constants used to describe the potency of irreversible inhibitors: k_{inact} and K_I (Eq. 3). K_I represents the inhibitor concentration needed for half the maximum possible rate of covalent bond formation [67,68]. It is essential to note that this is constant is distinct from K_i, which is related to the dissociation of the E-I complex and is unaffected by covalent bond formation. The k_{inact} is a first-order rate constant that characterizes the maximum possible rate of covalent bond formation. Thus, the entire rate of covalent bond formation, starting from free unbound enzyme to the enzyme-inhibitor complex, can be

expressed as the ratio of k_{inact}/K_I [67]. High values in the order of $10^5/10^6$ represent high potency compounds.

Equation 3: Relationship between k_{obs}, k_{inact} and K_I

$$k_{obs} = \frac{k_{inact}[I]}{K_I + [I]}$$

1.3.2 Competitive Inhibitors

Reversible inhibitors are classified depending on their inhibitory mechanism. They can have a competitive, noncompetitive, or mixed mechanism of inhibition [65]. In this work, reversible competitive inhibitors were studied; thus, this subsection will be based on this mechanism of inhibition.

Following the scheme below, reversible competitive inhibitors compete with the substrate for the enzyme's active site. Consequently, a competitive inhibitor changes the enzyme's reaction's K_M without altering V_{max}. This change arises because more substrate molecules are required for the enzyme to become saturated.
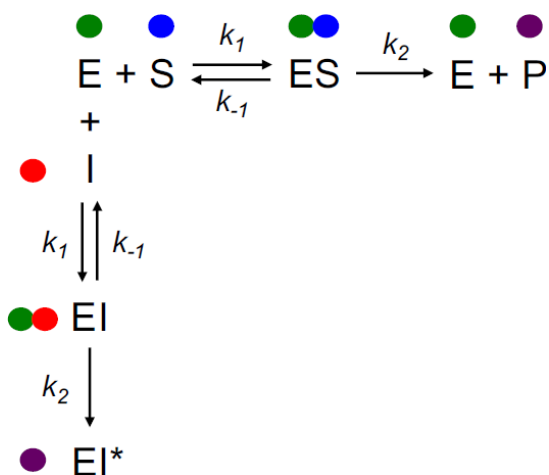
Source: Own author

The kinetic constant k_{-1} is named the inhibition constant (K_i) and is related to the dissociation of the enzyme-inhibitor complex. The K_i indicates the strong enzyme affinity for the inhibitor [69].

1.4 Covalent Reversible Inhibitors

Different classes of covalent inhibitors (CI) can effectively target cysteine proteases. CI are equipped with an electrophilic *warhead* that mediates their affinity for the biological target, taking advantage of the nucleophilic nature of the cysteine residue of CPs [70,71]. Covalent inhibitors offer several advantages over non-covalent ones, such as higher affinity, low dosing, reduced off-target effect, and prolonged residence time. Currently, up to 40 covalent drugs have received FDA approval [72]. Among these, some well-known drugs to treat COVID-19 disease include aspirin, penicillin, omeprazole, and paxlovid [23,70,71].

Different warheads can be used to modulate affinity with the CPs, like vinyl sulfones and epoxides (irreversible inhibitors), and nitriles, aldehydes, and oximes are examples of reversible warheads [25,53,69]. Therefore, in this work, different warheads were used to inhibit different CPs, giving a focus on the nitriles. The inhibitors are all peptide-mimetic molecules with a common scaffold depicted in figure 6AB for most compounds. Some compounds investigated against M^{pro} had a scaffold similar to Nirmatrelvir (Fig. 3). To discover the most suitable functional groups for affinity, variations were made in specific positions of the molecule (Fig. 6A).

Figure 6: (A) Inhibitor scaffold highlighting the P1, P2, and P3 positions. (B) Standard compound (Neq570) employed in the kinetics assays targeting cathepsins, with the interactions between the compound and the CP's subsites highlighted.
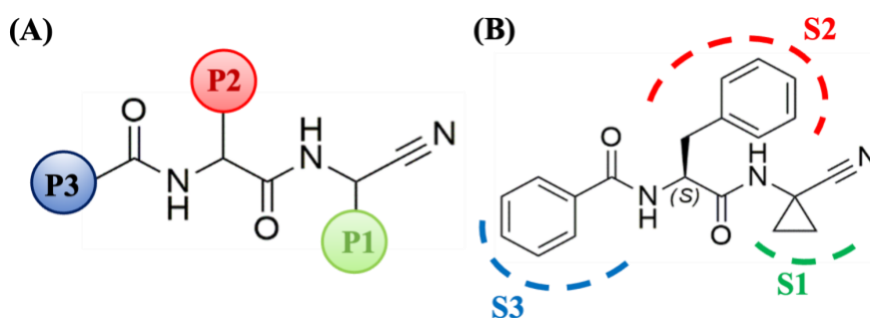
Source: Own author

Covalent inhibitors can bind to the target either reversibly or irreversible, depending on the specific *warhead* that will govern the reaction rate (Fig. 7) [70,73,74]. The binding process unfolds in two steps; first, an initial reversible binding, and then the formation of the covalent bond with the target [6,68].

Figure 7: Generic mechanism played by covalent inhibitors. E is the enzyme, I the covalent inhibitor, E*I is the non-covalent enzyme-inhibitor complex, and E-I is the covalently-bound complex between enzyme and inhibitor.
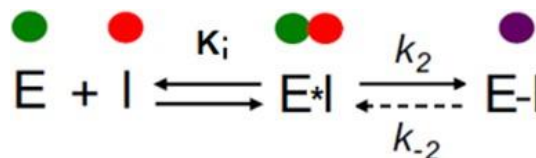
Source: Own author

Reversible covalent inhibitors exhibit finite values for the rate constants k_2 and k_{-2}, while in the case of irreversible inhibitors, k_{-2} approaches nearly zero. From a medicinal chemistry point of view, reversible covalent inhibitors are of greater interest compared to irreversible ones, primarily because of the elevated potential for toxicity associated with irreversible drugs.

1.5 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) is an analytical technique that directly measures the heat of a reaction associated with a binding event between two or more species in a solution[75]. The method allows one to obtain detailed information regarding the thermodynamic profile of a system, providing entropy (ΔS), enthalpy (ΔH), Gibbs free energy (ΔG), and an association constant (K_a), and stoichiometry (n) of the reaction in a single experiment[76,77]. Moreover, unlike non-calorimetric techniques, it is possible to estimate the heat binding capacity (ΔCp) through several experiments at different temperatures[77].

During a usual ITC experiment, a known concentration of the macromolecule is stored in a cell, in which the heat changes are measured in an adiabatic environment. The ligand is stored in a syringe and, throughout the experiment, is gradually titrated in small aliquots into the cell containing the macromolecule (Fig. 8). Every injection results in a heat change as more macromolecules bind to the ligand. Eventually, the ligand will surpass the macromolecule by a molar ratio of two to three-fold. Any alterations in the observed heat will be linked to other components in the solution, such as buffer molecules and solvent [78].

Figure 8: Schematic representation of isothermal titration calorimetry equipment (left) and a usual ITC result isotherm obtained for an exothermic reaction (right).
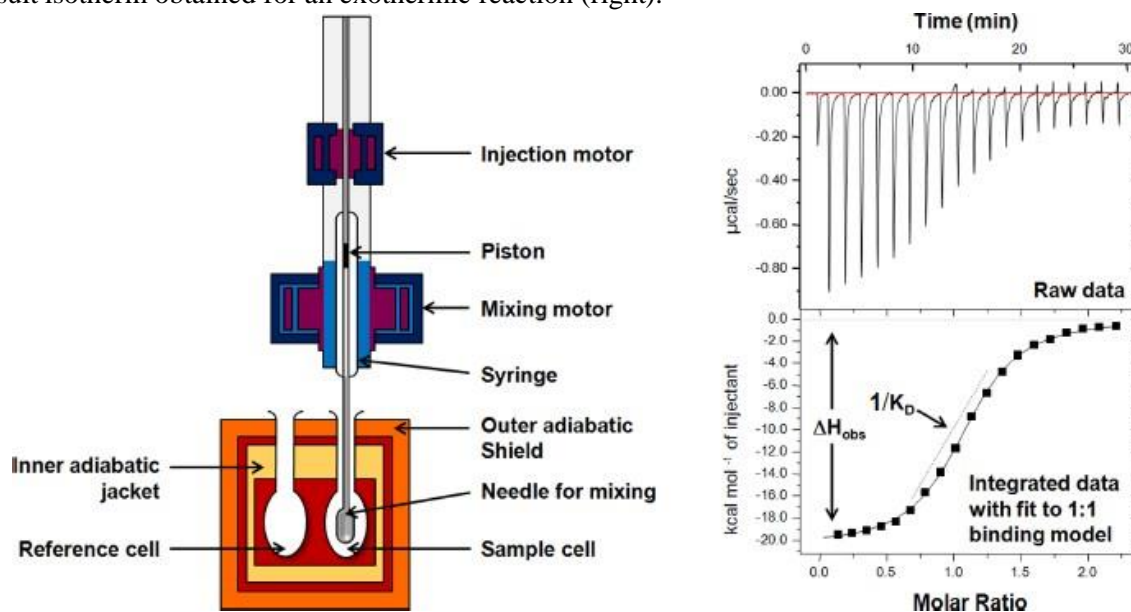
Source: Adapted from Geschwindner, S.; Ulander, J.; Johansson, P., Ligand Binding Thermodynamics in Drug Discovery: Still a Hot Tip?, **J. Med. Chem**., v. 58, p. 6321 – 6335, 2015.

In an ITC experiment, the obtained sigmoid will vary depending on the c value. The c value, called the Wiseman parameter, is the product of the receptor concentration and the binding constant K_a (Eq. 4) [79]. According to the parameter value, it is possible to classify the type of interaction between protein-ligand as tight-binding when $c > 100$, moderate binding ($10 < c < 100$), and low-binding when $c < 10$ [79]. Thus, if $c > 100$, the inclination of the sigmoid is too high, which can cause problems in the K_a determination. If $c < 10$, the curve will have a low inclination, and values of n, K_a, and ΔH can be miscalculated. Therefore, to avoid such problems, Velasquez-Campo *et al*.[80] described an indirect measurement that can be used for tight-binding ligands ($c > 100$). In the indirect method, the protein is incubated with a lower-affinity inhibitor whose thermodynamic parameters are known, and this solution is then titrated with the high-affinity ligand. The data obtained with this method is adjusted using a competition model available in the equipment software, and the thermodynamic parameters are easily obtained.

Equation 4: Determination of the Wiseman "c" parameter.

$$c = nK_a[M]_t$$

Where n is the number of binding sites per macromolecule (M), and $[M]_t$ is the total concentration.

Furthermore, the ITC technique has become a critical approach used to understand the binding process between macromolecule and ligand, giving detailed information about the enthalpy and entropy contributions of the binding process. Thus, it is a powerful technique to aid drug design.

Moreover, the ITC studies will focus on achieving the desirable postulated thermodynamic profile for SARS-CoV-2 M^{pro}-ligand interaction, in which all thermodynamic contributions are favorable (Fig. 9A). In order to achieve the desired profile, structural variations will be made to the inhibitor scaffold shown in Figure 6A.

Figure 9: a) Desirable thermodynamic profile achieved for the HIV-1 protease inhibitor KNI-764. b) The crystallographic structure obtained for HIV-1 protease bound to KIN-764 (PDB code 1MSM).
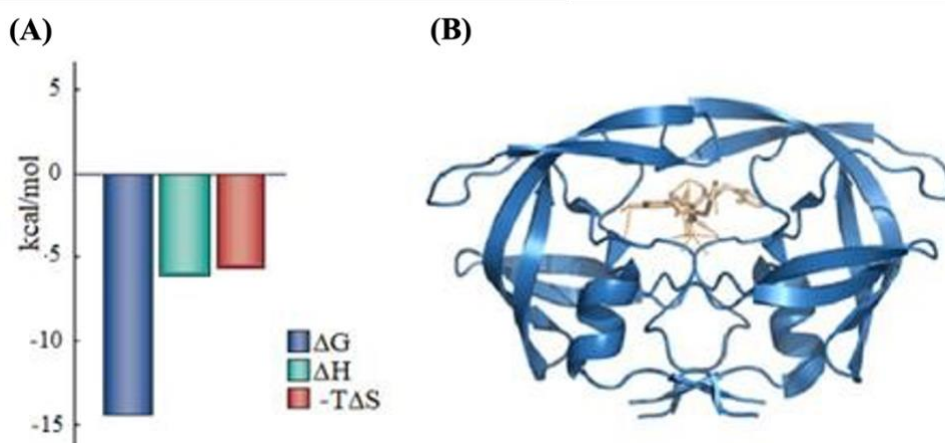
Source: adapted from Claveria-Gimeno, R., Vega, S., Abian, O. & Velazquez-Campoy, A. A look at ligand binding thermodynamics in drug discovery. **Expert Opin. Drug Discov.,** v. 12, p. 363–377, 2017.

1.5.1 *Enthalpy-Entropy Compensation*

Frequently, a structural change in a molecule will not result in significant variation in the ligand's affinity with a macromolecule. This is a known phenomenon called enthalpy-entropy compensation (EEC) [81,82]. The EEC happens when a structural modification in ligand results in a change in the binding enthalpy contribution $\Delta\Delta H \equiv \Delta H_2 - \Delta H_1$ that is partially or entirely displaced by a similar change in the binding entropy component $T\Delta\Delta S \equiv (T\Delta S_2) - (T\Delta S_1)$ [82].

Understanding the enthalpy-entropy phenomenon can be used to identify favorable functional groups aided by matched molecular pair analysis (MMPA), which can help design new drug candidates.

1.6   Structure-activity Relationship Analysis

Drug discovery is a multidisciplinary field that involves the optimization of compounds for various crucial factors, including enhanced affinity, improved bioavailability, and safety profiles, to transform them into promising drug candidates. A common approach for optimizing compounds is the analysis of structure-activity relationships (SAR).

SAR refers to the connection between a molecule's chemical or three-dimensional structure and biological activity. The analysis plays a fundamental role in drug discovery, from initial screening to lead optimization. SAR analysis can be classified into two main categories: those reliant on regression models, such as Quantitative SAR (QSAR), and those that are based on physicochemical approaches [83]. By comprehending SAR for a set of compounds, researchers gain insights into the chemical space of the desired molecule, leveraging this knowledge for further enhancements in physicochemical attributes or activity/selectivity indices [83,84].

1.6.1   Matched Molecular Pairs and Activity Cliffs

Another approach usually used in combination with SAR is the matched molecular pair (MMP) analysis. The MMP evaluates the compound's properties and potency associated with a single-located structural change. If the observable change in the affinity of the pairs is equal or higher than 2.0 log units, the MMPs are called activity cliffs. If no meaningful change is observed, the pairs are called bioisosteres. MMPs can vary in complexity, ranging from straightforward modifications like changing hydrogen for a chlorine atom to more intricate transformations that involve a big part of the molecule [85].

Activity cliffs (ACs) are typically characterized by structurally similar compounds that, despite their similarity, exhibit significant variations in their potency against the same target [86]. ACs are essentially representative of chemical modifications that exerct substantial influence on biological activity. Thus, they hold a special significance in SAR analysis and the process of optimizing compounds [86,87].

ACs can be classified in fingerprint-based and substructure-based. The substructure-based can be further classified in: topology-cliffs (Fig 10), chirality cliffs, R-groups cliffs, and scaffold cliffs [86].

Figure 10: Example of a topology activity cliff where a fluorine atom at the *meta* position (1) was changed to the *para* position of an aromatic ring. The modification led to a change in pK_i of 2.0 log units.
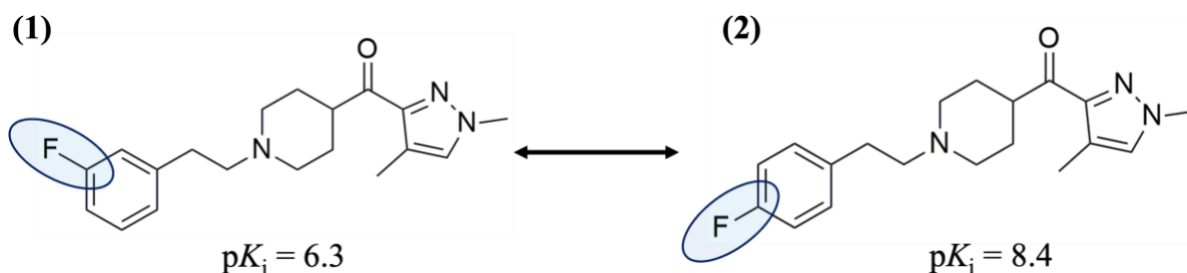
Source: adapted from Stumpfe, D.; Hu, H.; Bajorath, J. Evolving Concept of Activity Cliffs. **ACS Omega**, v. 4, n. 11, p. 14360–14368, 2019

Identifying ACs based on substructures is facilitated through the use of MMPs. Aplication of MMPs to the identification of ACs has led to the introduction of MMP-based ACs, often referred to as "MMP-cliffs". Therefore, the use these analysis were employed throughout this work to analyze a series of covalent inhibitors against different CPs.

2   OBJECTIVES

2.1   General Objectives

Evaluate the effect of mutations in the specificity of SARS-CoV-2 M^{pro} against different substrates and inhibitors. Employ matched-molecular pair and structure-activity relationship analysis with a range of covalent reversible inhibitors to evaluate the influence of different moieties for affinity.

2.2   Specific objectives

- Express and purify the WT M^{pro} and variants;
- Investigate the catalytic activity of the WT M^{pro} in the presence of kosmotropes;
- Explore the effects of the M^{pro} mutations on the catalytic activity and specificity of this protease towards different substrates;
- Study the protease's affinity against a range of peptide mimetic inhibitors;
- Evaluation of MMP/SAR analysis of the inhibitors against the WT M^{pro};
- Investigate the inhibitory capability of a range of inhibitors against the human cathepsins L, S, and B;
- Evaluation of MMP/SAR analysis of the inhibitors against the cathepsins;
- Selectivity analysis between the human cathepsins L, S, and B;
- Express the WT M^{pro} in a mammalian cell line to investigate the inhibitors' lipophilicity using an activity-based probe.

3   MATERIAL AND METHODS

The inhibitors used in this work were synthesized at the NEQUIMED/USP laboratories and at the Patrick G. Johnston Centre for Cancer Research, Queen's University of Belfast at Ph.D. professor Rich Williams laboratory. The substrate Ac-Abu-Tle-Leu-Gln-ACC (QS1) and the activity-based probe BODIPY-PEG(4)-QS5-VS[88] were synthesized in Professor Ph.D. Marcin Drag's laboratory at the Wrocław University of Science and Technology in Poland. The methodologies described in sections 3.1, 3.2, part of 3.6 and 3.8 have been recently published [89].

3.1   Mutagenesis of SARS-CoV-2 M^{pro} Variants

The SARS-CoV-2 M^{pro} plasmid was constructed following the design by Zhang *et al.* [5] and inserted into a pGEX-6p-1 vector for subsequent expression in *E. coli*. Mutations C145A, K90R (Beta1, Beta2), P132H (Omicron), and A193V (Beta2) were constructed using overlap PCR mutagenesis. This process involved two internal, partially complementary, mutagenic primers (forward, primer-1, 3, 5 and 7) and (reverse, primer-2, 4, 6 and 8) (Appendix). The two flanking primers were (forward, primer-9) and (reverse, primer-10) (Appendix). PCR was conducted using a GeneAmp PCR System 2400 (Perkin-Elmer) with Phusion polymerase (Thermo Scientific). The first round of PCR generated two fragments employing primer-9/primers-1,3,5,7 and primer-10/primers-2,4,6,8. The resulting PCR products were purified using agarose gel, and the specific bands were excised and further purified using a QIAquick PCR purification kit (Qiagen). The purified products were then combined and utilized as the template for a second round of PCR with the two flanking primers. The second PCR product underwent purification and was digested with *BamHI/ XhoI*, followed by another purification step, and finally ligated into pGEX-6P-1 restricted with *BamHI/ XhoI*. *E. coli* strain DH5α competent cells were transformed with the ligated vector. The full sequence of the insert was checked by DNA sequencing (Eton BioScience, Washingtons) to verify the success of the mutagenesis and the absence of any other mutations. Finally, the expression plasmid was subcloned into *E. coli* strain BL21(DE3) for further expression of the proteins.

3.2   Expression and Purification of SARS-CoV-2 M^{pro}

The proteins were expressed in *E. coli* BL21 (DE3) competent cells. Cultures were grown in 2x YT media supplemented with 100 μg/mL ampicillin (Sigma-Aldrich). Once the cells reached an OD_{600} of 0.6, protein expression was induced using 0.5 mM of isopropyl-D-thiogalactoside (Sigma-Aldrich). The induction process was carried out for 5 hours at 37ºC and 250 rpm. Subsequently, the cells were harvested by centrifugation at 9954 x g and 4ºC for 15 minutes. The resulting cell pellets were resuspended in lysis buffer A (20 mM Tris, 150 mM NaCl pH 7.8) at a volume of 5 mL per gram of pellet and sonicated for 5 min. The lysate was then centrifuged at 4ºC, 20000 x g for 1 hour, and the resulting supernatant was loaded onto a chromatography column containing Ni^{2+}-chelating Sepharose resin (GE Healthcare Life Science, Washingtons). The column was subsequently washed with 50 mL of buffer A followed by 10 mL of buffer B (20 mM Tris, 500 mM NaCl pH 7.8). The M^{pro}-His was obtained by imidazole elution using different percentages of buffer C (20 mM Tris pH 7.8, 150 mM NaCl, 500 mM Imidazole). Fractions were collected and analyzed by SDS-PAGE. Finallt, the protein samples were preserved at -80ºC for storage.

3.3   Trichloroacetic acid precipitation

Protein precipitation by trichloroacetic acid (TCA) was performed to concentrate protein samples in order to remove contaminants, such as salts (e.g., sodium citrate), so that the protein samples can be analyzed by SDS-PAGE or western blotting. In this procedure, a cold 30% TCA solution was used to process the protein samples. Half the sample volume of TCA solution was added to each sample. These mixtures were then incubated on ice for 30 minutes, followed by centrifugation for 30 minutes. After centrifugation, the supernatant containing the purified proteins was separated, and the pellet was subjected to a wash with 10% cold TCA solution. This washing step was followed by another round of centrifugation for 10 minutes. The final wash was conducted using pure acetone, and the samples were subjected to a final centrifugation for 5 minutes. The resulting protein pellets were dried at 37°C. The final pellet was resuspended in 20 μL of sample buffer for SDS-PAGE. Centrifugations were performed at 4 ºC and 16,000 xg.

3.4   SDS-Page

The SDS-PAGE was used to separate proteins according to their apparent molecular mass and evaluate purity. Samples were mixed with the sample buffer containing 1 mM of DTT

and heated at 95 ºC for 5 minutes to complete the denaturation of the proteins. The samples and the precision plus protein™ unstained standard (Bio-Rad) were added to a Bolt™ 4-12% Bis-Tris Plus gels (Invitrogen) and run at 200 V for 25 minutes using Bolt™ MES-SDS buffer (Invitrogen). The gels were stained with InstantBlue® Coomassie protein stain (Abcam-Sigma-Aldrich) for 30 to 60 min under continuous shaking or used directly for western blotting. The gels were distained with milli-Q water.

3.5   His-tag Removal by PreScission Protease

The removal of the His-tag from M^{pro} was performed using the PreScission protease (Cytiva) to evaluate the effects on its activity. The cleavage of the His-tagged M^{pro} was done using 20 µM of the enzyme and adding different concentrations of prescission (5 units – 0.01 units) into the His-tagged M^{pro} solution. The final solution (M^{pro}-His + prescission) was then incubated for 1 hour at 37ºC. The samples were analyzed through SDS-PAGE and assayed to evaluate the enzyme activity.

3.6   Enzyme Kinetics Assay

The assays were performed in a CLARIOstar microplate reader (BMG Labtech) and a Biotek Synergy HT microplate reader in a 96-well white/black plate (Corning). The assay followed the release of 7-amino-4-carbamoylmethylcoumarin (ACC) fluorophore from the substrate Ac-Abu-Tle-Leu-Gln-ACC (QS1) [88], provided by professor Ph.D. Marcin Drag from the Wrocław University of Science and Technology in Poland. The excitation and emission wavelengths for the substrate are 355 nm and 460 nm, respectively. The substrate hydrolysis rates (RFU/s) were determined and analyzed according to the methods below, which describe equilibrium and rate constants for the inhibition of M^{pro} by the various inhibitors.

3.6.1   Effect of Kosmotropes on the M^{pro} Activity

To assess the activity of the WT M^{pro}, a series of kosmotropic salts were employed: sodium citrate (NaCitrate), ammonium citrate (NH$_4$Citrate), Sodium sulfate (NaSO$_4$), ammonium sulfate ((NH$_4$)$_2$SO$_4$) and sodium chloride (NaCl). The assay was conducted in a buffer containing 20 mM PIPES pH 7.2, 100 mM NaCl, 1 mM EDTA, and 4 mM DTT. Enzyme samples (0.2 µM) were incubated in the buffer alongside varying concentrations of kosmotropes

in a range from 0 to 1.4 M. The final substrate concentration was set at 20 μM. The enzymatic assay was carried at 37 ºC. The fluorescence generated due to the release of the ACC fluorophore was measured over approximately 15 minutes. To determine the rate of substrate hydrolysis, only the linear portion of each progress curve measured in RFU/s, was considered.

3.6.2 Determination of the Michaelis-Menten Constant (K_M) for WT SARS-CoV-2 M^{pro} and Variants

The WT M^{pro} and variants were prepared in the assay buffer (20 mM PIPES pH 7.2, 100 mM NaCl, 1 mM EDTA, 0.7 M sodium citrate, 0.1% Triton X100, and 4 mM DTT). The final concentration of M^{pro} (25 nM) was titrated with serial dilutions of substrate. The rate of hydrolysis was monitored for approximately 15 min at 37ºC. Data was analyzed using GraphPad Prism 8.

3.6.3 Jump Dilution Assay

M^{pro} and inhibitors were incubated for 30 min at 100 nM and 200 nM in assay buffer. Subsequently, a 10-fold dilution was performed, leading to a final concentration of [M^{pro}] = 10 nM and [I] = 20 nM. The reaction started with adding the mixture to the QS1 substrate with a final concentration of 20 μM. The reaction was followed for approximately 30 minutes.

3.6.4 Determination of the Inhibition Constant (K_i) for WT SARS-CoV-2 M^{pro} and Variants

Inhibitors were prepared in assay buffer (20 mM PIPES, 100 mM NaCl, 1 mM EDTA, 0.7 M sodium citrate, 4 mM DTT, pH 7.2), with starting concentrations varying from 10 to 0.1 μM. The inhibitors were preincubated with M^{pro} (25 nM) for 30 min at 37ºC. The reaction was started with the addition of the QS1 substrate with a final concentration of 20 μM. The inhibition contants were determined using the Morrison equation (5) in the GraphPad Prism 8 software.

Equation 5: Morrison Equation. Y is the enzyme activity, X concentration of inhibitor and Et is the total enzyme concentration.
$$Y=Vo*(1-(((((Et+X+(Ki*(1+(S/Km))))-(((Et+X+(Ki*(1+(S/Km))))^2)-4*Et*X)^0.5))/(2*Et)))$$

3.6.5 Determination of k_{inact}/K_I for Time-dependent Inhibitors

SARS-CoV-2 M^{pro}-His (25 nM) were added into wells containing a mixture of different concentrations of irreversible inhibitors, starting at 10 μM, and substrate QS1 (20 μM). The reaction was followed for 30 min. The progress curves were analyzed using equation 6 in the GraphPad Prism 8 software.

Equation 6: Equation used in GraphPad Prism for the determination of k_{obs} for irreversible inhibitors. P is the product formation, v_o is the initial reaction rate, v_o is the steady-state velocity and t is the time

$$[P] = vs.t + (vo - vs)\frac{(1 - e^{-k_{obs}t})}{k_{obs}}$$

3.6.6 Determination of the Kinetic Constants for hCatL, hCatB, and hCatS

Enzyme kinetics assays were conducted at room temperature using Corning® 96-well black flat-bottom microplates, with an excitation wavelength of 360 nm and emission wavelength of 460 nm for the substrate Z-Phe-Arg-AMC. The enzymes were activated in their respective assay buffer (Table 1) supplemented with 5 mM DTT (Sigma-Aldrich) and 0.014% of Triton X-100 (Sigma-Aldrich). Table 1 provides information on enzyme concentrations, activation time, and temperature used in the assays. The rate of substrate hydrolysis was followed for 5 min. Each assay was carried out in triplicates and covered eight different substrate concentrations.

Table 1: Parameters used in the kinetics assays for different cathepsins.

Enzyme	[Enzyme] nM	Substrate	Assay buffer	Activation time	Activation Temperature
*h*CatL	1.9	Z-FR-AMC	100 mM sodium acetate pH 5.5	20 min	0 °C
*h*CatS	2.0	Z-FR-AMC	100 mM sodium citrate pH 6.0	1 hour	0 °C
*h*CatB	1.0	Z-FR-AMC	100 mM sodium phosphate pH 6.0	30 min	0 °C

3.6.7 Determination of the Inhibition Constant (K_i) for hCatL, hCatS, and hCatB

The enzyme activation steps and concentrations are the same as in Table 1. However, substrate concentrations in the assay were fixed and equal to the K_M of each enzyme, so $[S] = K_M$.

Following the activation step, the inhibitors were introduced to the enzyme solution and incubated for 5 minutes. The reaction started with the addition of the substrate solution. Subsequently, the rate of substrate hydrolysis was calculated using Gen5TM Biotek software. The apparent inhibition constant K_i^{app} was calculated through a non-linear data fitting using Origin 2020 software. The equation used for the calculations was $V_s = V_0/(1 + [I]/K_i^{app})$, with V_s representing the steady-state velocity, V_0 the velocity in the absence of inhibitor, and $[I]$ denoting the inhibitor concentration. Each experiment was conducted in triplicate.

The true inhibition constant, K_i, was derived by the correction of K_i^{app} using the Cheng and Prusoff [90] equation: $K_i = K_i^{app}/(1 + [S]/K_M)$. Notably, none of the tested compounds exhibited fluorescent at the wavelengths employed in the assay.

3.7 Differential Scanning Fluorimetry

Differential scanning fluorimetry assays (DSF) were performed in a qPCR system Mx3000P (Agilent). The solution was prepared in 20 mM Tris pH 7.8 buffer using SYPRO Orange protein gel stain (Thermo Fischer) at a final concentration of 5X. M^{pro}'s final concentration was 25 µM. Different percentages of DMSO (Sigma-Aldrich) were evaluated. Denaturation curves were obtained at a ranging temperature of 25 to 75 °C with a rate of 1°C/ cycle; fluorescence was measured at the end of each cycle. The data were analyzed using Boltzmann fitting on Origin 2020 Software.

3.8 Cleavage of Protein-based Substrate

Protein substrate cleavage (catalytic mutant M^{pro}) was assessed by examining the impact of various concentrations of WT M^{pro} and variants through SDS-PAGE. The catalytic mutant M^{pro} was added to protein solutions at a fixed concentration of 2.0 µM. These solutions were then incubated for 1 hour at 37 ºC. Following the incubation, the reaction was stopped by the addition of 30% TCA. The resultant samples were loaded onto an SDS-PAGE gel to separate the reaction products. The catalytic efficiency was determined using equation (7) [91] by determining the M^{pro} concentration at which 50% of substrate cleavage occured (EC_{50}). The gels were scanned at 700 nm using a LI-COR Odyssey CLx.

Equation 7: Determination of k_{cat}/K_M for the cleavage of the protein-based substrate. EC_{50} is the active M^{pro} concentration necessary to cleave half the substrate during the incubation period t.

$$k_{cat}/K_M = \frac{ln2}{(EC_{50})t}$$

3.9 Calorimetry Experiments

The calorimetry experiment was performed using a PEAQ-ITC (Malvern). Before the experiments, M^{pro} was dialyzed in an amicon ultra-10 10-kDa membrane for buffer exchange and concentration. The solution was centrifuged in an Eppendorf 5804R equipment with 4500 rcf and a temperature of 4 °C. Protein concentration was measured in mg mL^{-1} in a DeNovix DS-11+ spectrophotometer using a molar extinction coefficient of 32890 M^{-1} cm^{-1}. The final buffer was 50 mM HEPES pH 7.5 and 150 mM Na$_2$SO$_4$.

3.9.1 Isothermal Titration Calorimetry

The standard parameters used in the direct measurements are depicted in table 2. Parameters are corrected depending on the inhibitor's affinity for the target.

Table 2: Standard PEAQ ITC parameters used in the experiments.

PEAQ ITC parameters	
Temperature	25 °C
Reference Power	5 μcal s^{-1}
Rotation Speed (syringe)	750 rpm
Number of Injections	19
Injection Volume	2 μL
Space Between Injections	120 s
Cell Total Volume	200 μL
Syringe Total Volume	40 μL
Delay	120 s
SARS-CoV-2 M^{pro} concentration	40 μM

The number and volume of injections, as the concentrations, were adapted depending on the inhibitor's strength. The thermodynamic parameters, such as ΔG, ΔH, -TΔS, and K_d, were calculated using Microcal PEAQ-ITC Analysis software.

3.9.2 Displacement Titration Method

A displacement titration must be performed to determine the thermodynamic parameters of high-affinity molecules to reduce accuracy problems in the parameter's determinations. To overcome these problems, the protein is incubated with a weak ligand that will be displaced by a stronger ligand titrated in the cell solution.[80] The thermodynamic parameters of the weak inhibitor must be known so that the competitive analysis method can be used.

The parameters of the experiment are the same as shown in table 2. Data were analyzed using Microcal PEAQ-ITC Analysis software.

3.10 Cell Culture

The human embryonic kidney cells HEK293T were cultured in an Isotemp incubator (Fisher Scientific) at 37ºC and 5% CO_2 in tissue culture-treated T25, T75 flasks and Petri dishes. The media used for the culture of the cells was DMEM (Dulbecco's Modification of Eagle's Medium) with 4.5 g/L glucose and sodium pyruvate without L-glutamine + 10% FBS (Fetal Bovine Serum) + PSG (Penicillin-Streptomycin-Glutamine 100X) (Gibco).

Adherent cells were treated with 5 mL of trypsin/EDTA 0.25% (Gibco) and incubated for 5 min at 37 ºC. Trypsin was inactivated by adding 5 mL of fresh media, and the solution was centrifuged for 6 min at 21 ºC and 130 rcf to obtain the cell pellet. Approximately 9 mL was removed from the centrifuge tube and the pellet was resuspended with 10 mL of fresh media. The cells were counted using a trypan blue solution of 0.4% (Sigma) with a hemacytometer (Fisher Scientific) to obtain the number of cells in cell mL^{-1}. After counting, the cells were added to a culture flask in a total cell concentration, guaranteeing a 50/70% cell confluency for transfection.

3.10.1 SARS-CoV-2 M^{pro} Transfection in HEK 293 T Cells

HEK293T cells were transfected with the vector pCDNA4 that encodes for SARS-CoV-2 M^{pro} (Appendix) to evaluate if mammalian cells can express the target protein. Cells were transfected with the SARS-CoV-2 M^{pro} vector and an empty vector as control, using lipofectamine™ 2000 (Invitrogen). Two solutions were prepared: A) 3 mL Opti-MEM (Gibco) + 120 μL of lipofectamine 2000; B) 3 mL Opti-MEM + vector (total DNA of 24 μg). The solutions A and B were mixed together and incubated for 20 min at room temperature (RT). Meanwhile, the cell media was refreshed with a non-PSG DMEM media. The transfection

mixture is added to the cells drop-like, following an incubation for 24 hours. After 20-24 hours of transfection, the transfection mixture was replaced by fresh media containing PenStrep.

3.10.2 Western Blot

Cells were lysed using RIPA buffer containing a series of protease inhibitors: 20 μM of 3,4-DCI, 10 μM E-64, 10 μM pepstatin A, and 10 μM Iodoacetamide. Lysates were cleared by centrifugation and the resulting supernatants were quantified using the BCA protein assay (Pierce ™) before performing a Western blot (WB). First, an SDS-PAGE gel with an Odyssey® molecular weight marker (LI-COR BioScience, Washingtons) was obtained and transferred to a nitrocellulose membrane using Bolt™ Transfer Buffer (Invitrogen) for 60 min at room temperature and 10 V. After the transfer, the membrane was stained with a 0.1 % Ponceau S solution (Fluka BioChemika) for 1 minute to evaluate if the transfer occurred successfully. The Ponceau S was entirely removed by water and Tris-buffered saline (TBS-T) pH 7.4 (Thermo Fisher) for 5 minutes. Afterward, the membrane was incubated in 5 % bovine serum albumin (BSA) + TBS-T buffer for 1 hour with continuous shaking at room temperature to block unspecific proteins. After blocking, the membrane was incubated with the primary antibody DYKDDDDK-Tag Rabbit ab, which recognizes a FLAG-tag (LI-COR BioScience, Washingtons). The incubation was carried out in 5% BSA + TBS-T buffer overnight at 4 ºC with continuous shaking. The membrane was washed for 3x5 minutes with TBS-T buffer at room temperature to remove the unbound primary antibody. The secondary antibody IRDye 800CW Donkey anti-Rabbit (LI-COR BioScience, Washingtons) was added to the membrane in TBS-T buffer, incubated for 1 hour, and then washed with TBS-T buffer to remove unbound secondary. Also, to evaluate the protein's activity, the activity-based probe (ABP) Bodipy-PEG (4)-Abu-D Tyr-Leu-Gln-VS was used. The ABP was incubated with the lysate for 30 min before SDS-PAGE and western-blot. The target protein was detected using an ODYSSEY CLx imaging equipment (LI-COR).

4   RESULTS AND DISCUSSION

4.1   WT SARS-CoV-2 M^{pro} and VOCs Expression and Purification

The M^{pro} mutations identified in the variants were performed by site-directed mutagenesis of the wild-type (WT) M^{pro} construct (Fig. 11A). The beta variant showcases the K90R and A193V mutations, situated at the interface of domain I, while the A193V mutation is positioned within the loop connecting domains II and III. The omicron variant mutation P132H is present within domain II of the protein, as illustrated in Figure 11B.

Figure 11: (A) SARS-CoV-2 M^{pro} construct, featuring the auto-cleavage site during expression (red arrow). The table provides details on the M^{pro} variants, including their respective mutations within the protein sequence. (B) Dimeric structure of M^{pro}, with the catalytic dyad C145 and H41 highlighted in protomer A (red), and the protein's mutations highlighted in protomer B (green). PDB: 6WTM
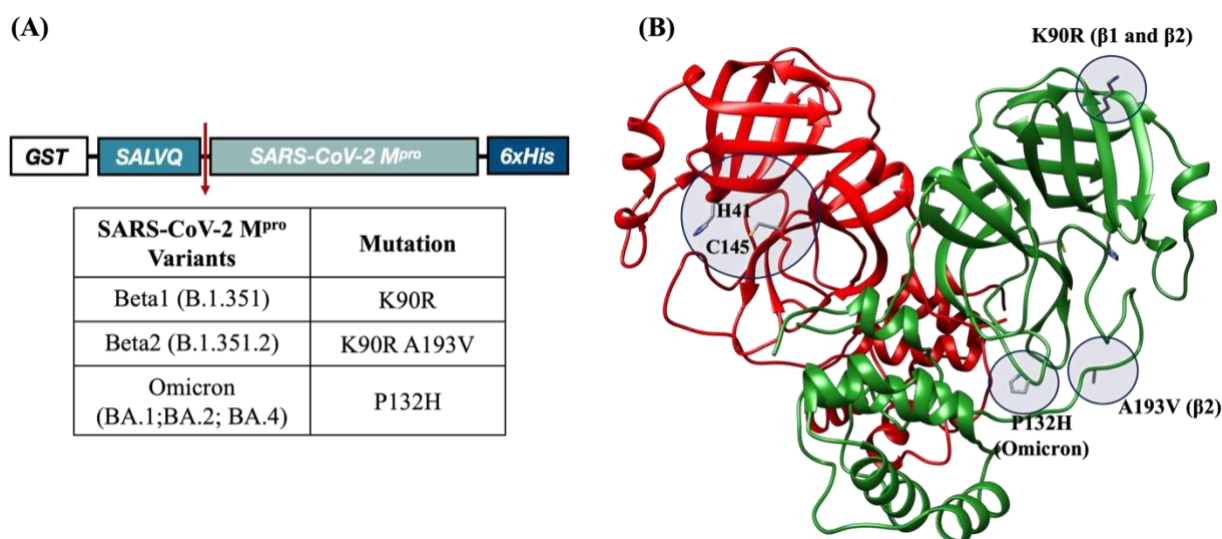
SARS-CoV-2 M^{pro} Variants	Mutation
Beta1 (B.1.351)	K90R
Beta2 (B.1.351.2)	K90R A193V
Omicron (BA.1;BA.2; BA.4)	P132H

Source: Accepted Article Rocho, F. R., *et al*., Differential specificity of SARS-CoV-2 main protease variants on peptide versus protein-based substrates, **The FEBS Journal**, 2023

The purified proteins were obtained with the histidine tag (6x His) in a yield ranging from 8.0 to 10 mg of protein for a 2 L expression volume. The SDS-PAGE electrophoresis was used to verify the protein's purity and to validate its correct structures by molecular sizes; all proteins shared the same molecular mass of 34.8 kDa. Figure 12 shows the SDS-PAGE gel for the WT M^{pro} (See appendix for SDS-PAGE of the M^{pro} variants).

Figure 12: (A) Protein gel of WT M^pro-His purification step using Ni column. A gradient of imidazole was used starting at 7.8 mM (-) until 500 mM (+) (2-fold dilution). (B) Gel (a) with the fractions of 250 and 500 mM of imidazole diluted 5-fold.
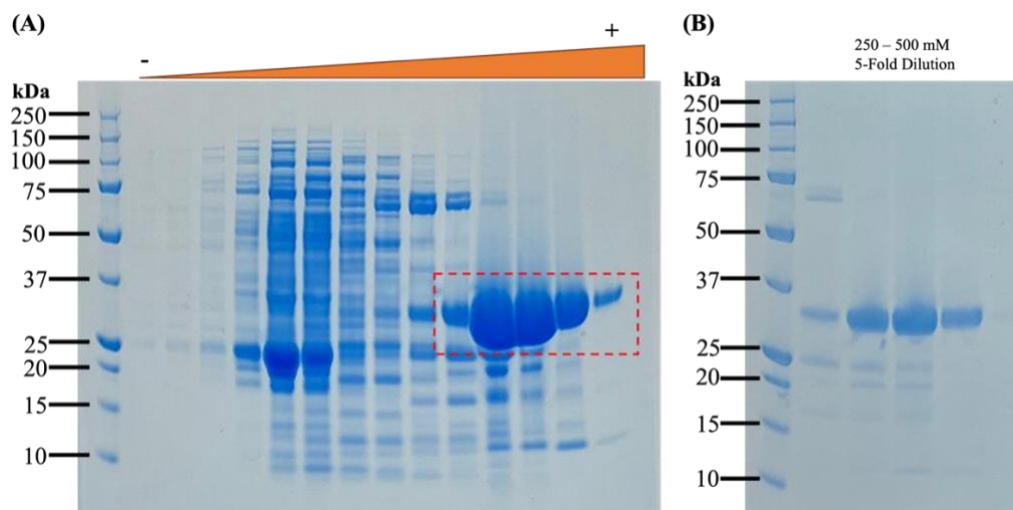
The M^pro catalytic activity relies on dimerization, so a study was conducted to assess whether the presence of the His tag at the protein's C-terminus might negatively affect the dimerization/activity of the protease. For that, an assay was performed titrating the WT M^pro in the presence of different concentrations of PreScission protease. With this kind of assay, it was possible to identify the optimum concentration of PreScission needed to cut WT M^pro effectively. The His tag removal was accompanied by SDS-PAGE (Fig. 13A). The solution containing the authentic WT M^pro was tested for activity with the QS1 substrate in which the obtained activity was compared with the WT M^pro-His (Fig. 13B). As the data shows, the removal of the His tag does not lead to a significant increase in the enzyme activity. Thus, all the subsequent assays were performed with His tagged M^pros.

Figure 13: (A) SDS-PAGE gel of WT M^{pro} titration with PreScission after 60 min incubation at 37 ºC. WT M^{pro} – 33.8 kDa; WT M^{pro}-His – 34.8 kDa; PreScission – 46 kDa. (B) Activity assay of WT M^{pro} against the substrate QS1. The protease was tested with and without the His Tag in assay buffer 20 mM PIPES pH 7.2, 100 mM NaCl, 1 mM EDTA, 4 mM DTT and 0.1% Triton X-100.
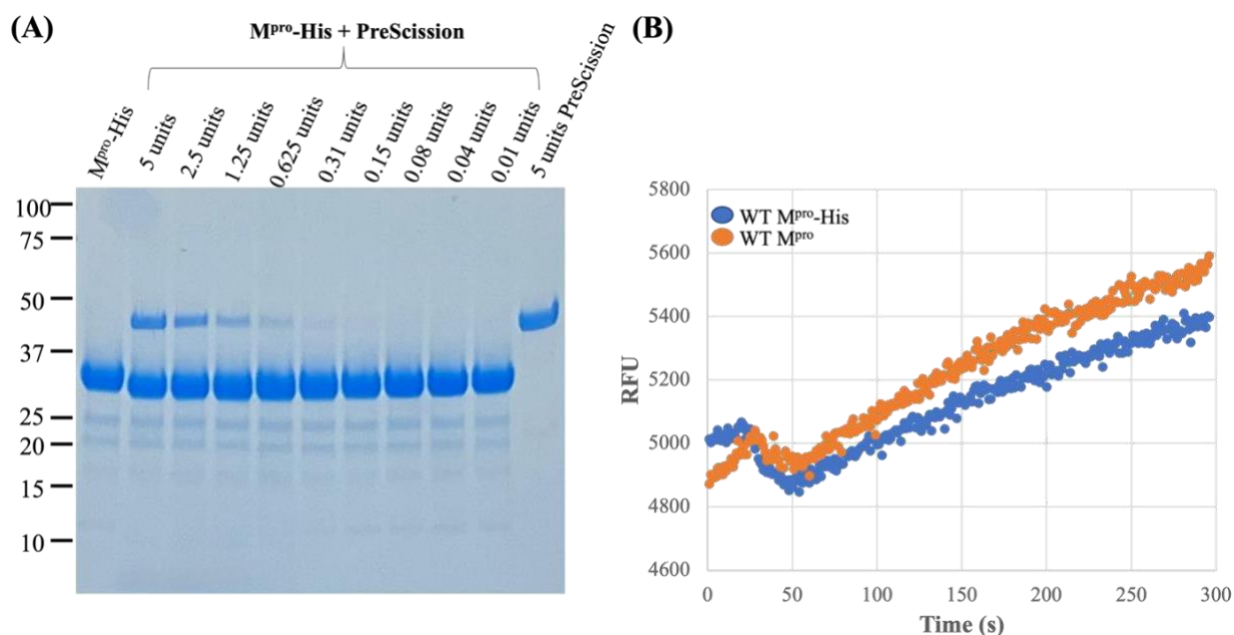
4.1.1 Effect of Kosmotropes in the WT M^{pro} Catalytic Activity

To achieve 100% dimerization and satisfactory catalytic activity, high concentrations of the enzyme are required, as M^{pro} is an obligate dimer. Therefore, with the aim of enhancing the protease's catalytic activity, an assay involving the use of different kosmotropes was employed. This approach allowed for the improvement in catalytic efficiency of M^{pro} without the need of high concentrations of the protease. Kosmotropic salts are ions that stabilize proteins, leading to the stabilization of dimeric conformations [92]. These salts have commonly been utilized as a standard approach to explore the *in vitro* dimeric state of enzymes, including other viral proteases and caspases [93,94].

The catalytic activity of M^{pro} is significantly increased in the presence of high concentrations of some kosmotropes (Fig. 14A), such as sodium citrate and sodium sulfate. The most effective kosmotropic salt was sodium citrate, a known surrogate dimerization reagent [91,95]. The highest activity increase was observed when using 1.4 M of sodium citrate, which led to an increase of 36.5-fold for M^{pro} (Fig. 14B) compared with the activity obtained in its usual Tris buffer pH 7.8. Although the proteases will be tested against a range of inhibitors, the optimum concentration of 1.4 M of sodium citrate was not used in the assays. This is based on the fact that a concentration of 1.4 M of sodium citrate decreased the solubility of the inhibitors.

Thus, using solubility tests with the inhibitors, a concentration of 0.7 M of sodium citrate was chosen for the subsequent assays. Therefore, the assay buffer was established as 20 mM PIPES pH 7.2, 100 mM NaCl, 1 mM EDTA, 0.1% Triton-X100, 4 mM of DTT supplemented with 0.7 M sodium citrate.

Figure 14: (A) Rate of QS1 cleavage at varying concentrations of kosmotropes for the WT M^{pro}. (B) Fold increase in activity of WT M^{pro} at a salt concentration of 1.4 M. The fold increase was determined by comparing the protease's activity in its usual buffer 20 mM Tris pH 7.8, 150 mM NaCl, 1 mM EDTA and 4 mM DTT with the buffer 20 mM PIPES pH 7.2, 100 mM of NaCl, 1 mM EDTA and 4 mM DTT supplemented with different concentrations of the kosmotropes.
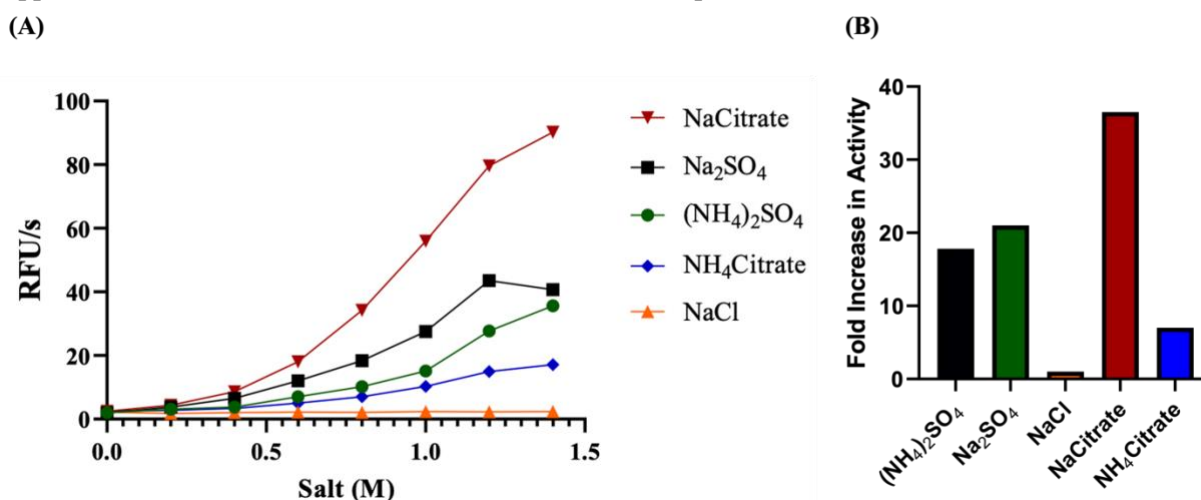
Source: Accepted Article Rocho, F. R., *et al.*, Differential specificity of SARS-CoV-2 main protease variants on peptide versus protein-based substrates, **The FEBS Journal**, 2023

Thus, having established the assay buffer, an assay was conducted aimed at determining both the active enzyme concentration and the impact of sodium citrate on the catalytic activity of WT M^{pro} using the QS1 substrate.

4.1.2 *Determining the Active Concentration of the Protease*

Typically, enzyme concentrations are determined by measuring the sample's absorbance at 280 nm (A280), which considers the overall protein content rather than the amount of active protein. However, this approach can lead to inaccuracies due to autoproteolysis when working with proteases. Thus, to determine the active enzyme concentration of the samples, the M^{pro}s were titrated against a tight binding inhibitor for better accuracy.

As M^{pro} and inhibitors bind in a 1:1 reaction, one can determine the active enzyme concentration in a sample by titrating the enzyme with different inhibitor concentrations. The inhibitor concentration ([I]) must be close to or equal to the enzyme concentration in order to

obtain a quasi-linear relationship when plotting fractional velocity (RFU/s) against [I]. Under these conditions, the fractional velocity will approach zero and remain constant at higher inhibitor concentrations. In this case, the data can be fit to a straight line extended to the *x*-axis, yielding the active enzyme concentration in the assay (Fig. 15). Thus, the tight binding inhibitor Nirmatrelvir was used to titrate the different enzymes. Nirmatrelvir was synthesized at the NEQUIMED/USP laboratories receiving the code Neq1183.

Figure 15: WT M^{pro} titration against Neq1183. (A) fraction enzyme activity against different concentrations of Neq1183. The dashed line represents the linear portion of the graph used to calculate the active enzyme concentration. (B) The linear portion of A) giving an active concentration of the enzyme in the assay equal to 14 nM. Assay conditions were Neq1183 = 0.1 µM, QS1 = 20 µM, and M^{pro} = 25 nM.
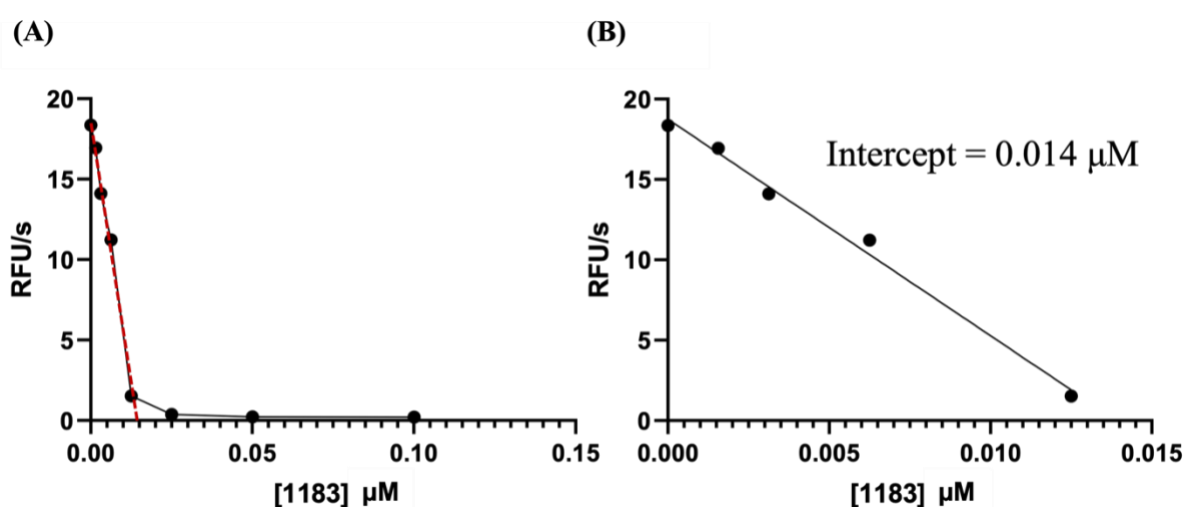
4.1.3 Determining the Kinetic Constants for WT Mpro and VOCs on a peptide Substrate

Seeking to comprehend how sodium citrate influences the catalytic activity of WT M^{pro}, the protease was assayed against the QS1 substrate with two different concentrations of sodium citrate, 1.0 and 0.7 M. The kinetic constants were compared with those obtained using the conventional 20 mM Tris buffer pH 7.8. As expected, the use of sodium citrate revealed changes in both K_M and k_{cat}. The K_M values decreased as sodium citrate concentration increased, suggesting an increase in the protein's affinity for the substrate. Similarly, there was a substantial 45-fold improvement in k_{cat} with increased sodium citrate. Hence, the kosmotrope improved the protease's catalytic activity (k_{cat}/K_M) by 160-fold. These findings consistently support the idea that sodium citrate contributes to heightened binding and catalysis, implying that its presence facilitates a more effective stabilization of the protein's active conformation.

Table 3: Evaluation of sodium citrate effects on the catalytic activity of WT M^{pro}. Kinetic constants for WT M^{pro} were determined using assay buffer containing different sodium citrate concentrations and assayed with the QS1 substrate. Data shown in the table represent the mean and standard deviation from three technical replicates.

[Enzyme] (nM)	Sodium Citrate (M)	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
200	0	182 ± 13	0.140 ± 0.003	769
25	0.7	34 ± 4.9	1.90 ± 0.11	55882
10	1.0	51 ± 2.3	6.30 ± 0.04	123529

The M^{pro} variants were also assayed using the QS1 substrate to assess the impact of mutations on the protease's catalytic efficiency (Table 4). The kinetic constants obtained for the variants revealed that mutations in the M^{pro} sequence resulted in a minimally improved catalytic efficiency for the QS1 substrate. Notably, the Beta2 variant displayed at most a 1.5-fold increase in catalysis when compared to the WT M^{pro}.

Table 4: Kinetic constants for WT M^{pro} and variants against QS1 substrate. Kinetic constants were obtained in assay buffer. Values show the mean and standard deviation.

Enzyme	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
WT	34 ± 4.9	1.90 ± 0.11	55882
Beta1	55 ± 2.1	3.60 ± 0.02	65454
Beta2	46 ± 2.5	3.13 ± 0.02	68043
Omicron	44 ± 3.4	2.54 ± 0.08	57727

Despite not being a high increase in k_{cat}/K_M, the data shows that the virus was able to improve M^{pro} catalytic efficiency with the double mutation observed in the Beta2 variant. Moreover, the mutations found in the M^{pro} sequence shows to have minimal differences in catalysis on the peptide substrate. Implying that the active site environment of the protein did not change enough to affect the enzymes' function and its ability to recognize small molecules, such as the QS1 substrate.

4.1.4 *Determining the Catalytic Efficiencies for WT M^{pro} and VOCs on a Protein-based Substrate*

As the M^{pro} variants seems not to have an impact on the recognition of small molecules, an assay using a protein-based substrate was designed to evaluate the impact of the mutations on a substrate that mimics the natural polyprotein environment. Thus, an assay using a catalytic mutant M^{pro} that contains the catalytic C145 substituted by an alanine residue, was used as a substrate. The basis for the assay is that the active M^{pro} will hydrolyze the catalytic mutant M^{pro}

at its recognition sequence SALVQ↓S. Thus, as the catalytic mutant has a construct GST-SALVQ↓SM^{pro}, with de S residue being the N-terminal of the protein, the cleavage by an active M^{pro} will release the GST domain that can be visualized using SDS-PAGE (Fig. 16A). The SDS-PAGE gels were scanned at 700 nm, and a densitometric analysis of the bands was conducted. The relative band intensities were plotted against the active protease concentration to determine the amount of enzyme needed for 50% substrate cleavage during the incubation period (EC$_{50}$) (Fig. 16B).

Figure 16: (A) Densitometric analysis of the extent of cleavage of 2 μM protein-based substrate by using different concentrations of the omicron variant in assay buffer. Gel analysis was performed using Image Studio software. (B) The correlation between relative band intensities against omicron M^{pro} concentration gives the EC$_{50}$. Catalytic efficiency was determined using equation 7.
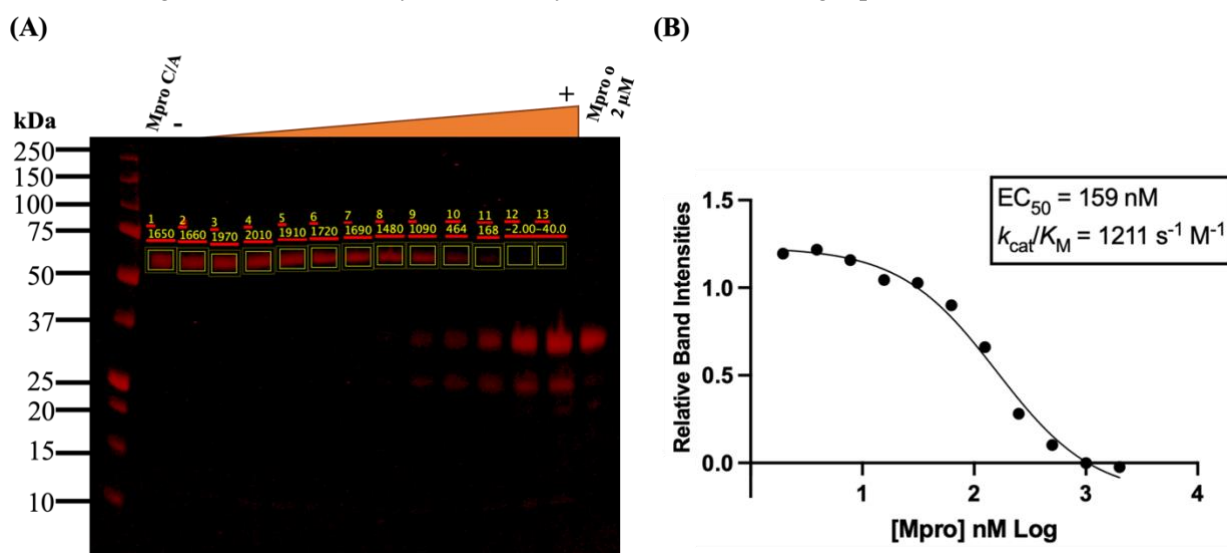
Source: Accepted Article Rocho, F. R., *et al*., Differential specificity of SARS-CoV-2 main protease variants on peptide versus protein-based substrates, **The FEBS Journal**, 2023

The catalytic efficiencies using a protein-based substrate differed considerably from those obtained with the peptide substrate. The Beta2 variant exhibited superior catalytic efficiency compared to the other variants, following the same pattern observed for the peptide substrate (Fig. 17). Notably, it displayed a significant 7-fold increase in cleavage rate when compared to the WT M^{pro}, particularly against the protein-based substrate. These findings imply that the mutations may exert a more pronounced influence on the protease's overall structure, particularly in the context of protein-protein interactions with its native substrates, the nsps. Thus, since the activity of the variants on the peptide substrate was very similar, one can hypothesize that inhibitors initially designed for the WT M^{pro} will be effective in all the variants. The drug Neq1183 (Nirmatrelvir) was tested against the M^{pro}s to test this hypothesis.

Figure 17: Comparison of the cleavage rates of the protein-based substrate by WT M^{pro} and variants. Cleavage was achieved through a dilution series of the M^{pro}'s and assessed via SDS-PAGE. The displayed k$_{cat}$/K_M values represent the mean and range based on two to three independent measurements and were determined using equation 7.
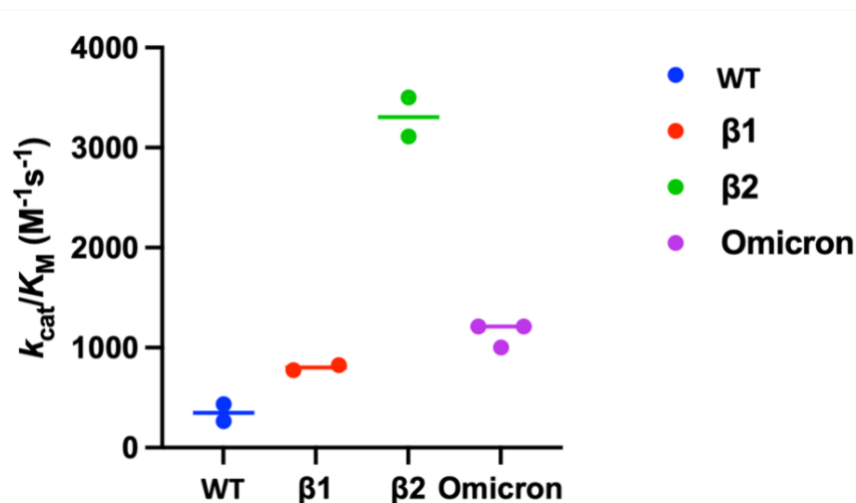
Source: Accepted Article Rocho, F. R., *et al.*, Differential specificity of SARS-CoV-2 main protease variants on peptide versus protein-based substrates, **The FEBS Journal**, 2023

4.1.5 Inhibition Constants Determination for WT M^{pro} and VOCs

Determining ligands' inhibition constants (K_i) is crucial in planning and discovering drug candidates. Through K_i data, it is possible to analyze the ligand structure and assess the importance of the functional groups in the interaction with the biomolecular target (structure-activity relationships) and to evaluate aspects such as selectivity between enzymes. Thereby, K_i determination aids medicinal chemists in achieving high-potent and selective inhibitors by guiding compound optimization.

The K_i is an equilibrium constant given by:

$$K_i = \frac{[E]\,[I]}{[EI]} = \frac{k_{-1}}{k_1}$$

Herein, K_i was determined through an indirect assay, in which the inhibitor and substrate are present in the enzyme solution. First, the inhibitor is incubated with the enzyme for 30 minutes at 37 °C, followed by adding the substrate solution to start the reaction. The fluorescence signals observed by-product formation will have a relative fluorescence unit (RFU) per time. Thus, the RFU corresponds to the rate of hydrolysis of the substrate molecule per unit of time. The obtained RFU/s data is plotted against inhibitor concentration, in which, by using different equations, the K_i can be calculated (Fig. 18).

Figure 18: Inhibition plot for K_i determination. Velocities (RFU/s) are plotted against [I]. Each portion of the graph will give different information, with the K_i being determined at the curvature of the graph.
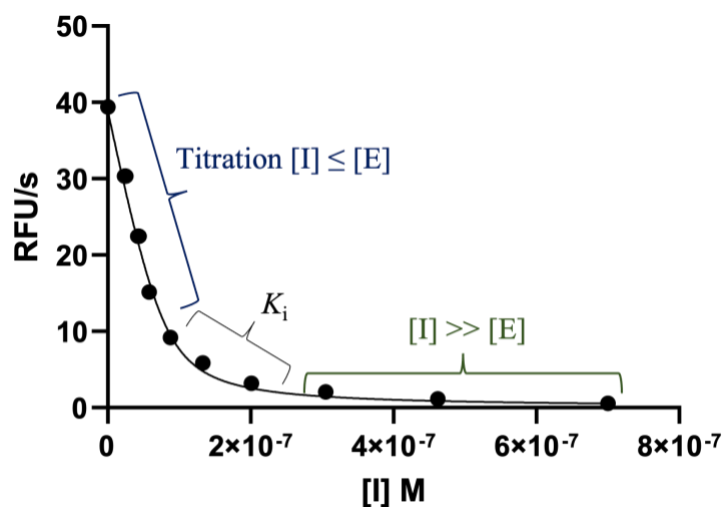
It is extremely important to have enough points at the plot's curvature, as shown in the figure above as it is this portion of the graph that will be used to calculate K_i. In this work, the K_i values were converted to pK_i (-log K_i) to better analyze the results. The higher the pK_i, the greater the compound's affinity. Therefore, the inhibitor Neq1183 and the inhibitors series provided by Prof. Ph.D. Rich Williams were tested against the WT M^{pro} and VOCs.

4.1.5.1 Inhibition Analysis of Neq1183 Against WT M^{pro} and VOCs

The K_i assay was performed for Neq1183 against the M^{pro}s (Fig. 19). As Neq1183 is a tight binding inhibitor, with a ratio of $[E]_0/K_i > 10$, the K_i cannot be calculated by conventional derivations of the Michaelis-Menten relationship. Under these conditions, the inhibitor will titrate the enzyme's active site, and the K_i will not be calculated accurately as there are not enough points at the curvature of the graph. In this case, one can approximate the K_i value of the inhibitor to be lower than the enzyme concentration in the assay.

Figure 19: K_i determination for Neq1183 against WT M^{pro}. The data was fitted using the Morrison equation. K_M was set as 20 μM and enzyme concentration in the assay was 25 nM.
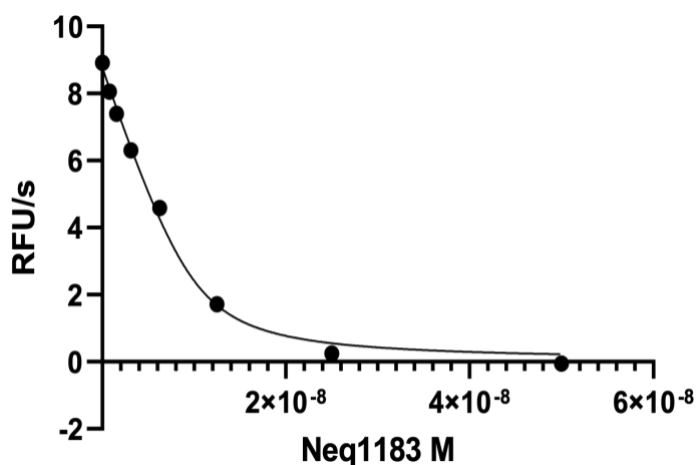
Despite the data obtained for Neq1183 inhibition against all M^{pro}s could be fitted into the Morrison equation, given a K_i of approximately 1 nM (pK_i = 9.0) for all the variants, the K_i obtained was not considered accurate due to the assay conditions that do not follow a steady-state kinetic. In this case, as the inhibitor had a similar behavior for all the M^{pro} variants and the enzyme concentrations used in the assay were all equal to 25 nM, it was considered that the K_i for Neq1183 against the M^{pro}s is below 25 nM, yielding a pK_i higher than 7.5, which is consistent with a previous publication [96]. The results support the hypothesis that the variants cannot distinguish between small molecules, as the mutations did not affect the protease's active site enough to build resistance to inhibition. Neq1183 was further investigated using isothermal titration calorimetry (ITC), but first, M^{pro} thermal stability in the presence of DMSO was examined.

4.1.6   Thermal Stability of SARS-CoV-2 M^{pro}

The thermal stability of SARS-CoV-2 M^{pro} was studied by differential scanning fluorimetry (DSF). The study evaluates M^{pro} thermal stability without DMSO and its stability with 5% and 10% of DMSO (Fig. 20).

The M^{pro} thermal stability in the absence of DMSO was shown to be T_m = 55.2 ° C; the result corroborates with others in the literature [97,98]. As for the T_m found in the presence of DMSO, only at a percentage of 10% a significant variation in the T_m was observed (ΔT_m = 1.5 °C). These results make it possible to understand the limitations of the protein solution regarding the use of DMSO. A certain percentage of DMSO has to be used to study the affinity

of M^{pro} with a range of compounds, as most of the compounds are not soluble in the buffer solution. Usually, 4 to 5 % of DMSO is used in the assays, resulting in a slight variation of the Tm ($\Delta Tm = 0.6$ °C) compared with the 10% of DMSO.

Figure 20: DSF experiment for M^{pro} with 0% (black), 5% (red), and 10% (blue) of DMSO.
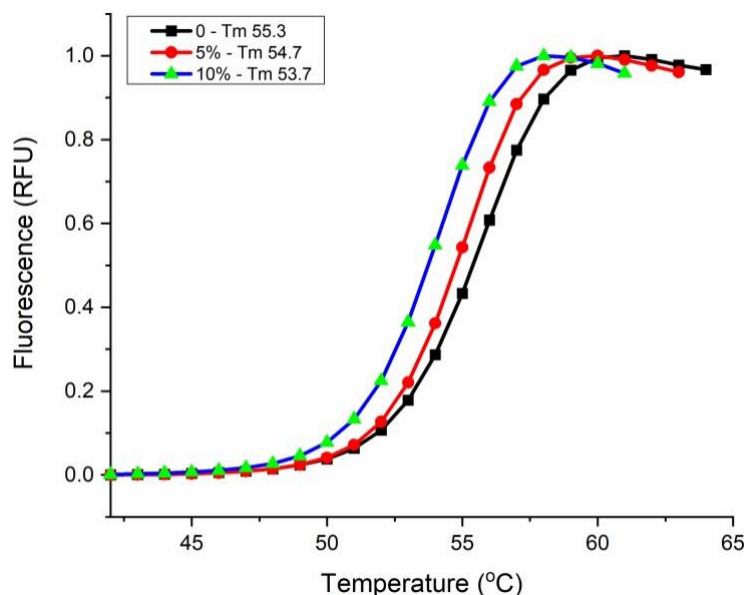
Source: own author

4.1.7 Determination of the Thermodynamic Signature of Neq1183 Against WT M^{pro}

The isothermal titration calorimetry (ITC) technique was used to understand better the bimolecular interaction between the inhibitor and SARS-CoV-2 M^{pro}. The information obtained in an ITC experiment can guide molecular design, as it gives crucial knowledge about the structure-activity relationship of the compounds. Since the free energy (ΔG) is obtained by the contributions of enthalpy (ΔH) and entropy ($-T\Delta S$), it is possible to rationalize the most significant interactions between the inhibitor and the enzyme. For instance, some non-covalent interactions, such as hydrogen bonds, will increase the enthalpy value, while the expulsion of water molecules within the active site or hydrophobic interactions will result in more favorable entropy [99]. Hence, with the values obtained through an ITC experiment, it is possible to modulate the enthalpic and entropic contributions to achieve more ligand stability in the active site (more negative ΔG).

Therefore, Neq1183 was tested using two different methods due to its high affinity for the target ($pK_i = 8.5$). The first was a direct measurement using an increased number of injections (78) of low volume (0.5 µL), and the second was using an indirect method, in which

the protein solution is incubated with a compound of lower affinity for the target (Neq1184, pK_i = 6.5). For the indirect method to be performed, the K_D of the low-affinity inhibitor must be known. Thus, an ITC experiment was done for Neq1184 against SARS-CoV-2 M^{pro} (Appendix). Figure 21 shows the isotherms obtained for the direct and indirect measurement of Neq1183 against SARS-CoV-2 M^{pro}. The thermodynamic parameters for all the inhibitors can be seen in Table 5.

Figure 21: Isotherm obtained with PEAQ-ITC equipment and the sigmoid obtained through the integration of the isotherms peaks for Neq1183 (A) and Neq1183 using an indirect assay (B) against SARS-CoV-2 M^{pro}.
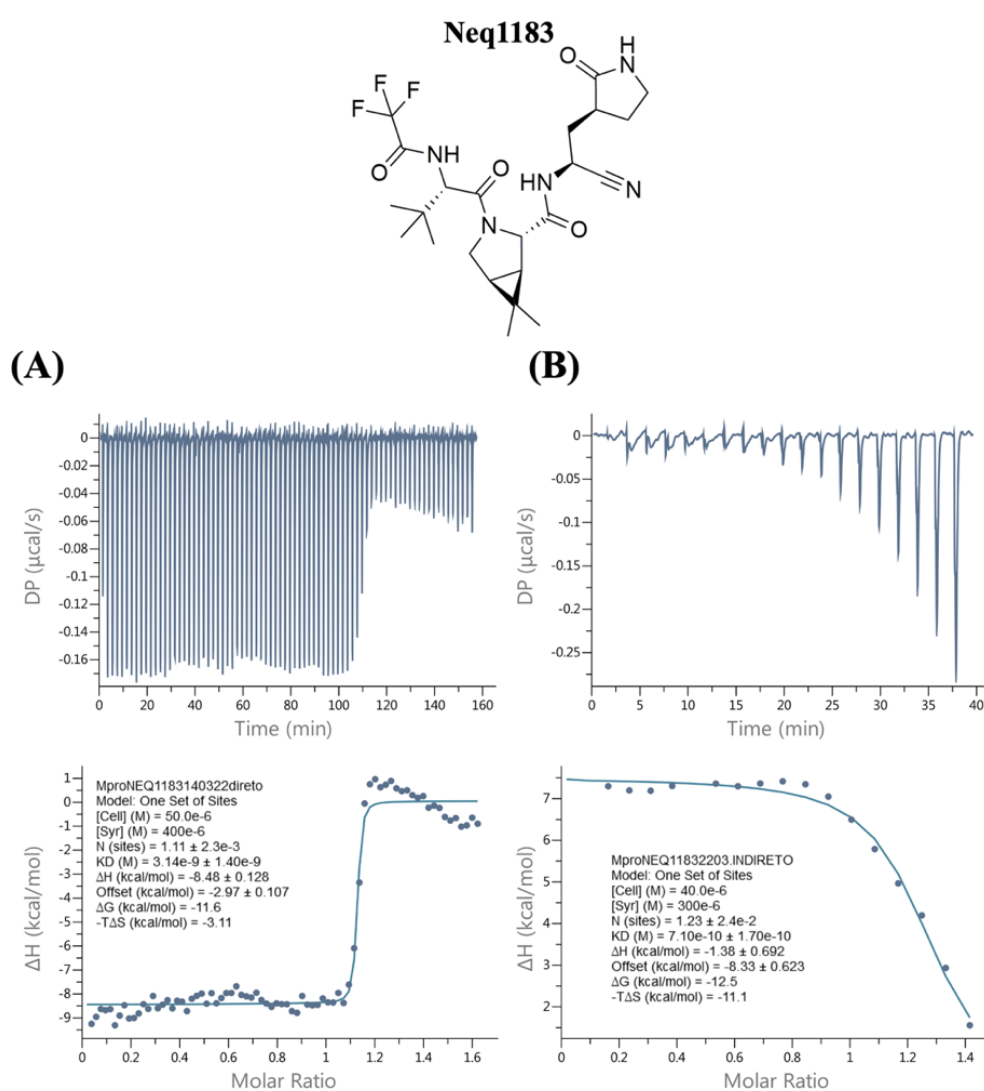
The obtained sigmoid for Neq1183 varied between the two different methods. For the direct measure, the sigmoid profile is as expected. In the beginning, the enzyme is free to interact with the inhibitor, and the assay ends with the saturation of the enzyme. This cannot be seen in the indirect measure; the protein is incubated with Neq1184, so the inhibitor already

occupies its active site. Neq1183 must move Neq1184 out of the enzyme's active site in the reaction. This makes the shape of the sigmoid change, first "saturated" with Neq1184, and then, due to Neq1183's strongest affinity for the enzyme, Neq1184 is moved away from the active site. Both methods had very close pK_D values but had different thermodynamic signatures (Table 5). The ΔG values for both methods are close, with a 1.1 kcal/mol difference. As for ΔH and $-T\Delta S$ values, a substantial difference was observed. The reaction is enthalpy-driven for the direct measurement, and for the indirect measurement, it is entropy-driven. One hypothesis for this discrepancy is that in the indirect measure, the presence of Neq1184 disturbs the environment. Thus, an increase in the system's entropy is observed, decreasing the enthalpy due to the enthalpy-entropy compensation phenomenon.

Table 5: Thermodynamic parameters obtained in a PEAQ-ITC experiment for the direct and indirect assay of Neq1183 and Neq1184 against SARS-CoV-2 M^{pro}.

Compound	K_D (nM)	K_i (nM)	ΔG (kcal/mol)	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	n	χ^2
Neq1183 (direct measurement)	3.42 ± 1.4	1.9 ± 0.001	-11.6	-8.48 ± 0.13	-3.11	1.11 ± 0.002	3.20×10^{-1}
Neq1183 (indirect measurement)	0.71 ± 0.17	1.9 ± 0.001	-12.5	-1.38 ± 0.69	-11.1	1.23 ± 0.002	4.70×10^{-2}
Neq1184	318 ± 34.2	360 ± 0.02	-8.86	-9.64 ± 0.11	0.78	1.04 ± 0.005	2.10×10^{-2}

The data depicted in Table 5 are in agreement with the pK_i values obtained since for Neq1183, a pK_i of 8.5 results in a ΔG of -11.64 kcal/mol, and for Neq1184, the pK_i of 6.5 will result in a ΔG of -8.87 kcal/mol, using the equation $\Delta G = -RT\ln K_i$. Still, the n values are around 1, indicating that the interaction between M^{pro} and the inhibitors is an equimolar reaction.

Comparing the thermodynamic signatures of the compounds, it is observed that Neq1183 has a positive contribution of $-T\Delta S$, whereas Neq1184 has a slightly negative contribution. Both inhibitors' interactions with the target are enthalpically driven and possess an exergonic reaction with negative values of ΔG, with Neq1183 forming a more stable complex with the enzyme. Probably, the reason for the high enthalpic contribution is due to the numerous hydrogen bonds that both compounds make with the M^{pro}, as it is possible to observe in the crystallized structure of the Nirmatrelvir (Neq1183) with the enzyme (Fig. 22). Since Neq1184 is a molecular pair of Neq1183, a similar result should be expected, but the difference in their structures results in a negative contribution of the entropic term ($-T\Delta S$) for Neq1184.

Figure 22: Cocrystalized structure of Nirmatrelvir (Neq1183) with M^{pro} in the covalent form, PDB: 7RFW. Hydrogen bonds are depicted as dashed black lines.
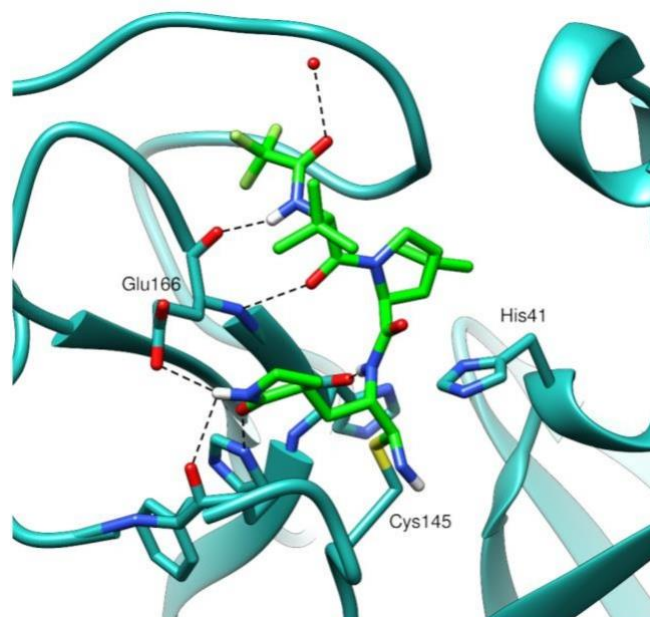
Furthermore, the Neq1183 thermodynamic signature was equal to the desired one (Fig. 9A), in which all the thermodynamic parameters contributed positively to the interactions with the target enzyme (Fig. 23A). However, for Neq1184, while enthalpy favors the free energy positively, even more than Neq1183, entropy destabilizes the complex (Fig. 23B). Lastly, the K_D and K_i values were very close, assuring that these values can be validated orthogonally.

Figure 23: Thermodynamic profiles for (A) Neq1183 and (B) Neq1184.
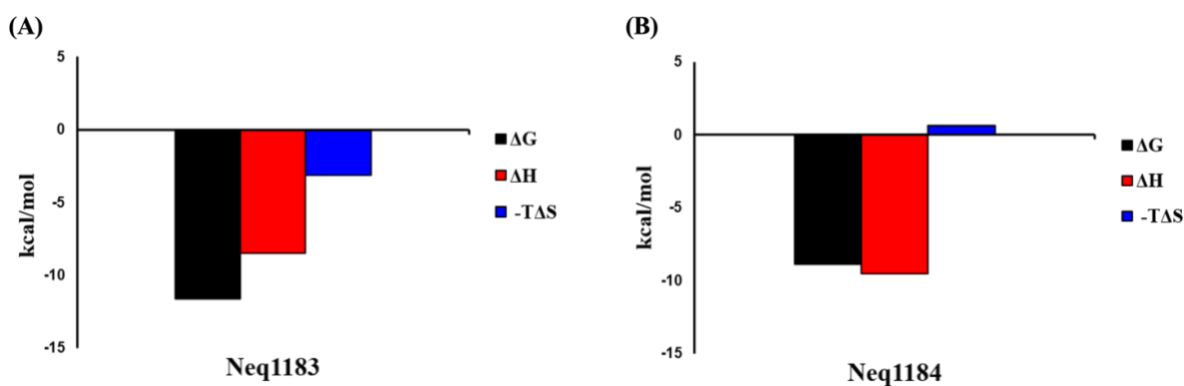
4.1.8   Kinetic Evaluation of a Series of Dipeptidyl Nitriles Against WT M^{pro} and VOCs

The K_i assays were performed with different dipeptidyl nitriles against all the M^{pro} variants to evaluate the effects of different moieties at the P3 position. The compounds tested in this section were synthesized at Prof. Ph.D. Rich Williams laboratories at the The Patrick G. Johnston Centre for Cancer Research, Queen's University of Belfast in the UK. The data obtained for the compound QD88 is shown in Figure 24, and the pK_is for the whole inhibitor series are shown in Table 6.

Figure 24: K_i determination for QD89 against WT M^{pro}. The data was fitted using the Morrison equation 5. K_M was set as 20 μM and enzyme concentration in the assay was 25 nM.
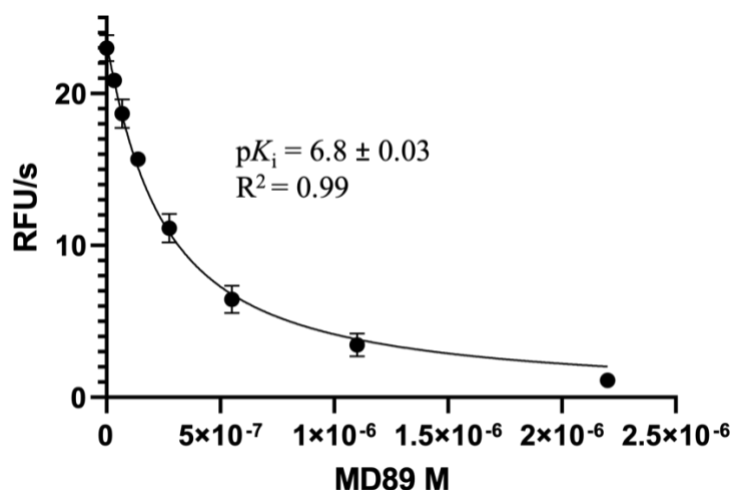
Table 6: pK_i determination for different dipeptidyl nitriles with WT M^{pro} and VOCs.

Inhibitors	pK_i			
	WT	Omicron	Beta1	Beta2
MD78	7.4 ± 0.04	7.6 ± 0.02	7.1 ± 0.02	7.6 ± 0.01
MD79	6.9 ± 0.03	6.8 ± 0.02	6.7 ± 0.05	6.8 ± 0.03
MD80	7.7 ± 0.01	7.7 ± 0.02	7.8 ± 0.03	7.9 ± 0.02
MD86	6.9 ± 0.02	6.7 ± 0.04	6.7 ± 0.02	6.9 ± 0.04
MD87	7.7 ± 0.04	7.6 ± 0.04	7.5 ± 0.04	7.8 ± 0.01
MD88	7.7 ± 0.01	7.7 ± 0.03	7.6 ± 0.01	7.8 ± 0.02
MD89	6.8 ± 0.03	6.8 ± 0.05	6.7 ± 0.04	7.0 ± 0.03
MD90	6.7 ± 0.03	6.8 ± 0.03	6.8 ± 0.04	7.0 ± 0.02
MD91	7.9 ± 0.05	7.8 ± 0.06	7.9 ± 0.04	8.0 ± 0.02
MD92	7.4 ± 0.03	7.9 ± 0.01	7.8 ± 0.03	8.1 ± 0.01

4.1.8.1   MMP/SAR Analysis Aiming P3/P4 Substitution

The inhibitors tested against the enzymes are all matched molecular pairs where a single atom or group was substituted for another at the P3/P4 position. Thus, facilitating a detailed investigation of their structural effects on inhibition. The P3 position carries a biphenyl group with a fluorine, chlorine, or methoxy group at the ring's *ortho, para and meta positions*. The SAR and MMP analysis considered variations in affinity associated with a structural change, denoted as ΔpK_i. The ΔpK_i is calculated through the differences in affinity between two molecules (e.g., $\Delta pK_i = pK_i$ (Y) - pK_i (X)). A ΔpK_i lower than 0.2 log units was considered insignificant and a ΔpK_i equal to or greater than 0.8 log units was considered significant for affinity or selectivity when compared between different targets.

The K_is (Table 6) did not suffer significant variations when comparing the different M^{pro} variants, supporting the results found in the previous sections, when analyzing the catalytic efficiencies of the different variants. An inhibitor designed for the first strain of SARS-CoV-2 will work for the other mutants with similar affinities. Despite the similarities in affinity for all the variants, we notice that the inhibitors performed slightly better for the Beta 2 variant carrying two mutations, with compound MD92 having a difference in pK_i of + 0.7 log units compared to the WT protease. Therefore, the compounds of Table 6 were analyzed regarding the effects of the structural changes made at the P3/P4 position for the WT M^{pro}.

Initially, two MMPs were analyzed, compounds MD78 → MD88 (Fig. 25A) and MD78 → MD92 (Fig. 25B).

Figure 25: MMPA for the inhibitors (A) MD78 → MD88, and (B) MD78 → MD92. The ΔpK_i values reflect structural modification targeting the WT M^{pro}. Orange arrows indicate a slight increase in affinity, while the gray arrow represents no change in affinity.

(A)

(B)
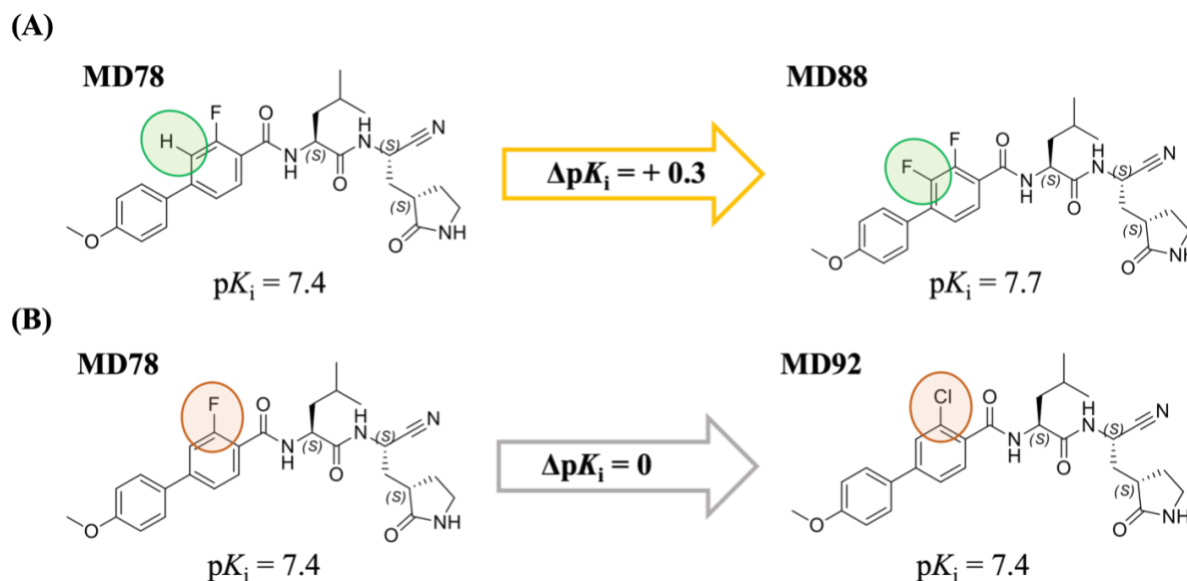
M^{pro} possesses a distinct active site region with a small S3 pocket readily exposed to the solvent. Consequently, the nature of the group at the P3 position of the inhibitor typically determines its accommodation within the S4 pocket. Therefore, given that all the inhibitors examined in this study feature a bulky P3 moiety, it is likely that they will predominantly interact with the S4 pocket of M^{pro}. In addition, all the inhibitors carry a γ-lactam at P1, which mimics the native P1 glutamine of the M^{pro} substrate, and leucine at P2 for better affinity (Fig. 3).

The pairs shown in Fig. 25 had none or very little gain in affinity with the substitutions made. In Fig. 25A, the change of a hydrogen to a fluorine atom led to no significant change in affinity. This bioisosteric substitution of hydrogen to a fluorine atom is a common approach used in medicinal chemistry to prevent metabolic oxidation. Likewise, when substituting a fluorine with a chlorine atom (Fig. 25B), no gain in affinity was observed.

Next, an MMP analysis was performed for the inhibitors MD79 → MD80 and MD91 → MD92 shown in the figure below.

Figure 26: MMPA for the inhibitors (A) MD79 → MD80, and (B) MD91 → MD92. The ΔpK_i values reflect structural modification targeting the WT M^{pro}. Green arrows indicate a significant increase in affinity, while the red arrow indicate a decrease in affinity.

(A)

MD79 $\Delta pK_i = + 0.9$ MD80

$pK_i = 6.9$ $pK_i = 7.8$

(B)

MD91 $\Delta pK_i = - 0.5$ MD92

$pK_i = 7.9$ $pK_i = 7.4$
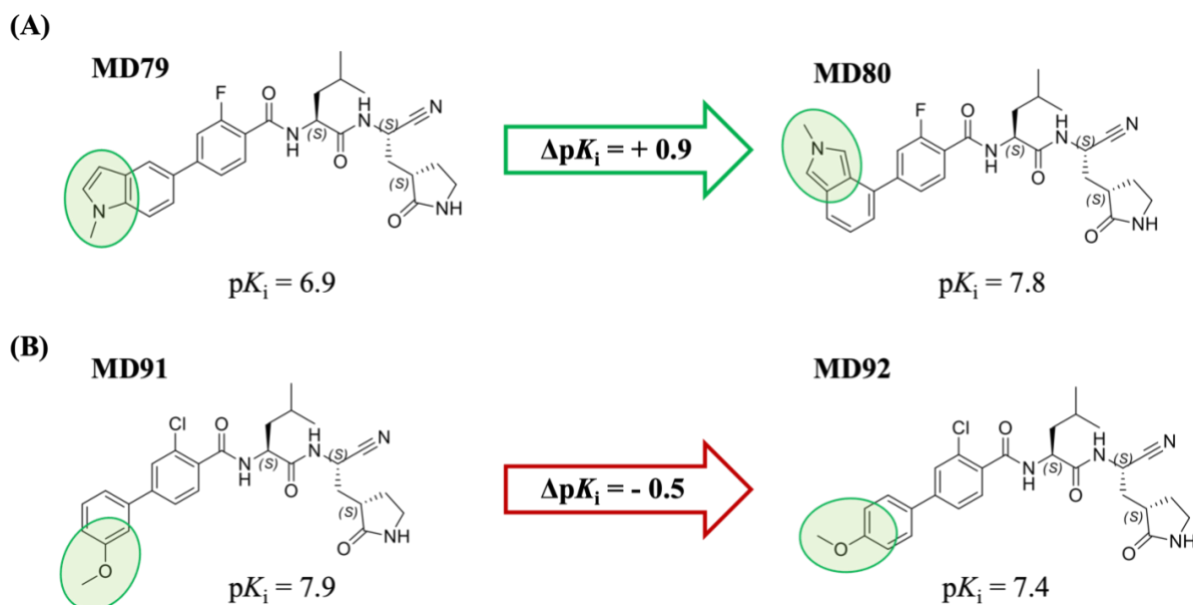
The MMP showed in Fig. 26A had a significant change in affinity of 0.9 log units when substituting an indole group with an isoindole. This suggests that the isoindole group likely forms more efficient interactions with the S4 pocket of the protein. Both groups must be performing π-alkyl interactions with Pro168 and Ala191. However, the conformation of the P3 group in the MD79 compound, suggest that the indol group could be pointing outside the S4 pocket, unlike MD80 where the change to the isoindole must have shifted the P3 moiety more towards the interior of the protein S4 binding pocket, resulting in a more favorable fit. In Fig. 26B, the methoxy group shifted from the *meta* to the *para* position of the aromatic ring, leading to a decrease in affinity. This decrease is probably related to a loss or less effective interaction with the Gln192 residue at the S4 pocket. In addition to the MMPA analysis, a double transformation cycle was done to identify structural modifications that enhance the compound's affinity for the target. This analysis focused on assessing additivity in the SAR analysis, as depicted in Fig. 27. Additivity in SAR was assessed by comparing a ΔpK_i value resulting from a double structural transformation of a compound pair with the sum of the ΔpK_i values obtained from the individual transformations within these pairs. This comprehensive additivity analysis accounts for all the processes involved when making structural modifications to a molecule, including alterations in the types of molecular interactions and their physicochemical properties.

Figure 27: Double transformation cycle for a series of compounds targeting WT M^{pro}. Green arrows indicate a gain in affinity and orange arrows no significant gain. All modifications made are at the P3/P4 position of the molecule.
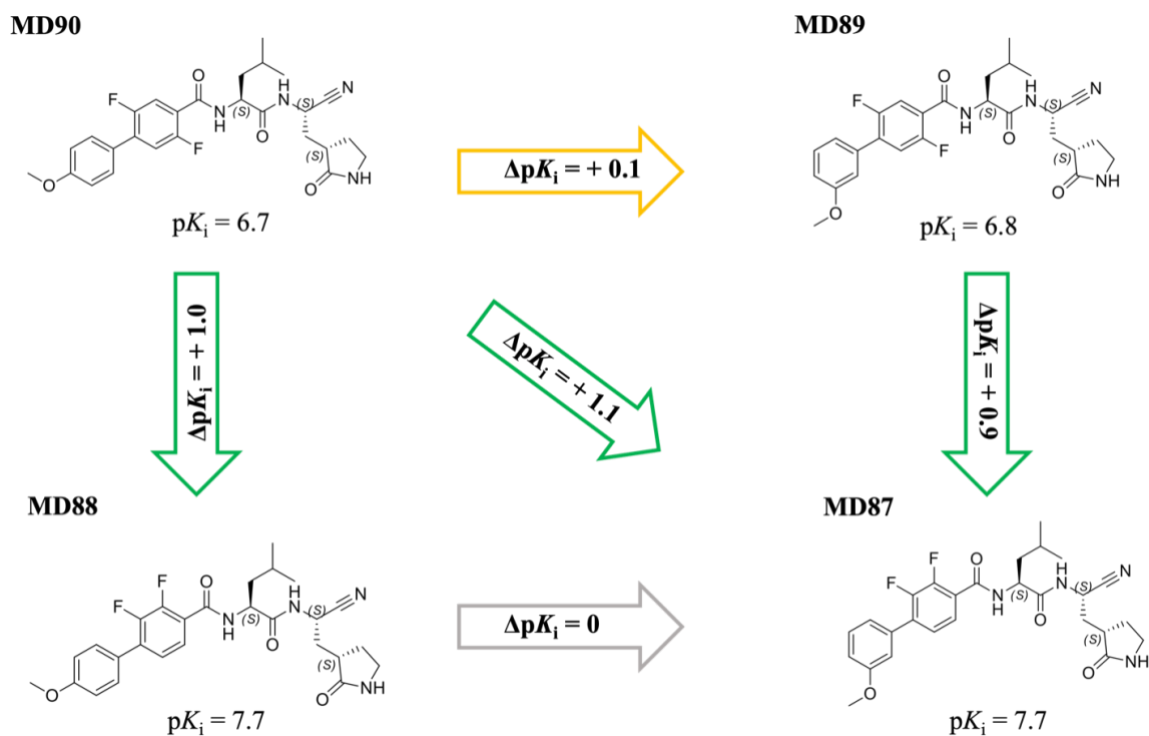
The impacts of structural transformations can be classified into three categories: additive, superadditive, or subadditive effects. A superadditive effect occurs when the ΔpK_i resulting from multiple structural transformations exceeds the sum of the individual pairs' ΔpK_i values. Conversely, a subadditive effect is observed when the combined affinities of individual structural transformations are lower than the sum of the individual pairs' ΔpK_i values. An additive effect occurs when the individual and multiple structural transformations yield the same affinity. Table 7 shows the differences in pK_i for all the VOCs.

Table 7: ΔpK_i values for WT M^{pro} and VOCs from the SAR analysis of Figure 27, the notation [X → Y] indicates a structural modification. Diagonal transformation is highlighted in green.

Structural Modification	WT	Omicron	Beta1	Beta2
MD90 → MD89	+ 0.1	0	+ 0.1	0
MD90 → MD88	+ 1.0	+ 0.9	+ 0.8	+ 0.8
MD89 → MD87	+ 0.9	+ 0.8	+ 0.8	+ 0.8
MD88 → MD87	0	- 0.1	- 0.1	0
MD90 → MD87	+ 1.1	+ 0.8	+ 0.7	+ 0.8

An additive effect was observed for the WT and Beta2 M^{pro} with the modification from MD90 to MD87 having the same change in ΔpK_i as the sum of the individual transformations MD90 $\rightarrow$ MD89 and MD90 $\rightarrow$ MD88. As for the Omicron and Beta1, the structural modifications led to a subadditive effect, with the modification from MD90 to MD87 having a lower ΔpK_i than that of the sum of the individual transformations MD90 $\rightarrow$ MD89 and MD90 $\rightarrow$ MD88. The addition of a fluorine atom at the *ortho* position of the aromatic ring leading to compound MD88 with two fluorine atoms had the most significant change in affinity (+ 0.9 log units). In medicinal chemistry, the presence of a fluorine atom exerts a significant influence on factors such as bioavailability, metabolic stability and toxicity [100]. Its noteworthy that the introduction of a fluorine atom onto an aromatic ring at the *ortho* position, typically results in an increase of potency [101]. The reasons behind this observation are broad, although not definitively established, as the fluorine atom can impact the properties of an aromatic ring in several ways. Thus, we can hypothesize that the modification MD90 $\rightarrow$ MD88 that is the key change for the gain in affinity, is due to the crucial substitution of a hydrogen for a fluorine atom at the *ortho* position. This substitution suggests that the remaining hydrogens of the ring, especially the *ortho* hydrogen, has a lower electronic density which makes the hydrogen more positive due to the fluorine's electronegativity.

Consequently, this lower electronic density may lead to the formation of a weak hydrogen bond between the hydrogen and the carbonyl group of the Glu166 main chain. Additionally, Budzik *et al.*[102] proposed that placing a fluorine atom in an *ortho* position relative to an amide fragment can induce the formation of an intramolecular hydrogen bond with the amide hydrogen. This contributes to stabilizing a favorable conformation of the molecule.

This discussion is primarily based on hypothesis derived from the analysis of M^{pro}'s active site region, as by the time of writing this work, no crystal structures had been obtained/analyzed for these compounds, even though they are being carried out.

4.1.8.1.1 SARS-CoV-2 M^{pro} and Human Cathepsin L Dual Inhibition

The dipeptidyl nitriles studied in the section above were tested against human cathepsin L to evaluate the compound's specificity towards a host protease. The resulting pK_is are presented in Table 8. It's evident that there is a lack of selectivity between the two proteases. However, this lack of selectivity may not be unfavorable in this scenario. In fact, it could be advantageous because the dual inhibition of these enzymes holds significance given cathepsin L's role in virus cell entry during infection.

Table 8: pK_i determination for different dipeptidyl nitriles against WT M^{pro} and human Cathepsin L.

Inhibitors	pK_i	
	WT M^{pro}	*h*CatL
MD78	7.4 ± 0.04	7.2 ± 0.01
MD79	6.9 ± 0.03	7.1 ± 0.02
MD80	7.7 ± 0.01	7.1 ± 0.04
MD86	6.9 ± 0.02	7.2 ± 0.02
MD87	7.7 ± 0.04	7.2 ± 0.03
MD88	7.7 ± 0.01	7.3 ± 0.01
MD89	6.8 ± 0.03	7.1 ± 0.02
MD90	6.7 ± 0.03	7.0 ± 0.07
MD91	7.9 ± 0.05	7.2 ± 0.01
MD92	7.4 ± 0.03	7.3 ± 0.02

Interestingly, all the inhibitors shared the same affinity towards *h*CatL, with minor variations in the pK_i. Both proteases exhibit distinct catalytic sites, with *h*CatL having a well-defined S3 pocket where inhibitors may not be performing essential interactions, in contrast to M^{pro}. This particular feature may explain why variations in the P3 position of the compounds do not affect affinity. A bulky and hydrophobic P3 moiety would likely engage with the Leu69 residue in the *h*CatL S3 pocket through van der Waals interactions. Consequently, all inhibitors would fit the S3 pocket and interact similarly with *h*CatL, a pattern not observed in the case of M^{pro}. Therefore, considering the differences between the proteases, the dual inhibition observed for these compounds is very promising.

4.1.9 Kinetic Evaluation of a Series of Peptide Mimetic Inhibitors Against WT M^{pro}

Kinetic characterization of peptide mimetic compounds was performed against the WT SARS-CoV-2 M^{pro}. The compounds were synthesized at the NEQUIMED laboratories by Ph.D. Felipe Cardoso Prado Martins. In this series of compounds, four different warheads were used, which differs from the previously studied inhibitors with nitrile as a warhead. Consequently, the inhibitors were analyzed regarding their reversibility prior to the determination of their potencies.

The reversibility or dissociation assay considers the dissociation of the enzyme-inhibitor (EI) complex, leading to the full recovery of enzyme activity. This assay can be performed through various methods, such as jump dilutions, dialysis, and ultrafiltration [66]. In this work, a jump-dilution assay was performed to assess the reversibility of the enzyme (M^{pro}) when bound to its respective inhibitor. In this experimental setup, M^{pro} and the inhibitor were incubated for 30 min. Subsequently, a 10-fold dilution was performed before examining the enzyme activity. The final concentrations in the assay were fixed as 10 nM of M^{pro} and 20 nM of inhibitor. This

assay design is due to the 1:1 binding ratio between M^{pro} and the inhibitors. Thus, in theory, equivalent concentrations of enzyme and inhibitor could have been used to analyze reversibility. However, an excess of inhibitor was used to ensure complete binding to the enzyme. Figure 28 shows the residual activity, represented in relative fluorescence units (RFU), observed for M^{pro} against four compounds bearing distinct warheads.

Figure 28: Dissociation assay with inhibitors bearing distinct warheads against the WT SARS-CoV-2 M^{pro}. The inhibitor Neq1223 bears an oxime as warhead, Neq1229 a vinyl sulphone, Neq1230 a fluorine vinyl sulfone, and Neq1236 an oxime carbamate warhead. The control is composed of M^{pro} and substrate. The assay was conducted for 30 min.
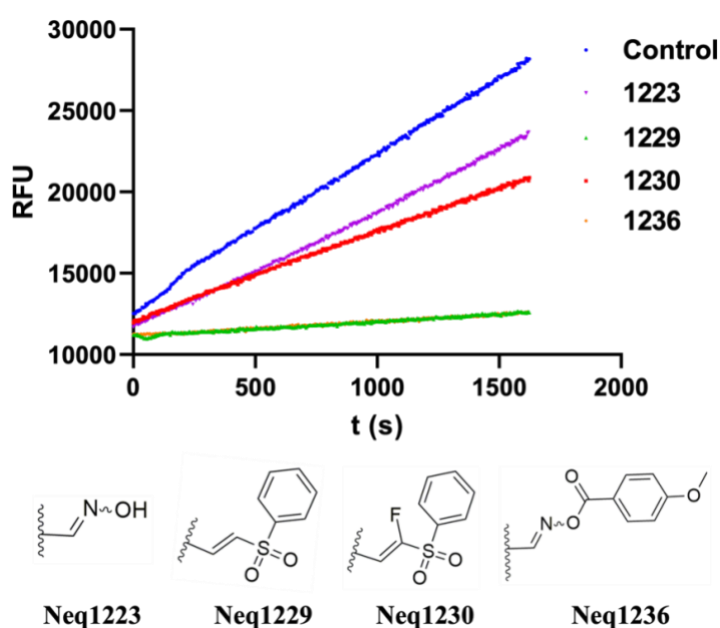
The compounds Neq1223 and Neq1230 displayed a reversible mechanism of inhibition. Notably, the obtained curves for these two compounds exhibit a similar behavior as the control despite its lower slopes. This is expected as a reversible inhibitor makes the enzyme less efficient kinetically. In addition, compounds Neq1229 and Neq1236 demonstrated an irreversible mechanism of inhibition, with close to no enzyme residual activity. These results were expected, especially for the oxime and vinyl sulfone warheads that are widely studied.

Given the distinct mechanism of inhibition displayed by these molecules, different assays were conducted to assess their affinity against M^{pro}. Thus, the standard K_i assay was performed with reversible inhibitors containing the oxime and the fluor vinyl sulfone. However, for the irreversible inhibitors, as K_i cannot be determined due to the ligand's nature, the constant k_{inact}/K_I was calculated instead.

4.1.9.1   Assessing the Potency of Irreversible Inhibitors

To calculate the k_{inact}/K_I for the irreversible inhibitors Neq1229 and Neq1236, the first step was to calculate the observed first-order rate constant k_{obs} (Fig 29A). This involved calculating k_{obs} at various inhibitor concentrations. Subsequently, by plotting k_{obs} against inhibitor concentration, as depicted in Figure 29B, the k_{inact}/K_I value could be obtained.

Figure 29: (A) progress curve for the inhibition of M^{pro} by Neq1229. k_{obs} was determined by using equation 2 for each inhibitor concentration. (B) The inhibitor's potency was obtained by linear regression of k_{obs} as a function of inhibitor concentration.
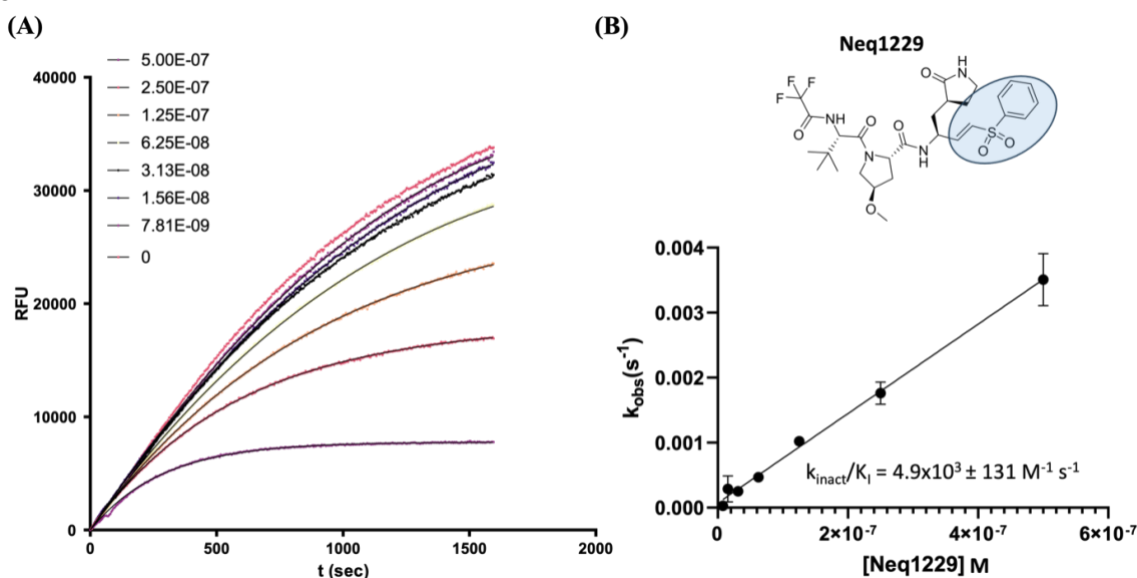
By analyzing the k_{obs} vs [I] plot, it becomes possible to discern the type of mechanism of inhibition displayed by Neq1229. The inhibitors Neq1229 and Neq1236 (appendix) exhibit a direct mechanism of inactivation, as evidenced by the linear relationship observed in the graph. Interestingly, both inhibitors display a notably high value of k_{inact}/K_I, with Neq1236 having the highest value of 7.8×10^4 $M^{-1}s^{-1}$. Since these two inhibitors share the same structure, with the only change being the warhead, this difference in potency can be attributed mainly to the variation of the warhead. Therefore, the oxime carbamate is a more effective warhead to inhibit M^{pro} than the classic vinyl sulfone. This finding offers exciting insights into the use of warheads to modulate potency. Hence, Neq1236 emerges as an attractive compound for potential development as an activity-based probe (ABP), owing to its irreversible warhead and high potency in the same range as published ABPs [103].

4.1.9.2   MMP/ SAR Analysis Aiming P2 Substitution

The inhibitors investigated in this section were tested mainly against the WT M^{pro}. This choice was based on the results obtained in the previous sections, where the compounds exhibit similar inhibitory effects on the variants as they do on the WT M^{pro}. The inhibitors investigated constitute matched molecular pairs, where a single atom or functional group was substituted for another at the P2 position. All compounds have an oxime or nitrile as warhead and, thus, are all reversible covalent inhibitors. The inhibitors share the same groups as Neq1183 (Fig. 4) at P1, P3 and P4. The P1 is a γ-lactam, and the P3 and P4 are a tert-butyl and a trifluoro acetamide, respectively. The P2 was varied with different moieties to evaluate the effects in potency. The data obtained for the compound Neq1223 is shown in Figure 30. Table 9 presents the calculated pK_is for this series of compounds against the WT M^{pro} and hCatL.

Figure 30: K_i determination for Neq1223 against WT M^{pro}. The data was fitted using the Morrison equation. K_M was set as 20 μM and enzyme concentration in the assay was 25 nM.
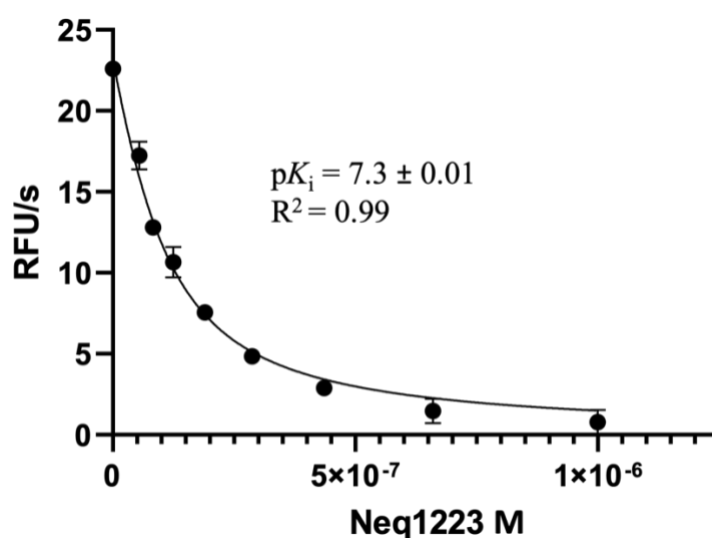
Table 9: pK_i determination for a range of peptide mimetic compounds against the WT M^{pro} and hCatL.

	pK_i or Residual Activity (%)[a]	
Inhibitors	**WT M^{pro}**	**hCatL**
Neq1220	6.8 ± 0.01	>95%
Neq1223	7.3 ± 0.01	60%
Neq1224	7.3 ± 0.08	>95%
Neq1227	5.6 ± 0.02	86%
Neq1228	7.0 ± 0.02	81%
Neq1230	8.1 ± 0.01	>95%
Neq1233	8.8 ± 0.01	76%
Neq1234	7.0 ± 0.03	>95%
Neq1235	6.7 ± 0.02	81%
Neq1248	8.9 ± 0.09	85%

No significant affinity was observed for the compounds against hCatL, showing high selectivity. This selectivity can be attributed to incorporating the proline-like moiety at the P2 position of the inhibitors. The broader and differently shaped S2 pocket of hCatL resulted in distinct or ineffective binding interactions between P2-S2 in this enzyme.

The MMP/SAR analysis for the compounds bearing an oxime as warhead from Table 9 is depicted in Figure 31. The P2 of these compounds was varied, in which different groups based on a proline amino acid were explored. As seen in the SAR, the P2 modification of Neq1220 to Neq1223 and 1224 led to a slight increase in potency of $\Delta pK_i = + 0.5$ log units. The compounds may be performing hydrophobic interactions with the S2 pocket of M^{pro}. In the case of Neq1223, the aromatic ring added to the proline-like group could interact with His41 through a π-stacking between the aromatic rings. The same must happen with Neq1224, with the alkene moiety added to the proline. This hypothesis is based on the mode of binding (MoB) displayed by Ensitrelvir, a potent M^{pro} inhibitor ($K_i = 9$ nM) that has a long and flexible aromatic moiety at P2 [104]. In addition, Neq1224 could be performing a hydrogen bond with the side chain of Gln189, which can also occur for Neq1220.

Figure 31: MMP/SAR analysis of compound Neq1220 with six compounds varying P2. The ΔpK_i values corresponds to the structural modification targeting the WT M^{pro}. Green boxes indicate an increase in affinity, while the orange and gray boxes indicate a moderate to no change in affinity.
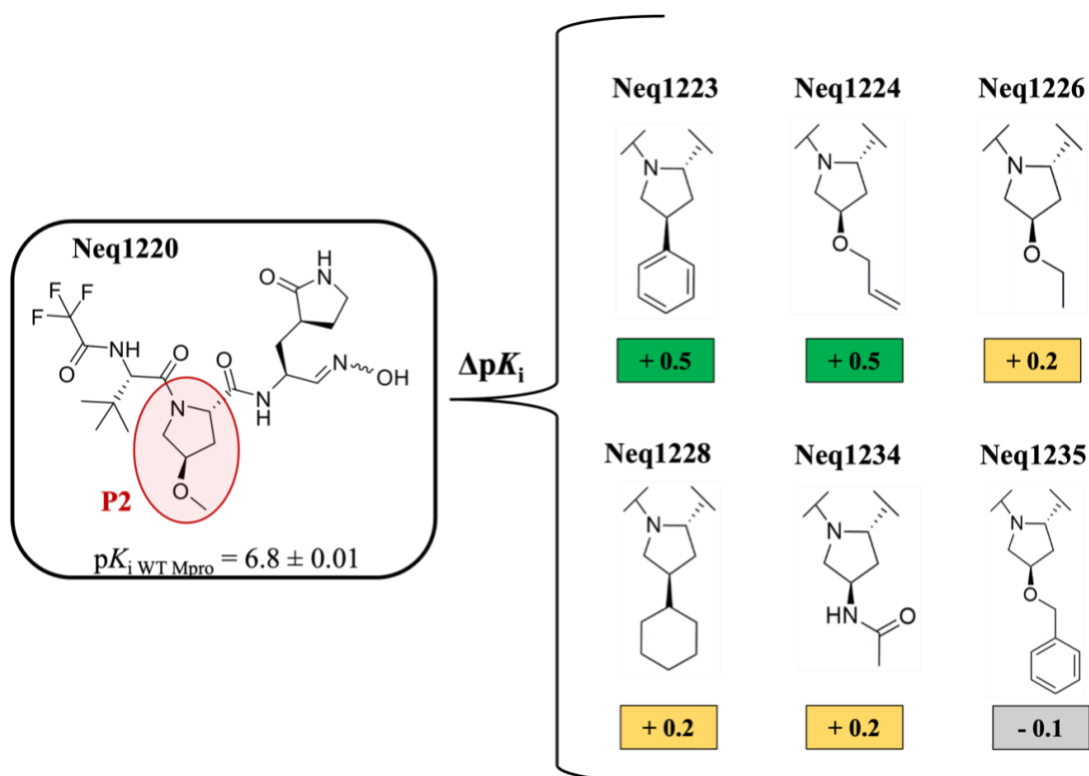
The modification of Neq1220 to Neq1226 and Neq1228 resulted in a moderate increase of $\Delta pK_i = +0.2$ log units. These results suggest that the loss of the double bond and associated aromaticity played a role in this moderate difference in potency. Interestingly, while no π-stacking interactions can happen, hydrophobic interactions appear to exert a more significant influence on affinity within the S2 pocket of the protein. Regarding Neq1234, the reduction in hydrophobic groups was compensated by introducing an amide, which can engage in hydrogen bonding with the side chain of Gln189. In contrast, the substitution Neq1220 → Neq1235, which involved elongating the group, did not alter affinity. Despite the presence of an aromatic ring, the addition of an aliphatic chain between the proline and the ring may have caused the ring to not be properly positioned in the S2 pocket. Consequently, this adjustment likely prevented π-stacking interactions, as observed in compound Neq1223. It's important to highlight that this section is based on hypothesis by analyzing M^{pro}'s active site residues as there are no available crystal structures depicting the interactions between these compounds and the protease. Furthermore, docking and molecular dynamics calculations are undergoing for these set of compounds, which can provide a better understanding of the interactions with the protease.

The oxime warhead presented to be an interesting approach for the development of M^{pro} inhibitors. However, since our group mainly focuses on nitriles, the substitution of the oxime warhead for a nitrile was made to evaluate the influence of the warhead in the affinity. For this substitution, compounds Neq1220 and Neq1223 were chosen (Fig. 32).

Figure 32: Warhead substitution of (A) Neq1220 → 1233, and (B) Neq1223 → Neq1248. Green arrows indicate an increase in pK_i.
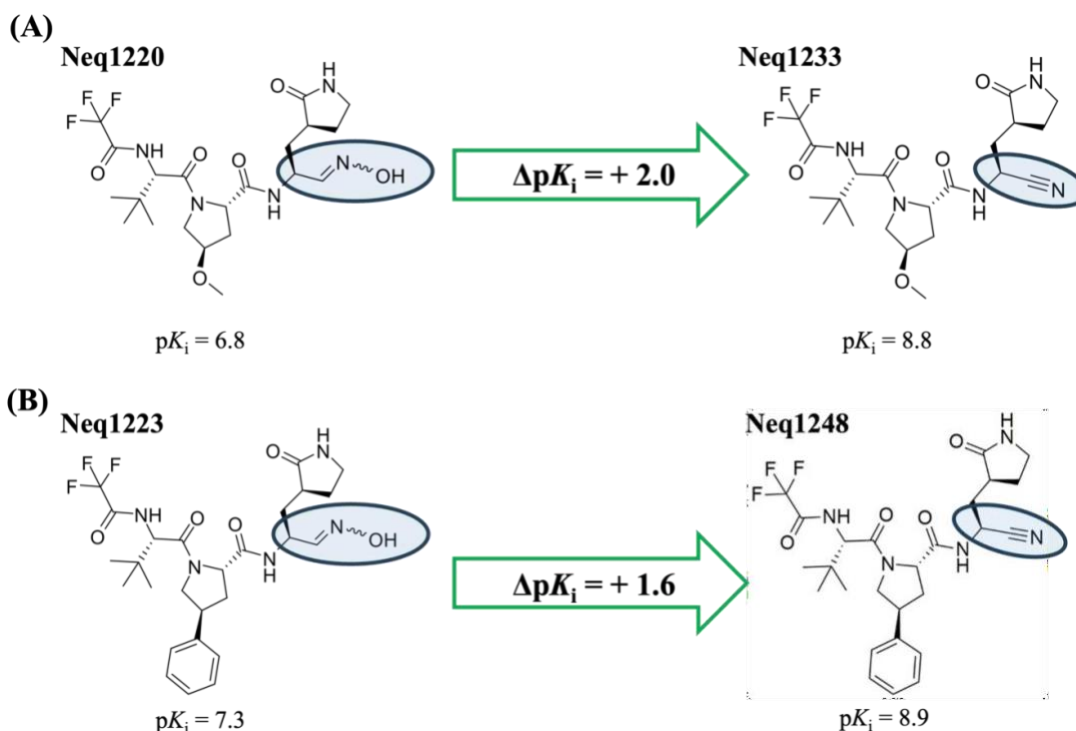
Replacing the oxime warhead for a nitrile, as seen for compounds Neq1220 to Neq1233, resulted in an activity cliff with a $\Delta pK_i = +2.0$ log units. However, as depicted in Fig. 32B, the warhead replacement, despite having a substantial change in affinity, did not lead to an activity cliff. Since both compounds share identical pK_i values but feature different substituents at P2, it is reasonable to infer that the gain in affinity is mainly due to the warhead substitution. The nitriles generally exhibit higher reactivity than oximes due to the nature of their chemical structure. This is mainly due to the lower electrophilicity of oximes when compared to nitriles that arises from the relative electronegativities of the atoms involved in the binding process with the target protein. In oximes, the presence of a double bond between carbon and nitrogen, along with the oxygen atom's influence, reduces the electrophilic character of the adjacent carbon atom. As for nitriles, the triple bond, that is highly polarized, allows for resonance structures that further stabilize the partial negative charge on nitrogen and enhance the electrophilicity of the carbon atom. Thus, for these reasons and the obtained results, the nitriles seem to be a better warhead choice for SARS-CoV-2 M[pro] inhibition.

Additionally, the best inhibitors from this series were selected for testing against the VOCs to support the hypothesis that the same inhibitor inhibits the WT M[pro] and VOCs equally. The resulted pK_i values for these compounds are reported in Table 10.

Table 10: pK_i determination for different peptide mimetic inhibitors against WT M^{pro} and VOCs.

Inhibitors	pK_i			
	WT	Omicron	Beta1	Beta2
Neq1220	6.8 ± 0.01	7.0 ± 0.02	6.9 ± 0.02	6.9 ± 0.01
Neq1223	7.3 ± 0.01	7.6 ± 0.03	7.4 ± 0.02	7.3 ± 0.03
Neq1233	8.8 ± 0.01	8.8 ± 0.02	9.0 ± 0.01	8.9 ± 0.02

The pKi values reported in Table 10 against the WT M^{pro} and VOCs for different inhibitors are as expected. The compounds Neq1220 and Neq1223 bearing the oxime warhead showed the same behavior as the nitriles against the VOCs.

These findings have demonstrated that the WT M^{pro} and its variants can be effectively inhibited using the same compound that can feature distinct warheads. These results present a promising opportunity for developing new antiviral drugs that can be effective against emerging variants of the SARS-CoV-2 virus.

4.2 Cell-based Assay

The WT SARS-CoV-2 M^{pro} was transfected into HEK 293 T cells. The experiment aimed to obtain an active M^{pro} in a cell line to further investigate some of the compound's lipophilicity. Following transfection, cell lysates containing 3 and 6 million cells were prepared and treated with specific protease/proteasome inhibitors. A BCA assay determined the total protein concentration within the lysates, yielding protein concentrations ranging from 9 to 11 mg. Thus, a western blot of the lysate was performed in order to identify the M^{pro} expression with specific antibodies (Fig. 33). The M^{pro} construct used to transfect has a sequence of M^{pro}-2xStrep-tag-3xFlag-tag (40.9 kDa) (Appendix). So, the first antibody recognized the Flag-tag and the second the Strep-tag.

Figure 33: HEK 293 T transfected, Untransfected, and with empty vector, cell lysates containing 3 and 6 million cells. The expressed M^{pro} is shown in green at 40.9 kDa.
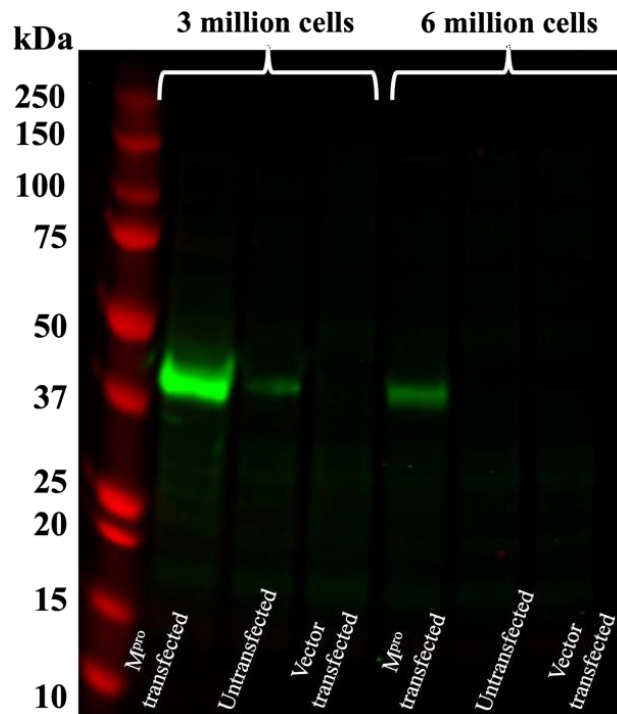
Upon examination of the western blot (WB) (Fig. 33), the expressed M^{pro} can be seen around 40.9 kDa, indicating that the transfection occurred successfully. Notably, the 3 million cell lysate exhibited a higher protein expression than the 6 million cell lysate, suggesting that lower cell confluency prior to transfection is more efficient.

To assess the activity of the expressed M^{pro} in the cells, an assay employing the activity-based probe (ABP) Bodipy-PEG (4)-Abu-D Tyr-Leu-Gln-VS (Fig. 34A) was designed. A lysate sample was incubated with 1 μM of the ABP for 30 min at 37 ℃. This incubation was carried out in PIPES pH 7.4 buffer, with and without 0.7 M of sodium citrate (Fig. 34BC), to evaluate, if, in case of inactivity of the protein, if the same could be activated in the presence of the salt. This would be observed with the binding of the probe.

Figure 34: Activity analysis of expressed M^pro via SDS-PAGE and WB using an ABP. (A) ABP Bodipy-PEG (4)-Abu-D Tyr-Leu-Gln-VS with the irreversible warhead vinyl sulfone highlighted in blue.(B) SDS-PAGE of the lysates with and without 0.7 M sodium citrate. The M^pro bound to the ABP can be seen in blue around 38 kDa. (C) WB of the lysates with and without citrate following addition of the antibodies. The M^pro bound to the ABP can be seen in blue around 38 kDa, and the M^pro bound to the antibody can be seen in red at 40.9 kDa.
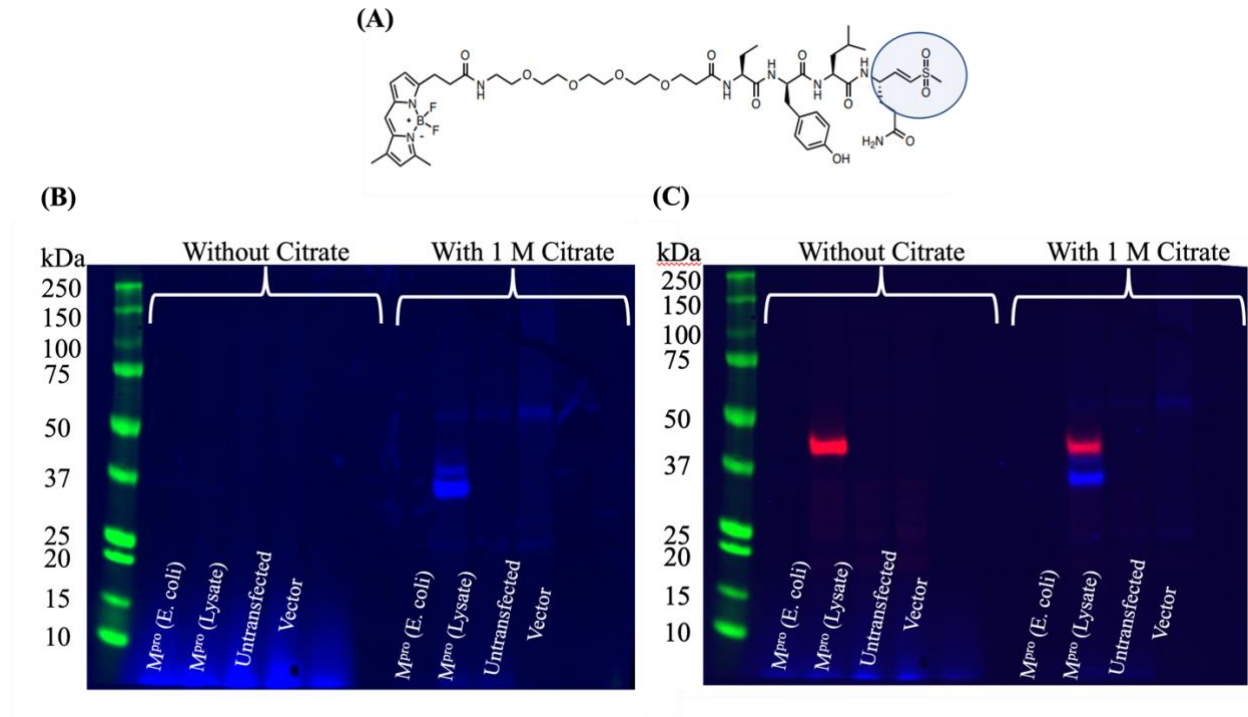
The activity assay of the expressed M^pro in HEK 293 T cell lysates showed that the enzyme was inactive. This can be seen in the SDS-PAGE gel and WB (Fig. 34 BC), where no signal of the M^pro bound to the ABP (blue signal) was observed when the buffer was not supplemented with sodium citrate. However, the signal of bound M^pro to the ABP was seen upon the addition of sodium citrate. Interestingly, when the protease was activated with sodium citrate, two signals appeared in the WB (Fig. 34C). The first signal in red, related to the full-length protease with all the tags (40.9 kDa), while the second signal in blue, appeared to be associated with the M^pro lacking the 3x Flag-tag (37.9 kDa). This auto-processing of the protease is interesting and can indicate that the number of tags used in the transfected construct can block the dimerization of M^pro, leading to the lack of activity observed. Therefore, the lipophilicity assays were not carried out due to the absence of an active protease and the impracticality of employing sodium citrate in cell-based assays.

4.3 Kinetic Studies for the Human Cathepsins L, S, and B

In addition to the SARS-CoV-2 M^{pro} studies, a comprehensive evaluation of a series of dipeptidyl nitriles against the human cathepsins L, S, and B was conducted to assess the selectivity of these compounds between the proteases. In order to determine the compound's inhibition constants, the kinetic parameters of these cathepsins were calculated using the fluorogenic substrate Z-FR-AMC, enabling an analysis of their activities.

4.3.1 *Kinetic Characterization for the Human Cathepsins L, S, and B*

The kinetic characterization of the human cathepsins L, S, and B are shown in Table 11. The obtained Michaelis-Menten plot is presented in Figure 35 for *h*CatL; for *h*CatS and *h*CatB, the plots are presented in the appendix section.

Figure 35: (A) Michaelis-Menten plot for *h*CatL against the substrate Z-FR-AMC. (B) Structure of Z-FR-AMC with the fluorogenic group AMC is highlighted in yellow, and the cleavage site by the *h*CatL is highlighted in red.
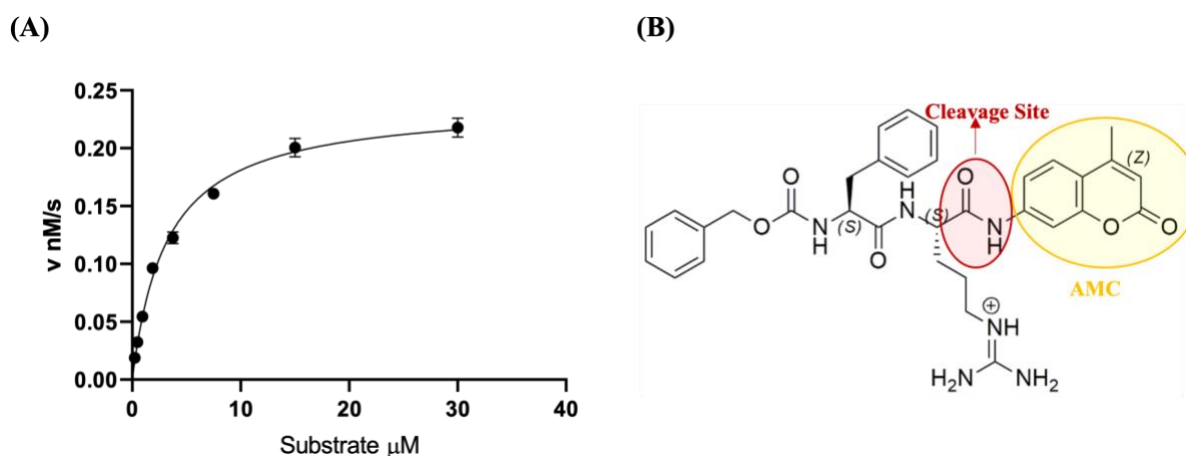
Source: own author

Table 11: Kinetic constants for the three human cathepsins on the Z-FR-AMC. The kinetic constants were obtained in assay buffer as described in the Material and Methods. The table shows the mean and standard deviation of three technical replicates.

Enzyme	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
*h*CatL[a]	3.20 ± 0.09	1.59 ± 0.12	496875
*h*CatS [a]	37.8 ± 1.7	1.31 ± 0.04	34656
*h*CatB [a]	96.8 ± 4.8	12.5 ± 1.13	129132

[a] [*h*CatL] = 1.9 nM; [*h*CatS] = 1.4 nM ; [*h*CatB] = 1.1 nM

The obtained kinetic values agree with those found in the BRENDA database. Cathepsin L and B showed a higher catalytic efficiency with the substrate than cathepsin S. This is interesting, considering that hCatL and hCatS are more similar than hCatB. Therefore, the enzymes were considered fully functional, and the inhibition assays were performed.

4.3.2 Inhibition Investigation for the Human Cathepsins L, S, and B

All compounds studied in this section are different dipeptidyl nitriles. The compounds are similar, although different moieties were explored at P1, P2, and P3. The pK_is calculated for this series are shown in Table 12.

Table 12: Dipeptidyl nitriles tested against the human cathepsins L, S, and B with their respective pK_i values and percentage of inhibition determined at 10 μM.

NEQUIMED Code	pK_i or Residual Activity (%)[a]		
	pK_i hCatL	pK_i hCatS	pK_i hCatB
Neq0400.1	6.4 ± 0.01	6.5 ± 0.09	85%
Neq0569	8.8 ± 0.02	7.4 ± 0.02	87%
Neq0675	7.2 ± 0.03	6.4 ± 0.02	5.4 ± 0.03
Neq0803	8.5 ± 0.02	7.6 ± 0.02	5.2 ± 0.01
Neq0985	8.2 ± 0.06	6.6 ± 0.09	73%
Neq1005	6.6 ± 0.02	5.5 ± 0.08	>95%
Neq1006	6.0 ± 0.01	5.8 ± 0.03	<95%
Neq1007	6.0 ± 0.01	5.5 ± 0.02	<95%
Neq1008	6.2 ± 0.02	6.2 ± 0.02	<95%
Neq1017	6.0 ± 0.01	6.2 ± 0.04	>95%
Neq1024	6.6 ± 0.06	7.1 ± 0.09	88%
Neq1025	6.4 ± 0.03	7.1 ± 0.05	83%
Neq1026	7.8 ± 0.06	7.1 ± 0.05	5.3 ± 0.05
Neq1027	7.5 ± 0.01	7.3 ± 0.07	5.3 ± 0.02
Neq1028	6.0 ± 0.01	6.1 ± 0.05	78%
Neq1029	5.9 ± 0.01	5.5 ± 0.04	67%
Neq1030	5.9 ± 0.01	5.9 ± 0.04	71%
Neq1032	6.5 ± 0.03	5.9 ± 0.01	80%
Neq1036	67%	5.8 ± 0.01	77%
Neq1037	5.7 ± 0.04	6.3 ± 0.09	84%
Neq1043	7.9 ± 0.02	5.9 ± 0.02	92%
Neq1044	9.8 ± 0.01	7.1 ± 0.01	6.3 ± 0.03
Neq1045	8.0 ± 0.06	6.2 ± 0.09	5.7 ± 0.08
Neq1056	5.7 ± 0.04	6.6 ± 0.06	>95%
Neq1057	88%	84%	93%
Neq1058	7.5 ± 0.05	6.2 ± 0.06	88%
Neq1059	7.4 ± 0.03	7.6 ± 0.08	60%
Neq1060	5.1 ± 0.09	>95%	75%
Neq1061	6.2 ± 0.01	5.8 ± 0.09	81%
Neq1062	8.3 ± 0.02	6.8 ± 0.03	5.3 ± 0.07
Neq1063	5.6 ± 0.07	5.7 ± 0.02	>95%

Neq1071	6.6 ± 0.02	6.2 ± 0.02	>95%
Neq1072	6.7 ± 0.01	5.9 ± 0.03	5.2 ± 0.09
Neq1073	5.8 ± 0.01	6.0 ± 0.03	>95%
Neq1075	6.5 ± 0.02	6.6 ± 0.07	79%
Neq1076	7.1 ± 0.04	7.1 ± 0.09	>95%
Neq1079	6.8 ± 0.02	7.0 ± 0.02	>95%
Neq1083	9.1 ± 0.03	8.7 ± 0.02	5.7 ± 0.04
Neq1087	7.4 ± 0.01	6.7 ± 0.02	6.6 ± 0.01
Neq1088	7.1 ± 0.09	6.4 ± 0.09	5.2 ± 0.01
Neq1091	7.0 ± 0.08	6.1 ± 0.04	5.4 ± 0.01
Neq1092	6.8 ± 0.01	6.1 ± 0.02	5.5 ± 0.03
Neq1093	8.1 ± 0.04	6.4 ± 0.09	6.9 ± 0.05
Neq1094	6.8 ± 0.05	5.6 ± 0.01	5.5 ± 0.08
Neq1096	7.9 ± 0.06	6.3 ± 0.01	5.8 ± 0.01
Neq1097	6.6 ± 0.04	5.7 ± 0.05	5.5 ± 0.01
Neq1099	9.4 ± 0.02	6.8 ± 0.04	6.3 ± 0.01
Neq1101	9.4 ± 0.02	7.6 ± 0.04	7.0 ± 0.09
Neq1102	8.0 ± 0.03	6.2 ± 0.04	5.8 ± 0.04
Neq1103	9.6 ± 0.02	6.9 ± 0.05	6.1 ± 0.09
Neq1105	8.1 ± 0.03	7.9 ± 0.01	7.0 ± 0.01
Neq1106	8.7 ± 0.02	7.8 ± 0.01	7.0 ± 0.03
Neq1107	5.4 ± 0.01	5.1 ± 0.03	>95%
Neq1108	5.6 ± 0.04	4.5 ± 0.05	>95%
Neq1109	6.0 ± 0.01	5.5 ± 0.01	5.7 ± 0.09
Neq1110	5.5 ± 0.01	4.9 ± 0.03	68%
Neq1118	6.7 ± 0.05	6.5 ± 0.04	4.9 ± 0.01
Neq1119	8.8 ± 0.02	7.1 ± 0.01	5.4 ± 0.01
Neq1120	7.1 ± 0.02	6.3 ± 0.01	62%
Neq1121	6.1 ± 0.04	6.4 ± 0.05	5.8 ± 0.03
Neq1122	5.8 ± 0.01	5.4 ± 0.05	87%
Neq1123	5.7 ± 0.06	6.1 ± 0.08	73%
Neq1124	5.8 ± 0.01	5.2 ± 0.08	83%
Neq1129	7.4 ± 0.05	6.9 ± 0.02	6.2 ± 0.01
Neq1130	7.3 ± 0.06	6.8 ± 0.06	6.3 ± 0.03
Neq1131	7.4 ± 0.02	7.0 ± 0.01	6.0 ± 0.05
Neq1132	6.8 ± 0.02	6.6 ± 0.03	5.9 ± 0.04
Neq1133	6.3 ± 0.02	6.4 ± 0.05	5.0 ± 0.02
Neq1134	5.2 ± 0.05	5.3 ± 0.04	90%
Neq1135	5.3 ± 0.04	5.4 ± 0.04	94%
Neq1136	5.0 ± 0.03	5.3 ± 0.04	4.1 ± 0.02
Neq1137	5.9 ± 0.03	5.3 ± 0.05	90%
Neq1138	5.2 ± 0.02	5.4 ± 0.05	85%
Neq1139	6.5 ± 0.04	7.0 ± 0.03	85%
Neq1140	6.3 ± 0.06	6.4 ± 0.03	93%
Neq1141	7.9 ± 0.05	7.6 ± 0.02	81%
Neq1142	6.7 ± 0.01	7.3 ± 0.09	5.6 ± 0.02
Neq1143	>95%	>95%	>95%
Neq1144	>95%	5.1 ± 0.03	>95%
Neq1146	5.0 ± 0.07	5.9 ± 0.04	6.0 ± 0.01
Neq1147	4.5 ± 0.02	7.0 ± 0.04	7.4 ± 0.01
Neq1148	8.0 ± 0.06	7.0 ± 0.06	5.3 ± 0.01
Neq1153	8.7 ± 0.07	8.9 ± 0.05	7.5 ± 0.02
Neq1164	4.6 ± 0.05	5.0 ± 0.03	>95%
Neq1165	>95%	>95%	>95%

Neq1166	6.3 ± 0.09	5.5 ± 0.04	87%
Neq1167	6.5 ± 0.03	7.7 ± 0.02	6.1 ± 0.01
Neq1168	6.6 ± 0.02	6.8 ± 0.09	6.2 ± 0.01
Neq1169	7.1 ± 0.02	7.2 ± 0.04	5.3 ± 0.01
Neq1170	>95%	>95%	>95%
Neq1171	8.0 ± 0.02	7.8 ± 0.05	5.1 ± 0.01
Neq1172	7.3 ± 0.07	7.3 ± 0.09	4.5 ± 0.02
Neq1173	79%	93%	>95%
Neq1174	6.9 ± 0.05	7.9 ± 0.03	4.6 ± 0.02
Neq1183	5.7 ± 0.04	85%	91%
Neq1184	>95%	>95%	82%
Neq1188	7.6 ± 0.03	7.9 ± 0.03	6.1 ± 0.02
Neq1189	6.3 ± 0.02	6.7 ± 0.01	5.5 ± 0.03
Neq1193	80%	>95%	94%
Neq1194	85%	>95%	>95%
Neq1195	85%	>95%	>95%

[a] Remaining activity at 10 μM inhibitor concentration obtained in triplicate measurements

A selectivity analysis among the cathepsins was conducted after determining the inhibition constants for the compounds.

4.3.3   Selectivity Study Amid the Human Cathepsins

Selectivity between CPs, especially between cathepsins, can be very hard to achieve [2,53]. The human cathepsins L, B, and S share high sequence identity in the active site region (Figure 36). While the S1 subsite is highly conserved between cathepsins L and S, the *h*CatB has a slightly larger subsite with a Glu122 residue. However, the S2 and S3 subsites demonstrate less conservation among these three enzymes, presenting an opportunity to develop selective inhibitors.

Figure 36: Crystallographic structure of the active site region for the cathepsins L, B, and S. The PDB codes for the enzymes are 2Y2J, 6AY2, and 3OVX, respectively.
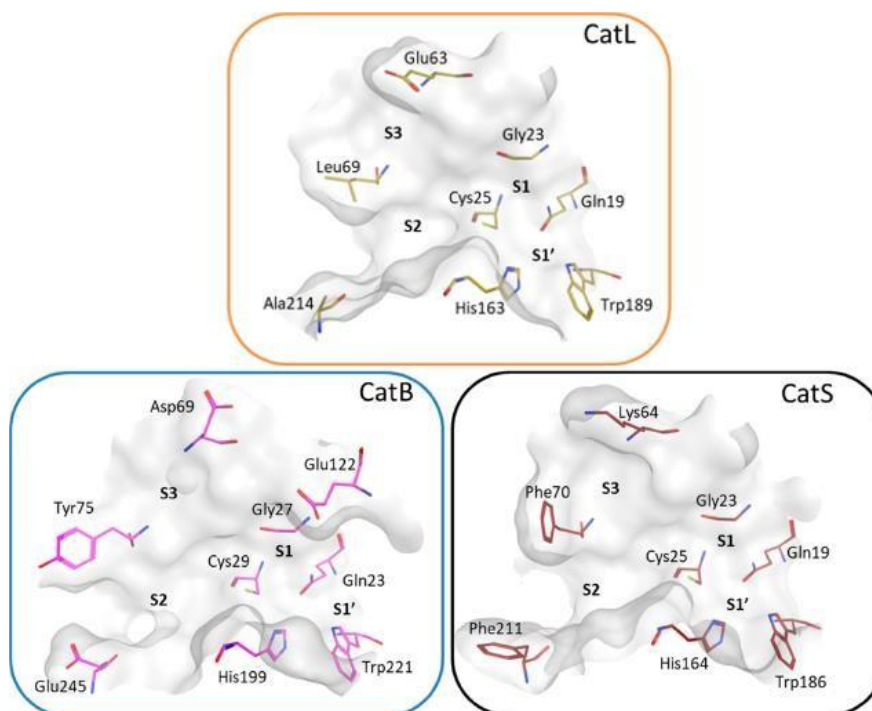
Source: Adapted from Cianni, L. *et al.* Can Cysteine Protease Cross-Class Inhibitors Achieve Selectivity? **J. Med. Chem.,** v. 62, p. 10497–10525, 2019

Therefore, an analysis between the pK_is of the cathepsins was performed. The analysis allows the identification of the preferable chemical fragments for each enzyme. Thus, it provides the knowledge to design new inhibitors. Human cathepsins L and S exhibit higher structural similarity than *h*CatB, so achieving selectivity between these two enzymes is more complicated. Since most of the compounds are more selective for *h*CatL and S over *h*CatB, the study regarding *h*CatB selectivity was not performed. Thus, the analysis between the human cathepsins L and S is depicted in Figure 37. It is essential to note that compounds with a pK_i value of less than 4.0 (cutting range) were not included in the study.

Figure 37: Selectivity analysis between human cathepsins L and S. The compounds circled in blue have differences in pK_i greater than or equal to two log units.
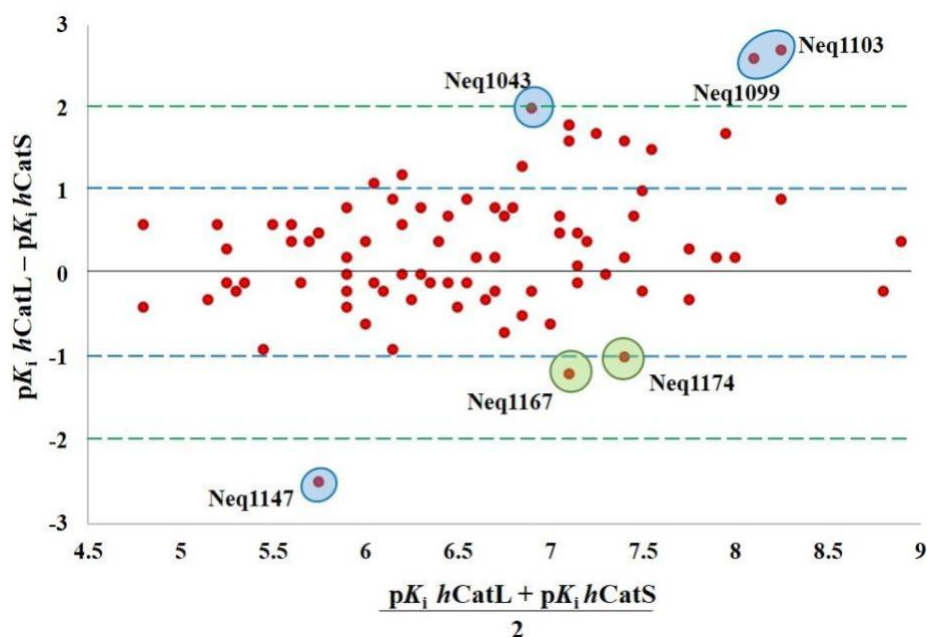
Source: own author

The y-axis shows the differences between the pK_i values obtained for the human cathepsins L and S, and the x-axis shows the mean values between the pK_is of the two enzymes. Compounds with high affinity for the CPs (x-axis) are displayed on the right side of the graph. The highest selectivity is observed on the y-axis, as the farther away from zero, the greater the selectivity. Positive values on the y-axis indicate selective compounds for hCatL and negative values for hCatS.

Figure 37 shows that most compounds are arranged at the top of the graph on the positive y-axis, indicating that most compounds are more selective for hCatL over hCatS. Many compounds display a difference in pK_i between one and two log units, and due to the significant number of compounds and their different structures, an SAR will not be performed. Compounds Neq1043, Neq1099, and Neq1103, in which the pK_i differences were more than two log units, are highlighted in blue circles. Structurally, the three compounds are similar in only one point of interaction with the protein, at the P2 position. The compounds possess an indole moiety at P2, which, as the analysis indicates, is a group that brings high selectivity for hCatL over cathepsins S and B, as shown in Figure 37 and Table 11. The indole is a large and bulky group, and as shown in Figure 36, hCatL's S2 subsite can support groups of that size and volume. The NH atom of the indole ring is performing a hydrogen bond with the Ala214 residue present in the subsite, leading to increased affinity.

As for hCatS, Neq1147, Neq1167, and Neq1174, they displayed a significant selectivity for this enzyme towards hCatL. These compounds have several structural differences, making it difficult to rationalize the reason for the observed selectivity; the only similarity is in the nitrile warhead. Therefore, further studies are needed to determine the reasons for the selectivity displayed by these compounds. A selectivity study between SARS-CoV-2 M^{pro} and the cathepsins was not performed due to the low number of pK_is obtained for M^{pro}.

5 CONCLUSION

The SARS-CoV-2 M^{pro} is an established biological target for coronavirus diseases. As viruses can mutate, understanding the effect of mutations in the sequence of M^{pro} is of great importance to achieve effective inhibition of the target.

The WT SARS-CoV-2 M^{pro} and variants expression was performed and confirmed using SDS-PAGE electrophoresis and enzyme kinetics. The catalytic efficiency of the proteins was improved by using kosmotropic salts, such as sodium citrate, by 36-fold when compared with a buffer with no salts. This shows that the kosmotropic salt stabilized the protein structure as a dimer, its catalytic active form, with no need to use high concentrations of protein in the assays.

The variants of concern (VOCs) showed little impact on recognizing peptide substrates/inhibitors, which was validated by testing the inhibitor Neq1183 and other peptide mimetic inhibitors against the WT and VOCs M^{pro}. The results suggest that the mutations did not alter the catalytic site environment enough to confer resistance to inhibition. This lack of alteration in the protein's ability to recognize small molecules suggests that inhibitors initially designed for the WT M^{pro} will likely remain effective against all the variants. Consequently, given the infrequent occurrence of mutations in its sequence and the minor impact demonstrated here on the protease's interaction with peptide-based molecules, M^{pro} will likely maintain its importance as a target of therapeutic interest.

Neq1183 was further analyzed using isothermal titration calorimetry against the WT M^{pro}. The outcomes showed a thermodynamic signature with all the thermodynamic parameters contributing positively to the interaction with the enzyme. Neq1184 was also studied with ITC, but it did not perform as well as Neq1183, having a detrimental entropic contribution, resulting in a less stable complex with the enzyme.

The MMP/SAR analysis did not identify atoms/groups at the P3 position in which the K_i had a significant change in affinity for M^{pro}. Despite the lack of the preferable groups used here, the compounds also inhibited the human cathepsin L, an essential protease involved in the SARS-CoV-2 infection. However, when dealing with the P2 position, bulky and aromatic groups showed a greater influence for affinity than compounds Neq1223 and Neq1224, although not significant. The use of different warheads demonstrated that the nitriles are, in fact, an excellent strategy for developing inhibitors for M^{pro}, as seen for Neq1183 (Nirmatrelvir). When substituting the oxime warhead for a nitrile, activity cliffs ($\Delta pK_i = +2.0$) were observed to favor the nitriles. Hence, using irreversible warheads, such as the oxime

carbamate and vinyl sulfones, presented significant affinity for the enzyme and can be further exploited to develop activity-based probes.

Additionally, the expression of M^{pro} in a mammalian HEK 293 T cell line was investigated, leading to an inactive expression of the protease. The results suggest the importance of small affinity tags at the protein's C-terminal, which will not interfere with dimerization.

Furthermore, three human cathepsins were investigated, aiming selectivity and high affinity for the human cathepsin L. Thus, a series of compounds were tested against human cathepsins L, S, and B, in which high affinity and selectivity (difference in pKi of 2 .0 log units) were observed for hCatL towards cathepsins S and B. An essential structural fragment was identified at position P2, the indole group that interacts with high affinity with the S2 subsite of the hCatL.

6 REFERENCES

[1]VERMA, S.; DIXIT, R.; PANDEY, K. C. Cysteine proteases: modes of activation and future prospects as pharmacological targets. **Frontiers in Pharmacology,** Lausanne, v. 7, 2016.

[2]SIKLOS, M.; BENAISSA, M.; THATCHER, G. R. J. Cysteine proteases as therapeutic targets: does selectivity matter? A systematic review of calpain and cathepsin inhibitors. **Acta Pharmaceutica Sinica B,** Beijing, v. 5, n. 6, p. 506–519, 2015.

[3]DRAG, M.; SALVESEN, G. S. Emerging principles in protease-based drug discovery. **Nature Reviews Drug Discovery,** London, v. 9, n. 9, p. 690–701, 2010.

[4]DEU, E.; VERDOES, M.; BOGYO, M. New approaches for dissecting protease functions to improve probe development and drug discovery. **Nature Structural & Molecular Biology,** New York, v. 19, n. 1, p. 9–16, 2012.

[5]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 α-ketoamide inhibitors. **Science,** Washington, p. eabb3405, 2020.

[6]HASANBASIC, S.; JAHIC, A.; KARAHMET, E.; SEJRANIC, A.; PRNJAVORAC, B. The role of cysteine protease in Alzheimer Disease. **Materia Socio Medica**, Sarajevo, v. 28, n. 3, p. 235, 2016.

[7]MCKERROW, J. H. The diverse roles of cysteine proteases in parasites and their suitability as drug targets. **PLOS Neglected Tropical Diseases,** San Francisco, v. 12, n. 8, p. e0005639, 2018.

[8]VIZOVIŠEK, M.; VIDMAR, R.; DRAG, M.; FONOVIĆ, M.; SALVESEN, G. S.; TURK, B. Protease specificity: towards in vivo imaging applications and biomarker discovery. **Trends in Biochemical Science,** Cambridge, v. 43, n. 10, p. 829–844, 2018.

[9]POWERS, J. C.; ASGIAN, J. L.; EKICI, Ö. D.; JAMES, K. E. Irreversible inhibitors of serine, cysteine, and threonine proteases. **Chemical Reviews,** Washington, v. 102, n. 12, p. 4639–4750, 2002.

[10]CIANNI, L.; SARTORI, G.; ROSINI, F.; DE VITA, D.; PIRES, G.; LOPES, B. R.; LEITÃO, A.; BURTOLOSO, A. C. B.; MONTANARI, C. A. Leveraging the cruzain S3 subsite to increase affinity for reversible covalent inhibitors. **Bioorganic Chemistry,** Maryland Heights, v. 79, p. 285–292, 2018.

[11]ANAND, K. Coronavirus main proteinase (3CLpro) structure: basis for design of Anti-SARS drugs. **Science,** Washington, v. 300, n. 5626, p. 1763–1767, 2003.

[12]JIN, Z. M.; DU, X.Y.; XU, Y., DENG, Y.; LIU, M.; ZHAO, Y. ; ZHANG, B.; LI, X.; ZANG, L.; PENG, C.; DUAN, Y.; YU, J.; WANG, L.; YANG, K. L.; LIU, F.; JIANG, R.; YANG, X.; VOCE, T.; LIU, X.; YANG, X.; FANGBAI; LIU, H.; GUDDAT, L. W.; XU, W.; XIAO, G. F.; QIN, C. F.; SHI, Z.; JIANG, H.; RAO, Z.;YANG, H.

Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. **Nature,** London, v. 582, n. 7811, p. 289–293, 2020.

[13]HILGENFELD, R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. **FEBS Journal,** West Sussex**,** v. 281, n. 18, p. 4085–4096, 2014.

[14]DAI, W.; ZHANG, B.; JUANG, X.M.; SU, H.; LI, J.; ZHAO, Y.; XIE, X.; JIN, Z. M.; PENG, J. J.; LIU, F.; LI, C.; LI, V.; FANGBAI; WANG, H. F.; CHENG, X.; CEN, X.; HU, S.; YANG, X.; WANG, J.; LIU, X.; XIAO, G. F.; JIANG, H.; RAO, Z.; ZHANG, L. K.; XU, Y.; YANG, H.; LIU, H. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. **Science,** Washington, p. eabb4489, 2020.

[15]NARWAL, M.; ARMACHE, J.-P.; EDWARDS, T. J.; MURAKAMI, K. S. SARS-CoV-2 polyprotein substrate regulates the stepwise Mpro cleavage reaction. **Journal of Biological Chemistry,** New York, v. 299, n. 5, p. 104697, 2023.

[16]WANG, F.; CHEN, C.; TAN, W.; YANG, K.; YANG, H. Structure of main protease from human coronavirus NL63: insights for wide spectrum anti-coronavirus drug design. **Scientific Reports,** London, v. 6, n. 1, 2016.

[17]MINKOFF, J. M.; TENOEVER, B. Innate immune evasion strategies of SARS-CoV-2. **Nature Reviews Microbiology,** London, 2023. doi:10.1038/s41579-022-00839-1.

[18]MACIP, G.; GARCIA-SEGURA, P.; MESTRES-TRUYOL, J.; SALDIVAR-ESPINOZA, B.; OJEDA-MONTES, M. J.; GIMENO, A.; CERETO-MASSAGUÉ, A.; GARCIA-VALLVÉ, S.; PUJADAS, G. Haste makes waste: a critical review of docking-based virtual screening in drug repurposing for SARS-CoV-2 main protease (M-pro) inhibition. **Medicinal Research Reviews,** New Jersey, v. 42, n. 2, p. 744–769, 2022.

[19]KNELLER, D. W.; PHILLIPS, G.; O'NEILL, H. M.; JEDRZEJCZAK, R.; STOLS, L.; LANGAN, P.; JOACHIMIAK, A.; COATES, L.; KOVALEVSKY, A. Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. **Nature Communications**, London, v. 11, n. 1, p. 3202, 2020.

[20]MILTNER, N.; KALLÓ, G.; CSŐSZ, É.; MICZI, M.; NAGY, T.; MAHDI, M.; MÓTYÁN, J. A.; TŐZSÉR, J. Identification of SARS-CoV-2 Main Protease (Mpro) cleavage sites using two-dimensional electrophoresis and in silico cleavage site prediction. **International Journal of Molecular Science,** Basel, v. 24, n. 4, p. 3236, 2023.

[21]BREWITZ, L.; DUMJAHN, L.; ZHAO, Y.; OWEN, C. D.; LAIDLAW, S. M.; MALLA, T. R.; NGUYEN, D.; LUKACIK, P.; SALAH, E.; CRAWSHAW, A. D.; WARREN, A. J.; TRINCAO, J.; STRAIN-DAMERELL, C.; CARROLL, M. W.; WALSH, M. A.; SCHOFIELD, C. J. Alkyne derivatives of SARS-CoV-2 main protease inhibitors including nirmatrelvir inhibit by reacting covalently with the nucleophilic cysteine. **Journal of Medicinal Chemistry,** Washington, v. 66, n. 4, p. 2663–2680, 2023.

[22]PREVITI, S.; ETTARI, R.; CALCATERRA, E.; DI MARO, S.; HAMMERSCHMIDT, S. J.; MÜLLER, C.; ZIEBUHR, J.; SCHIRMEISTER, T.; COSCONATI, S.; ZAPPALÀ, M. Structure-based lead optimization of peptide-based vinyl methyl ketones as SARS-CoV-2 main protease inhibitors. **European Journal of Medicinal Chemistry,** Issy les Moulineaux, v. 247, p. 115021, 2023.

[23]OWEN, D. R.; ALLERTON, C. M. N.; ANDERSON, A. S. S.; ASCHENBRENNER, L.; AVERY, M.; BERRITT, S.; BORAS, B.; CARDIN, R. D.; CARLOS, A.; COFFMAN, K. J.; DANTONIO, A.; DI, L.; ENG, H.; FERRE, R.; GAJIWALA, K. S.; GIBSON, S. B.; GREASLEY, S. E. HURST, B. L.; KADAR, E. P.; KALGUTKAR, A. S.; LEE, J. C.; LEE, J.; LIU, W.; MASON, S. W.; NOEL, S.; NOVAK, J. J.; OBACH, R. S.; OGILVIE. K.; PATEL, N. C.; PETTERSSON, M.; RAI, D. K.; REESE, M. R.; SAMMONS, M. F.; SATHISH, J. G.; SINGH, R. S. P.; STEPPAN, C. M.; STEWART, A. L.; TUTTLE, J. B.; UPDYKE, L.; VERHOEST, P. R.; WEI, L. Q.; YANG, Q.; ZHU, Y. An oral SARS-CoV-2 M [pro] inhibitor clinical candidate for the treatment of COVID-19. **Science,** Washington, p. eabl4784, 2021. doi:10.1126/Science, Washington.abl4784.

[24]ÁBRÁNYI-BALOGH, P.; KESERŰ, G. M. Warheads for designing covalent inhibitors and chemical probes. *In:* ADVANCES IN CHEMICAL PROTEOMICS. Amsterdam: Elsevier, 2022. p. 47–73.

[25]BONATTO, V.; LAMEIRO, R. F.; ROCHO, F. R.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A. Nitriles: an attractive approach to the development of covalent inhibitors. **RSC Medicinal Chemistry,** Cambridge, v. 14, n. 2, p. 201–217, 2023.

[26] SHINDO, N.; OJIDA, A. Recent progress in covalent warheads for in vivo targeting of endogenous proteins. **Bioorganic & Medicinal Chemistry,** Oxford, v. 47, p. 116386, 2021.

[27]JACOBS, J. L.; HAIDAR, G.; MELLORS, J. W. COVID-19: challenges of viral variants. **Annual Review of Medicine,** Palo Alto, v. 74, n. 1, p. 31–53, 2023.

[28]TURK, V. New embo members' review: lysosomal cysteine proteases: facts and opportunities. **The EMBO Journal,** Heidelberg, v. 20, n. 17, p. 4629–4633, 2001.

[29]OLSON, O. C.; JOYCE, J. A. Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. **Nature Reviews Cancer**, London, v. 15, n. 12, p. 712–729, 2015.

[30]KRAMER, L.; TURK, D.; TURK, B. The future of cysteine cathepsins in disease management. **Trends in Pharmacological Sciences**, Cambridge, v. 38, n. 10, p. 873–898, 2017.

[31]TURK, B. Targeting proteases: successes, failures and future prospects. **Nature Reviews Drug Discovery,** London, v. 5, n. 9, p. 785–799, 2006.

[32]DENNEMÄRKER, J.; LOHMÜLLER, T.; MAYERLE, J.; TACKE, M.; LERCH, M. M.; COUSSENS, L. M.; PETERS, C.; REINHECKEL, T. Deficiency for the cysteine protease cathepsin L promotes tumor progression in mouse epidermis. **Oncogene**, London, v. 29, n. 11, p. 1611–1621, 2010.

[33]DANA, D.; GARCIA, J.; BHUIYAN, A. I.; RATHOD, P.; JOO, L.; NOVOA, D. A.; PAROLY, S.; FATH, K. R.; CHANG, E. J.; PATHAK, S. K. Cell penetrable, clickable and tagless activity-based probe of human Cathepsin L. **Bioorganic Chemistry,** Maryland Heights, v. 85, p. 505–514, 2019.

[34]POREBA, M.; RUT, W.; VIZOVISEK, M.; GROBORZ, K.; KASPERKIEWICZ, P.; FINLAY, D.; VUORI, K.; TURK, D.; TURK, B.; SALVESEN, G. S.; DRAG, M. Selective imaging of cathepsin L in breast cancer by fluorescent activity-based probes. **Chemical Science,** Washington, v. 9, n. 8, p. 2113–2129, 2018.

[35]NISHIMURA, Y.; KAWABATA, T.; FURUNO, K.; KATO, K. Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes. **Archives of Biochemistry and Biophysics,** Maryland Heights, v. 271, n. 2, p. 400–406, 1989.

[36]DANA, D.; PATHAK, S. K. A review of small molecule inhibitors and functional probes of Human Cathepsin L. **Molecules,** Basel, v. 25, n. 3, p. 698, 2020.

[37]CIANNI, L.; FELDMANN, C. W.; GILBERG, E.; GÜTSCHOW, M.; JULIANO, L.; LEITÃO, A.; BAJORATH, J.; MONTANARI, C. A. Can cysteine protease cross-class inhibitors achieve selectivity? **Journal of Medicinal Chemistry,** Washington, 2019. doi:10.1021/acs.jmedchem.9b00683.

[38]SUDHAN, D. R.; SIEMANN, D. W. Cathepsin L targeting in cancer treatment. **Pharmacology & Therapeutics,** Philadelphia, v. 155, p. 105–116, 2015.

[39]LV, B.-J.; LINDHOLT, J. S.; WANG, J.; CHENG, X.; SHI, G.-P. Plasma levels of cathepsins L, K, and V and risks of abdominal aortic aneurysms: a randomized population-based study. **Atherosclerosis,** Shannon, v. 230, n. 1, p. 100–105, 2013.

[40]SUDHAN, D. R.; RABAGLINO, M. B.; WOOD, C. E.; SIEMANN, D. W. Cathepsin L in tumor angiogenesis and its therapeutic intervention by the small molecule inhibitor KGP94. **Clinical & Experimental Metastasis,** Dordrecht, v. 33, n. 5, p. 461–473, 2016.

[41]GOMES, C. P.; FERNANDES, D. E.; CASIMIRO, F.; DA MATA, G. F.; PASSOS, M. T.; VARELA, P.; MASTROIANNI-KIRSZTAJN, G.; PESQUERO, J. B. Cathepsin L in COVID-19: from pharmacological evidences to genetics. **Frontiers in Cellular and Infection Microbiology,** Lausanne, v. 10, p. 589505, 2020.

[42]QUILLES JUNIOR, J. C.; BERNARDI, M. D. L.; BATISTA, P. H. J.; SILVA, S. C. M.; ROCHA, C. M. R.; MONTANARI, C. A.; LEITÃO, A. Biological activity and physicochemical properties of dipeptidyl nitrile derivatives against pancreatic ductal adenocarcinoma cells. **Anti-Cancer Agents in Medicinal Chemistry,** Sharjah, v. 19, n. 1, p. 112–120, 2019.

[43]TABISH, T.; PRANJOL, M.; HORSELL, D.; RAHAT, A.; WHATMORE, J.; WINYARD, P.; ZHANG, S. Graphene oxide-based targeting of extracellular Cathepsin D and Cathepsin L as a novel anti-metastatic enzyme cancer therapy. **Cancers**, Basel, v. 11, n. 3, p. 319, 2019.

[44]PIŠLAR, A.; MITROVIĆ, A.; SABOTIČ, J.; PEČAR FONOVIĆ, U.; PERIŠIĆ NANUT, M.; JAKOŠ, T.; SENJOR, E.; KOS, J. The role of cysteine peptidases in coronavirus cell entry and replication: the therapeutic potential of cathepsin inhibitors. **PLOS Pathogens,** San Francisco, v. 16, n. 11, p. e1009013, 2020.

[45]SIMMONS, G.; GOSALIA, D. N.; RENNEKAMP, A. J.; REEVES, J. D.; DIAMOND, S. L.; BATES, P. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. **Proceedings of the National Academy of Sciences,** Oxford, v. 102, n. 33, p. 11876–11881, 2005.

[46]TANG, T.; BIDON, M.; JAIMES, J. A.; WHITTAKER, G. R.; DANIEL, S. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. **Antiviral Research,** Amsterdam, v. 178, p. 104792, 2020.

[47]OU, X.; LIU, Y.; LEI, X.; LI, P.; MI, D.; REN, L.; GUO, L.; GUO, R.; CHEN, T.; HU, J.; XIANG, Z.; MU, Z.; CHEN, X.; CHEN, J.; HU, K.; JIN, Q.; WANG, J.; QIAN, Z. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. **Nature Communications**, Oxford, v. 11, n. 1, p. 1620, 2020.

[48]ASHHURST, A. S.; TANG, A. H.; FAJTOVÁ, P.; YOON, M. C.; AGGARWAL, A.; CAMA, M. J. R.; STOYE, A.; BERETTA, L.; PWEE, D.; RELICH, A.; SKINNER, D.; LI, L.; MEEK, T. D.; MCKERROW, J. H.; GANCHO, V.; TSENG, C.T.; LARANCE, M.; TURVILLE, S.; GERWICK, W. M.; O'DONOGHUE, A. J.; PAYNE, R. J. Potent Anti-SARS-CoV-2 activity by the natural product Gallinamide A and analogues via inhibition of Cathepsin L. **Journal of Medicinal Chemistry,** Washington, v. 65, n. 4, p. 2956–2970, 2022.

[49]ZHAO, M.-M.; YANG, W.-L.; YANG, F.-Y.; ZHANG, L.; HUANG, W.-J.; HOU, W.; FAN, C.-F.; JIN, R.-H.; FENG, Y.-M.; WANG, Y.-C.; YANG, J.-K. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. **Signal Transduction and Targeted Therapy,** Chengdu, v. 6, n. 1, p. 134, 2021.

[50]MELLOTT, D.M.; TSENG, C.T.; DRELICH, A.; FAJTOVÁ, P.; CHENNA, B. C.; KOSTOMIRIS, D. H.; HSU, J.; ZHU, J.; TAYLOR, Z. W.; KOCUREK, K. I.; TATU, V.; KATZFUSS, A.; LI, L.; GIARDINI, M. A.; SKINNER, D.; HIRATA, K.; YOON, M. C.; BECK, S.; CARLIN, A. F.; CLARK, A. E.; BERETTA, L.; MANEVAL, D.; GANCHO, V.; FRUEH, F.; HURST, B. L.; WANG, H.; RAUSHEL, F. M.; O'DONOGHUE, A. J.; SIQUEIRA-NETO, J. L.; MEEK, T. D.; MCKERROW, J. H. A clinical-stage cysteine protease inhibitor blocks SARS-CoV-2 infection of human and monkey cells. **ACS Chemical Biology,** Washington, p. acschembio.0c00875, 2021. doi:10.1021/acschembio.0c00875.

[51]DHEILLY, E.; BATTISTELLO. E.; KATANAYEVA, N.; SUNGALEE, S.; MICHAUX, J.; DUNS, G.; WEHRLE, S.; SORDET-DESSIMOZ, J.; MINA, M.; RACLE, J.; FARINHA, P.; COUKOS, G.; GFELLER, D.; MOTTOK, A.; KRIDEL, R.; CORREIA, B. E.; STEIDL, C.; BASSAI, STERNBERG, M.; CIRIELLO, G.; ZOETE, V.; ORICCHIO, E. Cathepsin S regulates antigen processing and T cell

activity in Non-Hodgkin Lymphoma. **Cancer Cell,** Cambridge, v. 37, n. 5, p. 674-689.e12, 2020.

[52]ROCHO, F. D. R.; CIANNI, L.; BONATTO, V.; MARTINS, F. C. P.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A.; SHAMIM, A. Design, synthesis and stepwise optimization of nitrile-based inhibitors of cathepsins B and L. **Bioorganic & Medicinal Chemistry,** Oxford, v. 29, p. 115827, 2021.

[53]CIANNI, L.; FELDMANN, C. W.; GILBERG, E.; GÜTSCHOW, M.; JULIANO, L.; LEITÃO, A.; BAJORATH, J.; MONTANARI, C. A. Can cysteine protease cross-class inhibitors achieve selectivity? **Journal of Medicinal Chemistry,** Washington, v. 62, n. 23, p. 10497–10525, 2019.

[54]BIAN, B.; MONGRAIN, S.; CAGNOL, S.; LANGLOIS, M.-J.; BOULANGER, J.; BERNATCHEZ, G.; CARRIER, J. C.; BOUDREAU, F.; RIVARD, N. Cathepsin B promotes colorectal tumorigenesis, cell invasion, and metastasis. **Molecular Carcinogenesis,** New Jersey, v. 55, n. 5, p. 671–687, 2016.

[55]SCHMITZ, J.; LI, T.; BARTZ, U.; GÜTSCHOW, M. Cathepsin B inhibitors: combining dipeptide nitriles with an occluding loop recognition element by click chemistry. **ACS Medicinal Chemistry Letters,** Washington, v. 7, n. 3, p. 211–216, 2016.

[56]SCHMITZ, J.; GILBERG, E.; LÖSER, R.; BAJORATH, J.; BARTZ, U.; GÜTSCHOW, M. Cathepsin B: active site mapping with peptidic substrates and inhibitors. **Bioorganic & Medicinal Chemistry,** Oxford, v. 27, n. 1, p. 1–15, 2019.

[57]GUPTA, S.; SINGH, R. K.; DASTIDAR, S.; RAY, A. Cysteine cathepsin S as an immunomodulatory target: present and future trends. **Expert Opinion on Therapeutic Targets,** Oxfordshire, v. 12, n. 3, p. 291–299, 2008.

[58]MCGRATH, M. E.; PALMER, J. T.; BRÖMME, D.; SOMOZA, J. R. Crystal structure of human cathepsin S. **Protein Science,** Washington, v. 7, n. 6, p. 1294–1302, 1998.

[59]FOTI CUZZOLA, V.; PALELLA, E.; CELI, D.; BARRESI, M.; GIACOPPO, S.; BRAMANTI, P.; MARINO, S. Pharmacogenomic update on multiple sclerosis: a focus on actual and new therapeutic strategies. **The Pharmacogenomics Journal,** London**,** v. 12, n. 6, p. 453–461, 2012.

[60]AINSCOUGH, J. S.; MACLEOD, T.; MCGONAGLE, D.; BRAKEFIELD, R.; BARON, J. M.; ALASE, A.; WITTMANN, M.; STACEY, M. Cathepsin S is the major activator of the psoriasis-associated proinflammatory cytokine IL-36γ. **Proceedings of the National Academy of Sciences,** Oxford, v. 114, n. 13, p. E2748–E2757, 2017.

[61]CHEN, L.; LU, B.; YANG, Y.; ZHANG, W.; WANG, X.; ZHOU, H.; WEN, J.; YANG, Z.; HU, R. Elevated circulating cathepsin S levels are associated with metabolic syndrome in overweight and obese individuals. **Diabetes/Metabolism Research and Reviews,** Oxford, p. e3117, 2018. doi:10.1002/dmrr.3117.

[62]COPELAND, R. A. **Enzymes**: a practical introduction to structure, mechanism, and data analysis. New York: Wiley, 2000.

[63]NELSON, J. M.; VOSBURGH, W. C. Kinetics of invertase action. **Journal of the American Chemical Society,** Washington, v. 39, n. 4, p. 790–811, 1917.

[64]HENRI, V. Théorie générale de l'action de quelques diastases par Victor Henri [C. R. Acad. Sci. Paris 135 (1902) 916–919]. **Comptes Rendus Biologies**, Issy les Moulineaux, v. 329, n. 1, p. 47–50, 2006.

[65]VOET, D.; VOET, J. G. **Biochemistry**. New Jersey: John Wiley & Sons, 2011.

[66]COPELAND, R. A. **Evaluation of enzyme inhibitors in drug discovery**: a guide for medicial chemists and pharmacologists. New Jersey: John Wiley & Sons, 2015. v. 46.

[67]STRELOW, J. M. A Perspective on the kinetics of covalent and irreversible inhibition. **SLAS Discovery:** advancing life sciences R&D, Philadelphia, v. 22, n. 1, p. 3–20, 2017.

[68]VASHUM, Y.; PREMSINGH, R.; KOTTAISWAMY, A.; SOMA, M.; PADMANABAN, A.; KALAISELVAN, P.; SAMUEL, S. Inhibitory effect of cathepsin K inhibitor (ODN-MK-0822) on invasion, migration and adhesion of human breast cancer cells in vitro. **Molecular Biology Reports,** Dordrecht, v. 48, n. 1, p. 105–116, 2021.

[69]GEHRINGER, M.; LAUFER, S. A. Emerging and re-emerging warheads for targeted covalent inhibitors: applications in medicinal chemistry and chemical biology. **Journal of Medicinal Chemistry,** Washington, v. 62, n. 12, p. 5673–5724, 2019.

[70]ÁBRÁNYI-BALOGH, P.; PETRI, L.; IMRE, T.; SZIJJ, P.; SCARPINO, A.; HRAST, M.; MITROVIĆ, A.; FONOVIČ, U. P.; NÉMETH, K.; BARRETEAU, H.; ROPER, D. I.; HORVÁTI, K.; FERENCZY, G. G.; KOS, J.; ILAŠ, J.; GOBEC, S.; KESERŰ, G. M. A road map for prioritizing warheads for cysteine targeting covalent inhibitors. **European Journal of Medicinal Chemistry,** Issy les Moulineaux, v. 160, p. 94–107, 2018.

[71]SINGH, J.; PETTER, R. C.; BAILLIE, T. A.; WHITTY, A. The resurgence of covalent drugs. **Nature Reviews Drug Discovery,** London, v. 10, n. 4, p. 307–317, 2011.

[72]FARIDOON; NG, R.; ZHANG, G.; LI, J. J. An update on the discovery and development of reversible covalent inhibitors. **Medicinal Chemistry Research,** Cham, v. 32, n. 6, p. 1039–1062, 2023.

[73]DE CESCO, S.; KURIAN, J.; DUFRESNE, C.; MITTERMAIER, A. K.; MOITESSIER, N. Covalent inhibitors design and discovery. **European Journal of Medicinal Chemistry,** Issy les Moulineaux, v. 138, p. 96–114, 2017.

[74]SILVA, D. G.; RIBEIRO, J. F. R.; DE VITA, D.; CIANNI, L.; FRANCO, C. H.; FREITAS-JUNIOR, L. H.; MORAES, C. B.; ROCHA, J. R.; BURTOLOSO, A. C. B.; KENNY, P. W.; LEITÃO, A.; MONTANARI, C. A. A comparative study of warheads for design of cysteine protease inhibitors. **Bioorganic & Medicinal Chemistry Letters**, Oxford, v. 27, n. 22, p. 5031–5035, 2017.

[75]LINKUVIENĖ, V.; KRAINER, G.; CHEN, W.-Y.; MATULIS, D. Isothermal titration calorimetry for drug design: precision of the enthalpy and binding constant measurements and comparison of the instruments. **Analytical Biochemistry,** Philadelphia, v. 515, p. 61–64, 2016.

[76]CHIAD, K.; STELZIG, S. H.; GROPEANU, R.; WEIL, T.; KLAPPER, M.; MÜLLEN, K. Isothermal titration calorimetry: a powerful technique to quantify interactions in polymer hybrid systems. **Macromolecules,** Washington, v. 42, n. 19, p. 7545–7552, 2009.

[77]CLAVERIA-GIMENO, R.; VEGA, S.; ABIAN, O.; VELAZQUEZ-CAMPOY, A. A look at ligand binding thermodynamics in drug discovery. **Expert Opinion on Drug Discovery,** Oxfordshire, v. 12, n. 4, p. 363–377, 2017.

[78]GESCHWINDNER, S.; ULANDER, J.; JOHANSSON, P. Ligand binding thermodynamics in drug discovery: still a hot tip? **Journal of Medicinal Chemistry,** Washington, v. 58, n. 16, p. 6321–6335, 2015.

[79]TURNBULL, W. B.; DARANAS, A. H. On the value of c : can low affinity systems be studied by isothermal titration calorimetry? **Journal of the American Chemical Society,** Washington, v. 125, n. 48, p. 14859–14866, 2003.

[80]VELAZQUEZ-CAMPOY, A.; FREIRE, E. Isothermal titration calorimetry to determine association constants for high-affinity ligands. **Nature Protocols**, London, v. 1, n. 1, p. 186–191, 2006.

[81]FREIRE, E. Do enthalpy and entropy distinguish first in class from best in class? **Drug Discovery Today**, London, v. 13, n. 19–20, p. 869–874, 2008.

[82]CHODERA, J. D.; MOBLEY, D. L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. **Annual Review of Biophysics**, Palo Alto, v. 42, n. 1, p. 121–142, 2013.

[83]GUHA, R. On exploring structure–activity relationships. *In:* Kortagere, S., ed. **Silico models for drug discovery**. Totowa, Humana Press, 2013. v. 993, p. 81–94.

[84]WAWER, M.; BAJORATH, J. Local structural changes, global data views: graphical substructure−activity relationship trailing. **Journal of Medicinal Chemistry,** Washington, v. 54, n. 8, p. 2944–2951, 2011.

[85]BIRCH, A. M.; KENNY, P. W.; SIMPSON, I.; WHITTAMORE, P. R. O. Matched molecular pair analysis of activity and properties of glycogen phosphorylase inhibitors. **Bioorganic & Medicinal Chemistry Letters**, Oxford, v. 19, n. 3, p. 850–853, 2009.

[86]STUMPFE, D.; HU, H.; BAJORATH, J. Evolving concept of activity cliffs. **ACS Omega**, Washington, v. 4, n. 11, p. 14360–14368, 2019.

[87]STUMPFE, D.; BAJORATH, J. Exploring activity cliffs in medicinal chemistry: miniperspective. **Journal of Medicinal Chemistry,** Washington, v. 55, n. 7, p. 2932–2942, 2012.

[88]RUT, W.; GROBORZ, K.; ZHANG, L.; SUN, X.; ZMUDZINSKI, M.; PAWLIK, B.; WANG, X.; JOCHMANS, D.; NEYTS, J.; MŁYNARSKI, W.; HILGENFELD, R.; DRAG, M. SARS-CoV-2 Mpro inhibitors and activity-based probes for patient-sample imaging. **Nature Chemical Biology**, Oxford, v. 17, n. 2, p. 222–228, 2021.

[89]ROCHO, F. R.; SNIPAS, S. J.; SHAMIM, A; RUT, W.; DRAG, M.; MONTANARI, C. A.; SALVESEN, G. S. Differential specificity of SARS-CoV-2 main protease variants on peptide versus protein-based substrates. **The FEBS Journal**, West Sussex p. febs.16970, 2023. doi:10.1111/febs.16970.

[90]YUNG-CHI, C.; PRUSOFF, W. H. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. **Biochemical Pharmacology,** Philadelphia, v. 22, n. 23, p. 3099–3108, 1973.

[91]POP, C.; FITZGERALD, P.; GREEN, D. R.; SALVESEN, G. S. Role of proteolysis in Caspase-8 activation and stabilization. **Biochemistry,** Washington, v. 46, n. 14, p. 4398–4407, 2007.

[92]OKUR, H. I.; HLADÍLKOVÁ, J.; REMBERT, K. B.; CHO, Y.; HEYDA, J.; DZUBIELLA, J.; CREMER, P. S.; JUNGWIRTH, P. Beyond the hofmeister series: ion-specific effects on proteins and their biological functions. **The Journal of Physical Chemistry B,** Washington, v. 121, n. 9, p. 1997–2014, 2017.

[93]BOATRIGHT, K. M.; RENATUS, M.; SCOTT, F. L.; SPERANDIO, S.; SHIN, H.; PEDERSEN, I. M.; RICCI, J.-E.; EDRIS, W. A.; SUTHERLIN, D. P.; GREEN, D. R.; SALVESEN, G. S. A unified model for apical caspase activation. **Molecular Cell,** Cambridge, v. 11, n. 2, p. 529–541, 2003.

[94]SCHMIDT, U.; DARKE, P. L. Dimerization and activation of the herpes simplex virus type 1 protease. **Journal of Biological Chemistry,** New York, v. 272, n. 12, p. 7732–7735, 1997.

[95]OKAMOTO, D. N.; OLIVEIRA, L. C. G.; KONDO, M. Y.; CEZARI, M. H. S.; SZELTNER, Z.; JUHÁSZ, T.; JULIANO, M. A.; POLGÁR, L.; JULIANO, L.; GOUVEA, I. E. Increase of SARS-CoV 3CL peptidase activity due to macromolecular crowding effects in the milieu composition. **Biological Chemistry,** Berlin, v. 391, n. 12, 2010.

[96]GREASLEY, S. E.; NOELL, S.; PLOTNIKOVA, O.; FERRE, R.; LIU, W.; BOLANOS, B.; FENNELL, K.; NICKI, J.; CRAIG, T.; ZHU, Y.; STEWART, A. E.; STEPPAN, C. M. Structural basis for the in vitro efficacy of nirmatrelvir against SARS-CoV-2 variants. **Journal of Biological Chemistry,** New York, v. 298, n. 6, p. 101972, 2022.

[97]MA, C.; SACCO, M. D.; HURST, B.; TOWNSEND, J. A.; HU, Y.; SZETO, T.; ZHANG, X.; TARBET, B.; MARTY, M. T.; CHEN, Y.; WANG, J. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. **Cell Research**, London, v. 30, n. 8, p. 678–692, 2020.

[98]SACCO, M. D.; MA, C.; LAGARIAS, P.; GAO, A.; TOWNSEND, J. A.; MENG, X.; DUBE, P.; ZHANG, X.; HU, Y.; 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. **Science Advances**, Washington**,** v. 6, n. 50, p. eabe0751, 2020.

[99]BISSANTZ, C.; KUHN, B.; STAHL, M. A medicinal chemist's guide to molecular interactions. **Journal of Medicinal Chemistry,** Washington, v. 53, n. 14, p. 5061–5084, 2010.

[100]PIETRUŚ, W.; KURCZAB, R.; WARSZYCKI, D.; BOJARSKI, A. J.; BAJORATH, J. Isomeric activity cliffs—a case study for fluorine substitution of Aminergic G protein-coupled receptor ligands. **Molecules,** Basel, v. 28, n. 2, p. 490, 2023.

[101]PIETRUŚ, W.; KURCZAB, R.; STUMPFE, D.; BOJARSKI, A. J.; BAJORATH, J. Data-driven analysis of fluorination of ligands of Aminergic G protein coupled receptors. **Biomolecules**, Basel, v. 11, n. 11, p. 1647, 2021.

[102]BUDZIK, B.; GARZYA, V.; SHI, D.; WALKER, G.; LAUCHART, Y.; LUCAS, A. J.; RIVERO, R. A.; LANGMEAD, C. J.; WATSON, J.; WU, Z.; FORBES, I. T.; JIN, J. 2′ Biaryl amides as novel and subtype selective M1 agonists. Part II: further optimization and profiling. **Bioorganic & Medicinal Chemistry Letters**, Oxford, v. 20, n. 12, p. 3545–3549, 2010.

[103]LEMKE, C.; BENÝŠEK, J.; BRAJTENBACH, D.; BREUER, C.; JÍLKOVÁ, A.; HORN, M.; BUŠA, M.; ULRYCHOVÁ, L.; ILLIES, A.; KUBATZKY, K. F.; BARTZ, U.; MAREŠ, M.; GÜTSCHOW, M. An activity-based probe for Cathepsin K imaging with excellent potency and selectivity. **Journal of Medicinal Chemistry,** Washington, v. 64, n. 18, p. 13793–13806, 2021.

[104]NOSKE, G. D.; DE SOUZA SILVA, E.; DE GODOY, M. O.; DOLCI, I.; FERNANDES, R. S.; GUIDO, R. V. C.; SJÖ, P.; OLIVA, G.; GODOY, A. S. Structural basis of nirmatrelvir and ensitrelvir activity against naturally occurring polymorphisms of the SARS-CoV-2 main protease. **Journal of Biological Chemistry,** New York, v. 299, n. 3, p. 103004, 2023.

7   ATTACHMENT

2023 – Accepted Article

The FEBS Journal

Original Article

Differential specificity of SARS-CoV-2 main protease variants on peptide versus protein-based substrates

Fernanda R. Rocho ✉, Scott J. Snipas, Anwar Shamim, Wioletta Rut, Marcin Drag, Carlos A. Montanari, Guy S. Salvesen ✉

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi:10.1111/febs.16970.
Peer Review The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/febs.16970.

📄 PDF   🔧 TOOLS   ⬍ SHARE

Accepted Articles

Accepted, unedited articles published online and citable. The final edited and typeset version of record will appear in the future.

Metrics

Full text views: 16 ⓘ

Am score 2

⟳ Check for updates

Keywords

SARS-CoV-2 M^{pro} SARS-CoV-2 variants   Substrate specificity

Abstract

The SARS-CoV-2 main protease (M^{pro}) holds significant importance as a biological target in combating coronaviruses due to its importance in virus replication. Considering the emergence of novel SARS-CoV-2 variants and the mutations observed in the M^{pro} sequence, we hypothesized that these mutations may have a potential impact on the protease's specificity. To test this, we expressed M^{pro} corresponding to the original strain and variants Beta1, Beta2 and Omicron, and analyzed their activity on protein-based and peptide substrates. Although we observed differential activity on the protein-based substrate, there was very little difference when analyzed on the peptide substrate. We conclude that mutations on the M^{pro} sequence, despite having a minor effect on a peptide substrate cleavage, did not change the catalytic site environment enough to build resistance to inhibition. Therefore, we propose that inhibitors initially designed for the

**RSC
Medicinal Chemistry**

ROYAL SOCIETY
OF **CHEMISTRY**

REVIEW

Check for updates

Nitriles: an attractive approach to the development of covalent inhibitors

Vinícius Bonatto, [a] Rafael F. Lameiro,[a] Fernanda R. Rocho,[a]
Jerônimo Lameira, [ab] Andrei Leitão [a] and Carlos A. Montanari [*a]

Nitriles have broad applications in medicinal chemistry, with more than 60 small molecule drugs on the market containing the cyano functional group. In addition to the well-known noncovalent interactions that nitriles can perform with macromolecular targets, they are also known to improve drug candidates' pharmacokinetic profiles. Moreover, the cyano group can be used as an electrophilic warhead to covalently bind an inhibitor to a target of interest, forming a covalent adduct, a strategy that can present benefits over noncovalent inhibitors. This approach has gained much notoriety in recent years, mainly with diabetes and COVID-19-approved drugs. Nevertheless, the application of nitriles in covalent ligands is not restricted to it being the reactive center, as it can also be employed to convert irreversible inhibitors into reversible ones, a promising strategy for kinase inhibition and protein degradation. In this review, we introduce and discuss the roles of the cyano group in covalent inhibitors, how to tune its reactivity and the possibility of achieving selectivity only by replacing the warhead. Finally, we provide an overview of nitrile-based covalent compounds in approved drugs and inhibitors recently described in the literature.

Introduction

The nitrile (or cyano) group has several applications in diverse fields of chemistry, from superglues with methyl cyanoacrylate to drugs such as cimetidine. In the latter, nitriles have played a significant role, with over 70 approved drugs presenting this group in their chemical structure.[1] Within these approved drugs, 61 are small organic molecules, with 55 containing only one nitrile group in their scaffold and six presenting this group twice.

Drugs containing the cyano group are used to treat various diseases, ranging from viral infections to different types of cancer.[2] These drugs take advantage of the fact that nitriles can be an excellent group to improve the compounds' pharmacodynamic (PD) and pharmacokinetic (PK) profiles, as they can make different types of interactions with macromolecular targets and/or improve water solubility.[2–4]

The nitrile group (Fig. 1) has a linear geometry with a nitrogen atom bonded to an sp hybridized carbon atom through a triple bond. The C sp atom can act as an electrophile due to its electron deficiency, promoted by the high electronegativity of the nitrogen and high dipole moment in the triple bond. And by the lone pair of the

nitrogen atom, it is possible to interact with hydrogen bond donor groups.[3–5] Owing to its linear shape and low molecular volume, it can fit properly in the subsites of target proteins and perform lipophilic interactions *via* the triple bond's pi system.[3,4]

In addition to the nonbonded interactions, nitriles are a remarkable group that can form covalent adducts with proteins, mainly linked to a reactive cysteine or serine side chain.[6,7] The stability of the covalent bond between the ligand and the target can be modulated by different types of substituents in the vicinity of the nitrile, which also impacts its reactivity. The reactivity modulation is critical to the design of various types of inhibitors, from reversible to irreversible. In addition, the nitrile's electron-withdrawing property can be used in combination with other warheads, as is the case of cyano-acrylamides, in which the covalent bond is formed on the acrylamide β-carbon, while the cyano group increases the reactivity and promotes reversibility.[8,9] All these
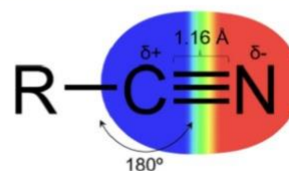
[a] Medicinal and Biological Chemistry Group, São Carlos Institute of Chemistry, University of São Paulo, Avenue Trabalhador Sancarlense, 400, 13566-590, São Carlos/SP, Brazil. E-mail: carlos.montanari@usp.br
[b] Institute of Biological Science, Federal University of Pará, Rua Augusto Correa S/N, Belém, PA, Brazil

Fig. 1 Representation of the nitrile (cyano) group.

REVIEW

Check for updates

A patent review on cathepsin K inhibitors to treat osteoporosis (2011 – 2021)

Fernanda R. Rocho [a], Vinícius Bonatto [a], Rafael F. Lameiro [a], Jerônimo Lameira [a,b], Andrei Leitão [a] and Carlos A. Montanari [a]

aMedicinal and Biological Chemistry Group, São Carlos Institute of Chemistry, University of São Paulo, São Carlos/SP, Brazil; bInstitute of Biological Science, Federal University of Pará, Belém, Brazil

ABSTRACT

Introduction: Cathepsin K (CatK) is a lysosomal cysteine protease and the predominant cathepsin expressed in osteoclasts, where it degrades the bone matrix. Hence, CatK is an attractive therapeutic target related to diseases characterized by bone resorption, like osteoporosis.

Areas covered: This review summarizes the patent literature from 2011 to 2021 on CatK inhibitors and their potential use as new treatments for osteoporosis. The inhibitors were classified by their warheads, with the most explored nitrile-based inhibitors. Promising in vivo results have also been disclosed.

Expert opinion: As one of the most potent lysosomal proteins whose primary function is to mediate bone resorption, cathepsin K remains an excellent target for therapeutic intervention. Nevertheless, there is no record of any approved drug that targets CatK. The most notable cases of drug candidates targeting CatK were balicatib and odanacatib, which reached Phase II and III clinical trials, respectively, but did not enter the market. Further developments include exploring new chemical entities beyond the nitrile-based chemical space, with improved ADME and safety profiles. In addition, CatK's role in cancer immunoexpression and its involvement in the pathophysiology of osteo- and rheumatoid arthritis have raised the race to develop activity-based probes with excellent potency and selectivity.

1. Introduction

Cathepsin K (CatK) is a lysosomal cysteine protease that belongs to the human cathepsin family; an eleven-membered family categorized according to their structures and catalytic mechanisms. The cysteine cathepsin family possesses a standard papain-like structure with a conserved active site region with a Cys-His-Asn catalytic triad (Figure 1) [1]. Cathepsins exist in several living organisms, such as bacteria, viruses, plants, and animals. They are known to be involved in a variety of diseases, such as many types of cancer [2–4], autoimmune disorders [5], and bone-related diseases [6].

The active site of CatK consists of 4-four well-characterized subsites (S1', S1, S2, S3 Figure 1), which are explored to design new ligands. These inhibitors may or may not present selectivity since the subsites share a high degree of structural homology with other cathepsins [7]. The S1 subsite contains the catalytic triad (Cys25, His162, and Asn182) responsible for the mechanism of action of the enzyme. The S2 pocket of CatK is the smallest among the enzymes of the cysteine cathepsin family, mainly because of the Tyr67 and Leu209 residues. The S3 subsite is often used to achieve selectivity over other cathepsins and is shaped by the Asp61 residue [7].

Consequently, inhibitors are designed to interact with CatK subsites to achieve high potency and selectivity over other cathepsins. The use of peptidomimetic compounds is a common approach, in which a reactive group, usually called warhead, is commonly linked to it (Figure 1) [8,9]. The warhead is a highly electrophilic moiety that participates in a nucleophilic attack promoted by the thiolate from the Cys25 residue, forming a covalent bond between the inhibitor and the enzyme. The other moieties of the inhibitors labeled as P1, P2, and P3 will participate in noncovalent interactions with the respective residues present in the S1, S2, and S3 subsites.

Interestingly, CatK is the only cathepsin highly expressed in osteoclasts, where the enzyme is present in the lysosome and cytoplasmic vesicles [10]. CatK is widely secreted by activated osteoclasts to degrade the bonde matrix, primarily type I collagen protein, constituting approximately 90% of the organic bone matrix [6]. The enzyme can also degrade type II collagen, the main matrix protein in cartilage [11]. Research with murine models reinforces the critical role of CatK in bone resorption [12–14]. Studies showed that mice with CatK deficiency could develop osteopetrosis of the long bones, in which inefficient osteoclasts activity was observed [13]. Additionally, a recent study showed that osteocytes could also express and secrete CatK, required for lactation-induced peri-lacunar resorption, to assure the right amounts of calcium in milk and aid skeletal development in offspring [15]. Therefore, the enzyme has become an attractive and essential biological target for treating bone-related diseases, primarily osteoporosis [8,16,17], which will be discussed throughout this review. Furthermore, it is essential to mention that despite CatK's role in osteoporosis, its implication goes beyond as the enzyme is also expressed in other cell types [16], which makes the protein a promising target for many diseases, such

CONTACT Carlos A. Montanari ✉ Carlos.Montanari@usp.br 🔗 Medicinal & Biological Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, 13566-590 São Carlos, SP, Brazil

RESEARCH ARTICLE

The Effect of Dipeptidyl Nitrile Derivatives on Pancreatic Ductal Adenocarcinoma Cells *In Vitro*

Sabrina Mendes Botelho[1], Fernanda dos Reis Rocho[1], Lorenzo Cianni[1], Carlos Alberto Montanari[1] and Andrei Leitão[1,*]

[1]*Medicinal & Biological Chemistry Group (NEQUIMED), São Carlos Institute of Chemistry (IQSC), University of São Paulo (USP) - Av. Trabalhador São-carlense, 400, São Carlos, SP, Brazil*

Abstract: *Aims:* This study aims to evaluate the bioactivity of dipeptidyl nitrile inhibitors of human cysteine cathepsins that could work as anticancer agents in a drug discovery and development project.

Background: Human lysosomal cysteine proteases promote cancer progression, migration, and metastasis, targeted by inhibitors.

Objective: Here, 19 cysteine protease inhibitors known as dipeptidyl nitriles were tested using MIA PaCa-2 pancreatic cancer cells and Balb/3T3 clone A31 non-tumoral mouse fibroblasts.

Methods: *In vitro* assays evaluated cell migration, colony formation, inhibition of the enzymatic activity in cell lysates, and combination therapy with gemcitabine.

Results: There were mixed results; the inhibitors reduced the number of colonies but did not affect the total area. Cells migrated despite enzyme inhibition by Neq0709 and Neq0712. As expected, the compounds were non-cytotoxic; they improved the potency of gemcitabine in the combined therapy assay, especially for Neq0707.

Conclusion: In summary, our findings revealed the complexity of dealing with the translation from biochemical to cell-based assays in the hit-to-lead step.

Keywords: Cell-based assays, SAR, cysteine protease inhibition, antineoplastic activity, colony formation assay, scratch healing assay, combination therapy.

1. INTRODUCTION

Pancreatic cancer has a high mortality rate because it is usually asymptomatic until advanced. For this reason, mean survival is only 35 months without treatment [1]. Surgery is usually an option; however, chemotherapy with gemcitabine is often provided with limited success [2]. Hence, there is an active search for novel targets and therapeutic approaches.

Cysteine proteases are targets for drug discovery owing to their substantial involvement in physiopathological processes [3]. In particular, cysteine cathepsins trigger angiogenesis and metastasis and guide the fate of cancer cells [4]. These enzymes are targets for anticancer drug discovery efforts [5], leading to the development of several types of inhibitors [6, 7]. Among them, our research group studied dipeptidyl nitrile derivatives to design highly potent inhibitors [8]. Dip-

eptidyl nitriles and derivatives are potent reversible covalent cathepsin inhibitors, and the selectivity may be exploited *via* subsite recognition (S3, S2, S1) by the substituent available in the corresponding position (P3, P2, P1) [9].

Based on our earlier results pointing to the cytostatic activity of dipeptidyl nitriles against the pancreatic cancer cell line MIA PaCa-2 [10], we increased the set of compounds by increasing the chemical diversity at positions P1, P2, and P3 (Fig. **1**) and using enantiomers in P_2 to evaluate the bioactivity of a broader set of chemicals.
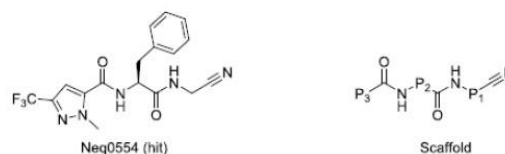
Fig. (1). Chemical structure of the dipeptidyl nitrile Neq0554 used as a hit for the scaffold-based analysis of derivatives.

* Address correspondence to this author at the Medicinal & Biological Chemistry Group (NEQUIMED), São Carlos Institute of Chemistry (IQSC), University of São Paulo (USP) - Av. Trabalhador São-carlense, 400, São Carlos, SP, Brazil; E-mail: andleitao@iqsc.usp.br

2021

Predicting the Relative Binding Affinity for Reversible Covalent Inhibitors by Free Energy Perturbation Calculations

Vinícius Bonatto, Anwar Shamim, Fernanda dos R. Rocho, Andrei Leitão, F. Javier Luque, Jerônimo Lameira,* and Carlos A. Montanari*

ACCESS | 📊 Metrics & More | 📖 Article Recommendations | 🗐 Supporting Information

ABSTRACT: Covalent inhibitors are assuming central importance in drug discovery projects, especially in this pandemic scenario. Many research groups have focused their attention on inhibiting viral proteases or human proteases such as cathepsin L (hCatL). The inhibition of these critical enzymes may impair viral replication. However, molecular modeling of covalent ligands is challenging since covalent and noncovalent ligand-bound states must be considered in the binding process. In this work, we evaluated the suitability of free energy perturbation (FEP) calculations as a tool for predicting the binding affinity of reversible covalent inhibitors of hCatL. Our strategy relies on the relative free energy calculated for both covalent and noncovalent complexes and the free energy changes have been compared with experimental data for eight nitrile-based inhibitors, including three new inhibitors of hCatL. Our results demonstrate that the covalent complex can be employed to properly rank the inhibitors. Nevertheless, a comparison of the free energy changes in both noncovalent and covalent states is valuable to interpret the effect triggered by the formation of the covalent bond on the interactions played by functional groups distant from the warhead. Overall, FEP can be employed as a powerful predictor tool in developing and understanding the activity of reversible covalent inhibitors.
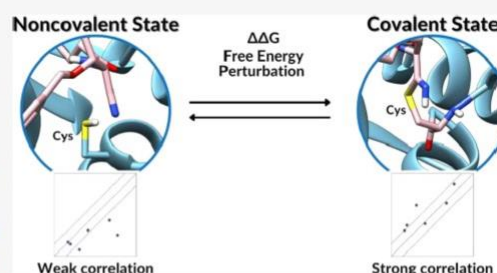
■ INTRODUCTION

Covalent inhibitors have received much attention over the past decade by discovering new drugs covalently bound to their targets, such as anticancer and antidiabetics agents.[1] Currently, the interest in this class of inhibitors has been revitalized due to the search for compounds targeting the new coronavirus main protease (Mpro) through covalent modification.[2–5] Noteworthy, Pfizer has started clinical trials for two distinct reversible covalent inhibitors to treat Covid-19, wherein the compounds inhibit this protease.[6,7]

The covalent inhibition mechanism involves the presence of a reactive group in the chemical scaffold of the inhibitor, suitably located to form a chemical bond with a specific residue upon the binding to the target protein.[8] The covalent bond formed in the inhibitor–enzyme complex can be reversible or irreversible. The appropriate choice of the warhead can control the reversibility of the covalent adduct.[9–11] Groups such as nitrile, aldehyde, and α-ketoamide are examples of warheads leading to the formation of a reversible covalent bond with their target.[12] In general, covalent inhibition is a two-step process, as depicted in Figure 1 for the nucleophilic attack of a cysteine residue to a nitrile-containing compound. The noncovalent complex formation enables the inhibitor's reactive electrophile to be located close to the enzyme's catalytic nucleophilic group in the binding pocket. The covalent
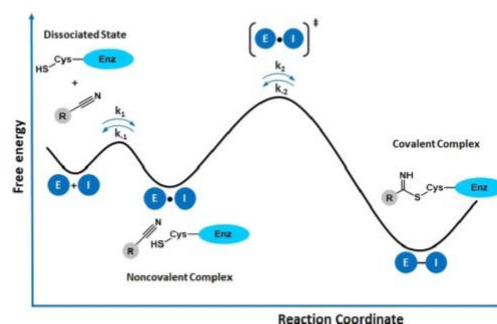

Figure 1. Schematic reaction involving a reversible covalent inhibitor (I) that incorporates a nitrile group as warhead and a reactive cysteine in the target enzyme (E) responsible for the nucleophilic attack. The rate constant k_{-2} determines the reversibility of the covalently bound adduct [E–I].

2021

Nitrile-based peptoids as cysteine protease inhibitors

Luana Alves, Deborah A. Santos[*], Rodrigo Cendron, Fernanda R. Rocho, Thiago K.B. Matos, Andrei Leitão[*], Carlos A. Montanari

Medicinal and Biological Chemistry Group (NEQUIMED), Institute of Chemistry of São Carlos, University of São Paulo, São Carlos/SP, Brazil

ARTICLE INFO

ABSTRACT

Peptidomimetics of the class of dipeptidyl nitrile analog peptoids were synthesized as inhibitors of mammalian cysteine proteases of the papain superfamily. The dipeptidyl nitrile side chains were attached to the peptide backbone's nitrogen atom, not to the α-carbons. Synthesized nitrile-based peptoid analogs that lack the hydrogen amide at P2-P3 are responsible for many of the secondary structure elements in peptides and proteins, making them resistant to proteolysis. The designed peptoids would lose a hydrogen bond with cruzain Asp161 decreasing the affinity toward the enzyme. A structure-activity relationship and matched molecular pair-based analysis between the dipeptidyl nitrile Neq0409 and its peptoid **4a** yielded the following cruzain affinities: $pK_i^{Neq0409} = 6.5$ and $pK_i^{4a} = 5.2$. respectively. A retrosynthetic matched molecular pair cliff (RMMP-cliff) analysis with a $\Delta pK_i^{Neq0409-4a}$ of 1.3 log is found for this transformation. These novel peptoids were then optimized, leading to compound **4i**, with high cruzain inhibition (pKi = 6.8). Cross-class cathepsin activity was observed for some of these novel compounds against cathepsins K, L and S, while other compounds presented a selective inhibition of cathepsin K (**4b**, **4c**, **4k**) over ten times higher than the other enzymes. The putative mode of binding was determined by using covalent docking, which also aided to describe the structure-activity relationship (SAR). Interestingly, none of the peptoids inhibited CatB to any appreciable extent. These results provide guidance to identify novel bioactive nitrile-based peptoids.

1. Introduction

Papain-like lysosomal cysteine cathepsins play crucial roles in normal cellular processes; however, overexpression and/or increased cathepsins activities can be associated with pathologies development.[1,2] Cysteine proteases (CP) are also identified in parasitic organisms, such as protozoa (*Trypanosoma, Leishmania,* and *Plasmodium*) and viruses.[3–7] Moreover, cysteine peptidases are validated targets for many diseases and the development of potent and selective inhibitors remains relevant.[8,9]

Non-peptide and peptidomimetics nitrile-based scaffolds have been recognized as CP inhibitors, in which the nitrile warhead promotes a covalent interaction with the CP catalytic site.[10,11] Dipeptidyl nitriles (DPN) have been recognized as cysteine proteases inhibitors. Many studies described the ligand's reversible-covalent interaction to the enzymes, where the selectivity can be improved by changing the peptide substituents by optimizing the interactions with subsites of each

target.[12–17] Azapeptide nitriles were described as promising warheads to CP ligands once the CαH replacement by nitrogen changed the kinetics to a time-dependent inhibition.[18,19] More recently, triazine nitriles were reported as cathepsin L and rhodesain potent inhibitors.[20–22]

The application of peptides as drugs has many advantages, such as recognizing the structures in biological targets and low toxicity; however, their low metabolic stability is the main disadvantage. Thus, amide isosteric replacement has attracted significant attention in drug discovery. Hence, peptidomimetics or peptides bioisosteres are frequently employed to modulate peptide properties, leading to the improvement of selectivity, stability, pharmacokinetics, among others.[23–25]

Peptoids are classified as *N*-substituted glycine oligomers, i.e., a peptide regioisomer, in which the α-carbon side chain of a peptide is moved to the amide nitrogen (Fig. 1). Consequently, they lose the asymmetric α-carbon and the hydrogen bond donor and increase their flexibility. Compared to the peptide counterpart, peptoids are often more stable to protease degradation and salt or pH variation; likewise,

8 APPENDIX

M^{pro} Constructs and Primers

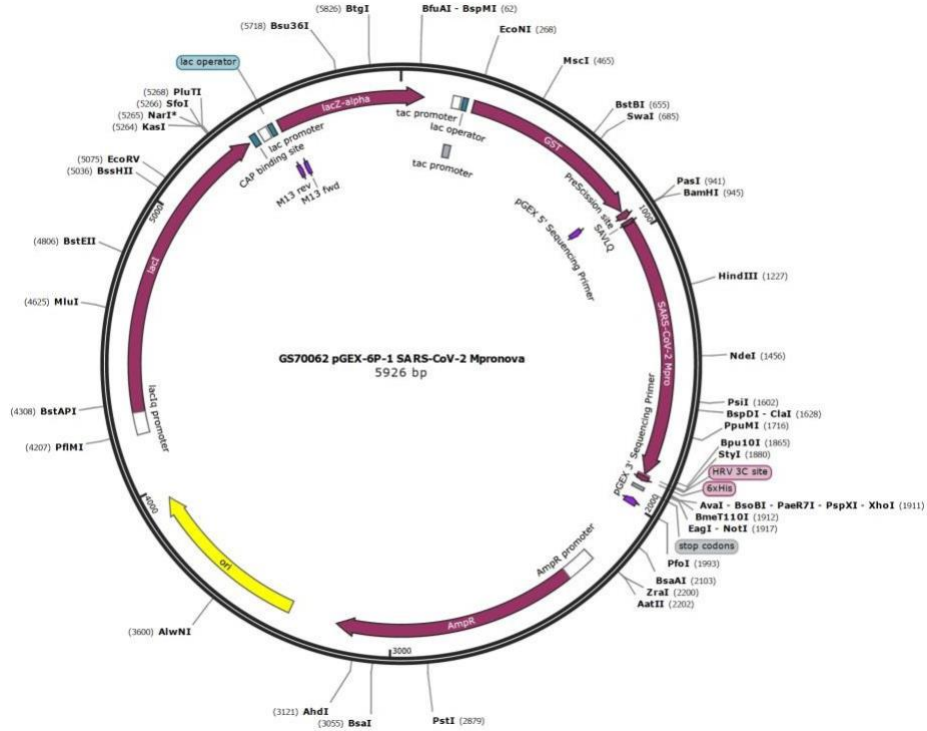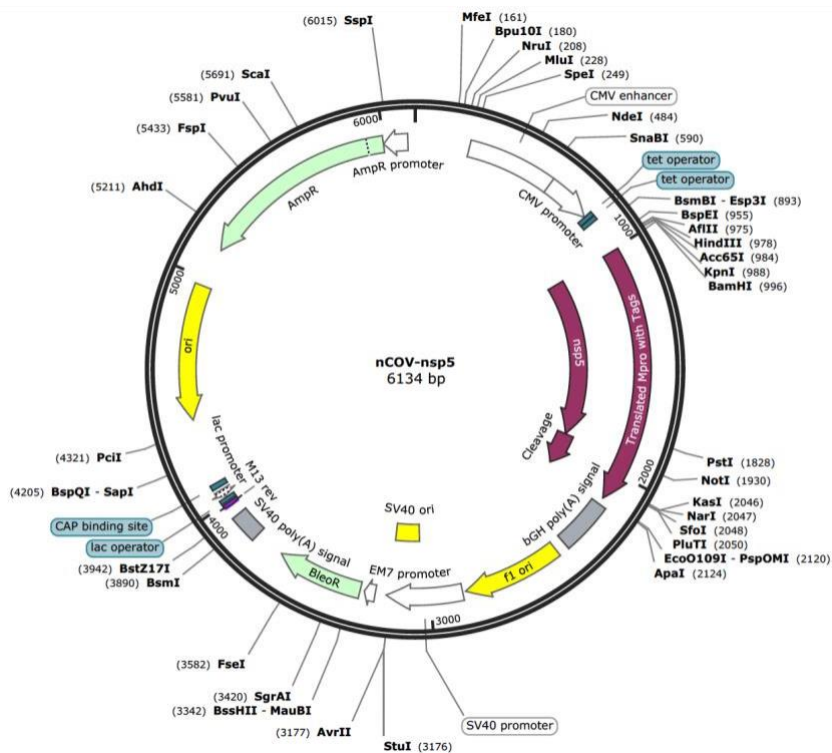**Figure S1:** WT M^{pro} plasmid construct. Generated using SnapGene software

Figure S2: WT M^{pro} mammalian construct for HEK transformation. Generated using SnapGene software

**Table S1:** Primers used for the M^{pro} PCR mutagenesis

SARS-CoV-2 M^{pro} Mutation	Forward (5´- 3´)	Reverse (5´- 3´)
C145A	Primer 1- CATTCCTTAATGGTTCAGCTGGTAGTGTTGG	Primer 2- CCAACACTACCAGCTGAACCATTAAGGAATG
K90R	Primer 3- GTGTACTTAAGCTTCGGGTTGATACAGCCAATCC	Primer 4- GGATTGGCTGTATCAACCCGAAGCTTAAGTACAC
A193V	Primer 5- GACAGGCAAACAGCACAAGTAGCTGGTACGGACAC	Primer 6- GTGTCCGTACCAGCTACTTGTGCTGTTTGCCTGTC
P132H	Primer 7- AATGTGCTATGAGGCACAATTTCACTATT	Primer 8- AATAGTGAAATTGTGCCTCATAGCACATT
Flanking Primers	Primer 9- GGGCTGGCAAGCCACGTTTGGTG	Primer 10- CCGGGAGCTGCATGTGTCAGAGG

SARS-CoV-2 M^{pro} Variants of Concern SDS-PAGE Purification

Figure S3: Protein gel of the purification step using Ni column for catalytic mutant Mpro-His (62.1 kDa). A gradient of imidazole was used starting at 7.8 mM (-) until 500 mM (+) (2-fold dilution).
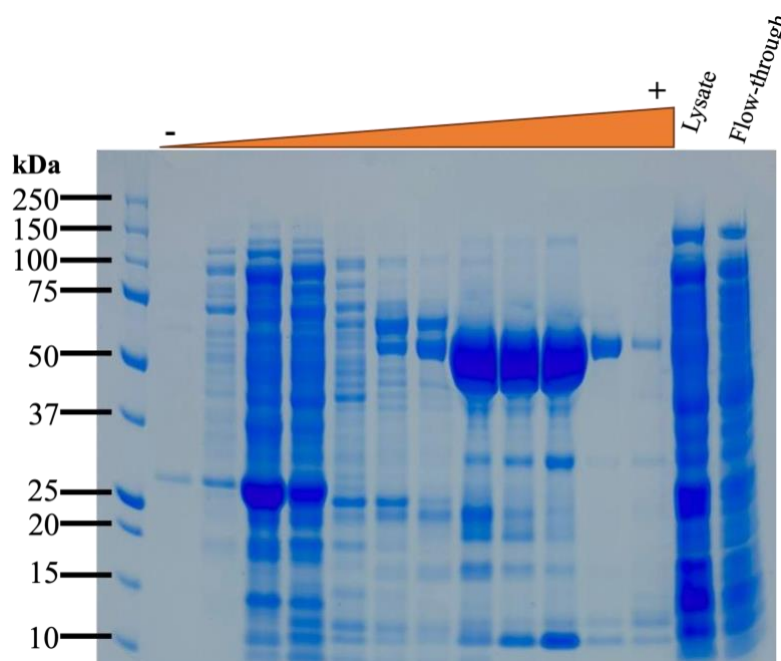
Figure S4: Protein gel of the purification step using Ni column for Beta variant Mpro-His (34.8 kDa). A gradient of imidazole was used starting at 7.8 mM (-) until 500 mM (+) (2-fold dilution).
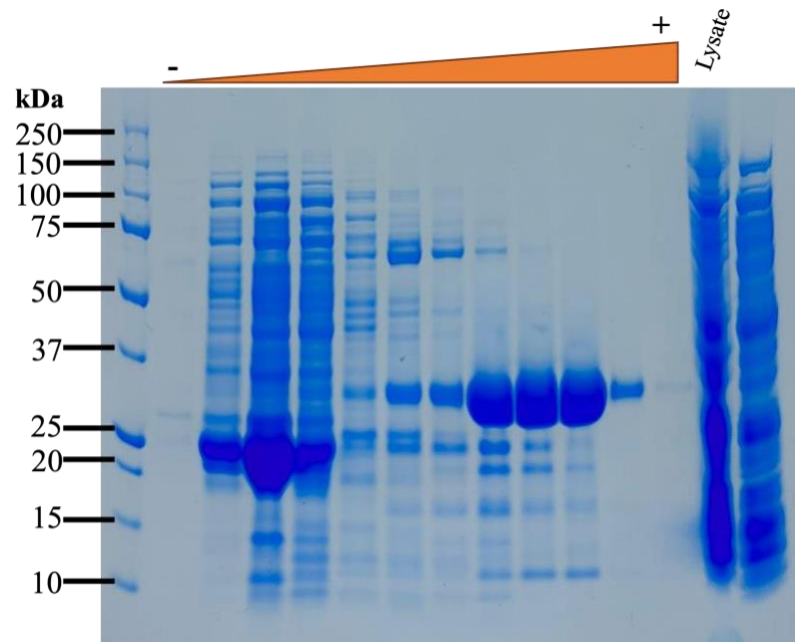

Figure S5: Protein gel of the purification step using Ni column for Mpro-His from the beta variant carrying to mutations (34.8 kDa). A gradient of imidazole was used starting at 7.8 mM (-) until 500 mM (+) (2-fold dilution). Samples were diluted 10x prior to addition to the gel.
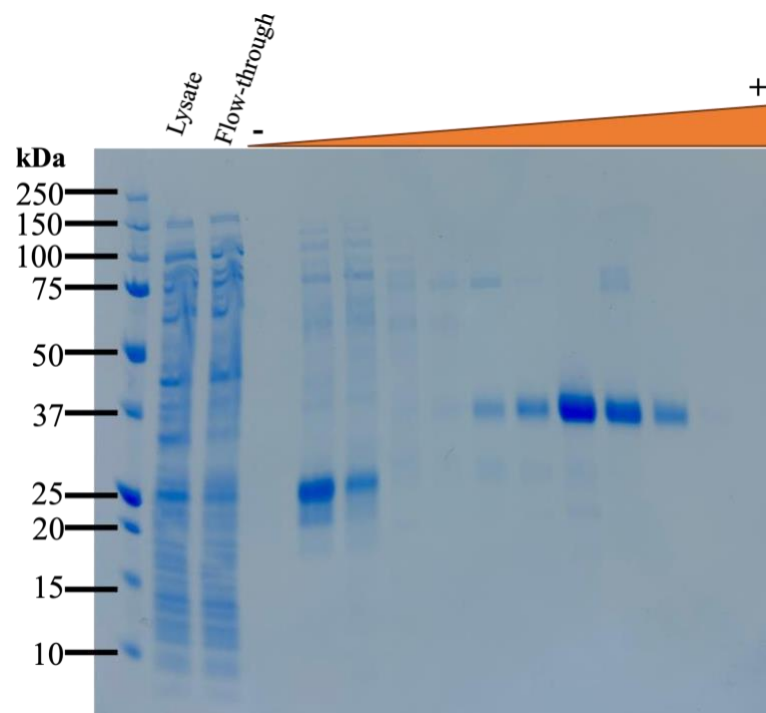
Figure S6: Protein gel of the purification step using Ni column for Mpro-His from the omicron variant (34.8 kDa). A gradient of imidazole was used starting at 7.8 mM (-) until 500 mM (+) (2-fold dilution).
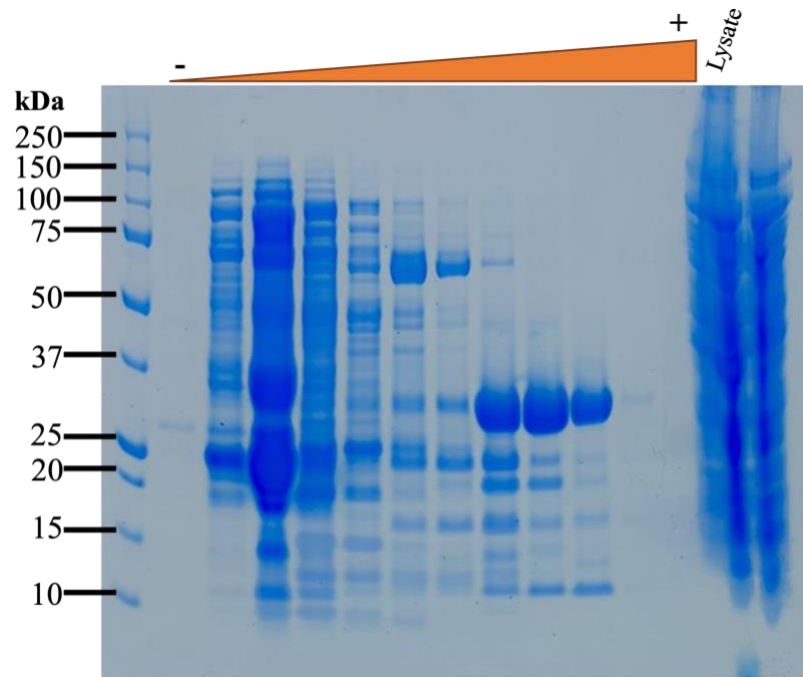
Figure S7: QS1 substrate structure with the fluorophore ACC highlighted in yellow, and the cleavage site is highlighted in red.
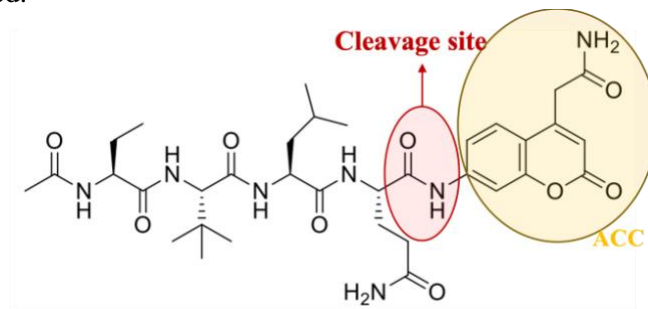
Kinetic Parameters for the WT SARS-CoV-2 M^pro and variants

Figure S7: Plot of reaction velocities against substrate concentration fitted to the Michaelis-Menten equation. (A) WT M^pro, (B) Beta1 M^pro, (C) Beta2 M^pro, and (D) Omicron M^pro.
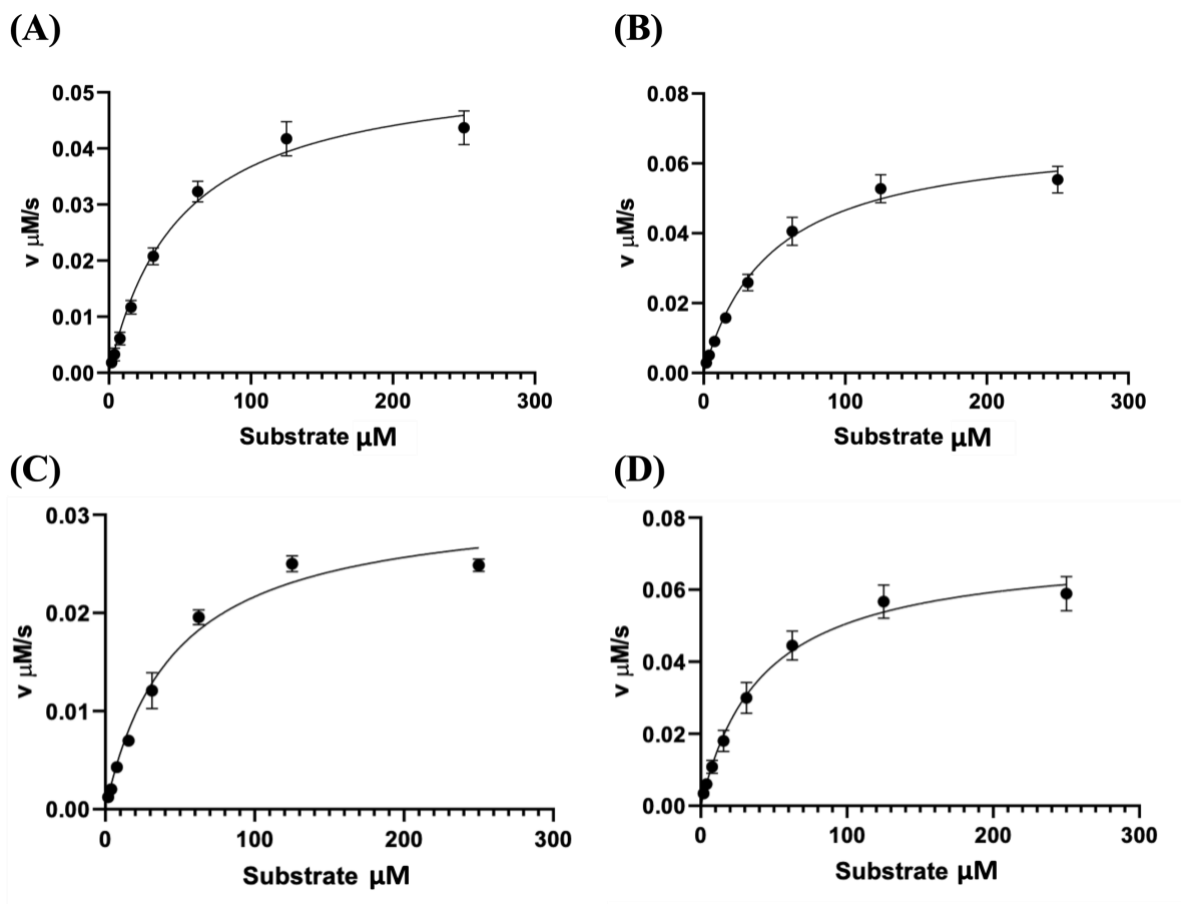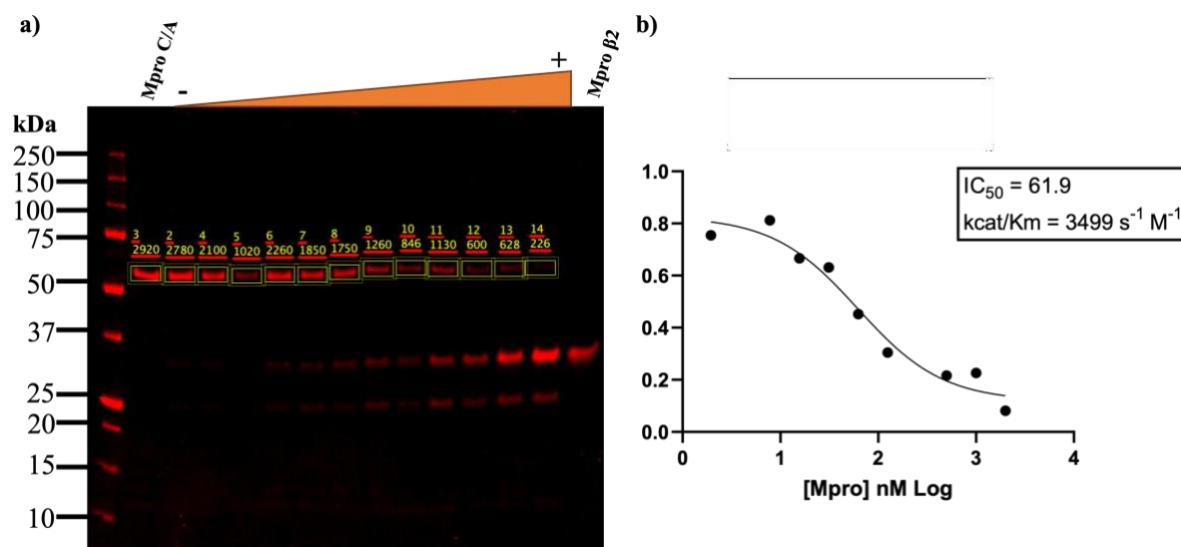

Protein-based Substrate Cleavage for WT M^pro and Variants

Figure S8: Protein substrate cleavage for the WT M^pro. a) The gel was scanned and analyzed using Image Studio software, and b) the catalytic efficiency was calculated using equation (1).
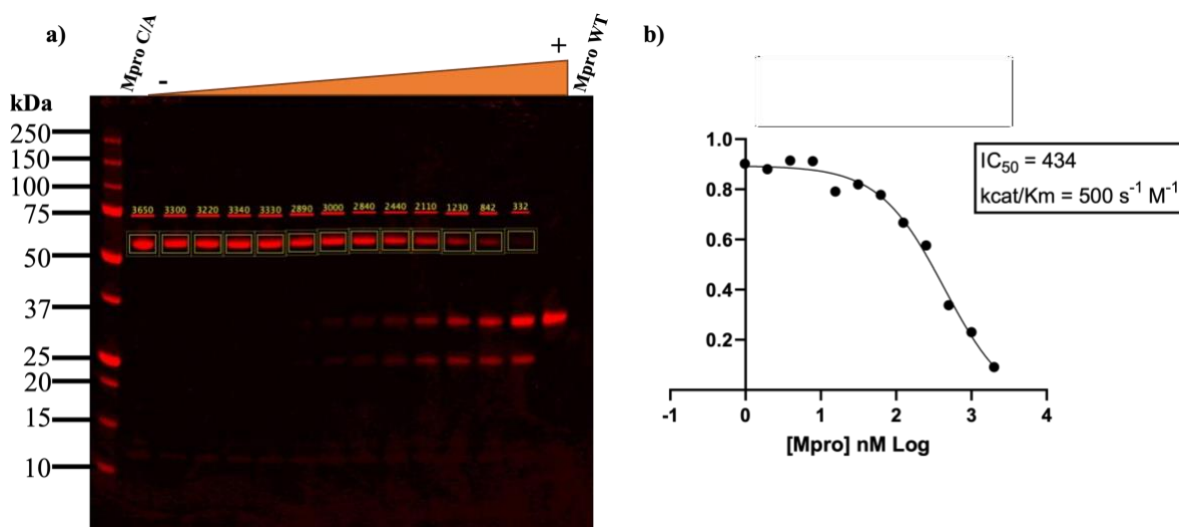
Figure S9: Protein substrate cleavage for the WT M^{pro} replicate. a) The gel was scanned and analyzed using Image Studio software, and b) the catalytic efficiency was calculated using equation (1).
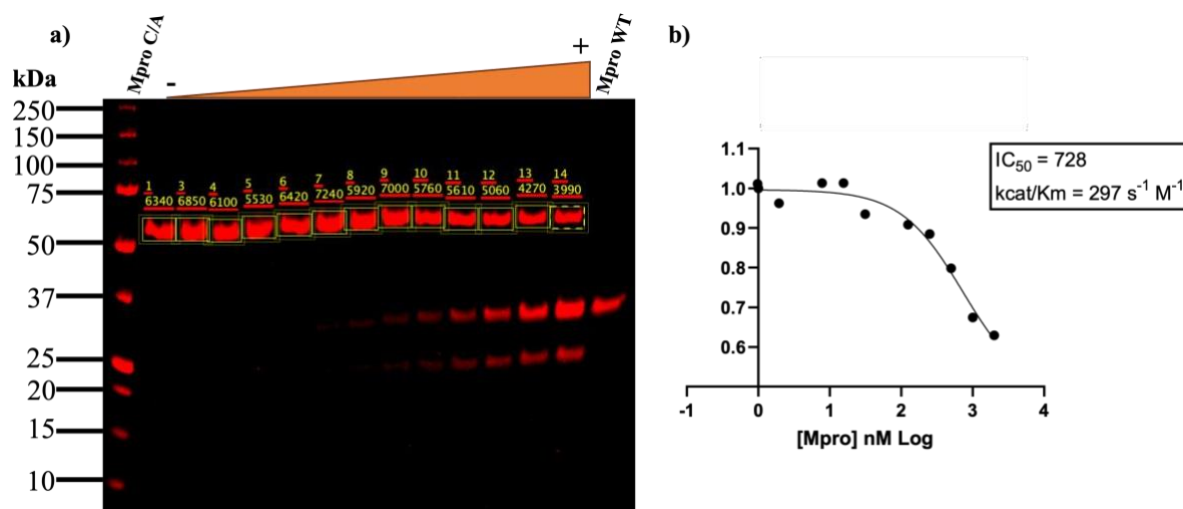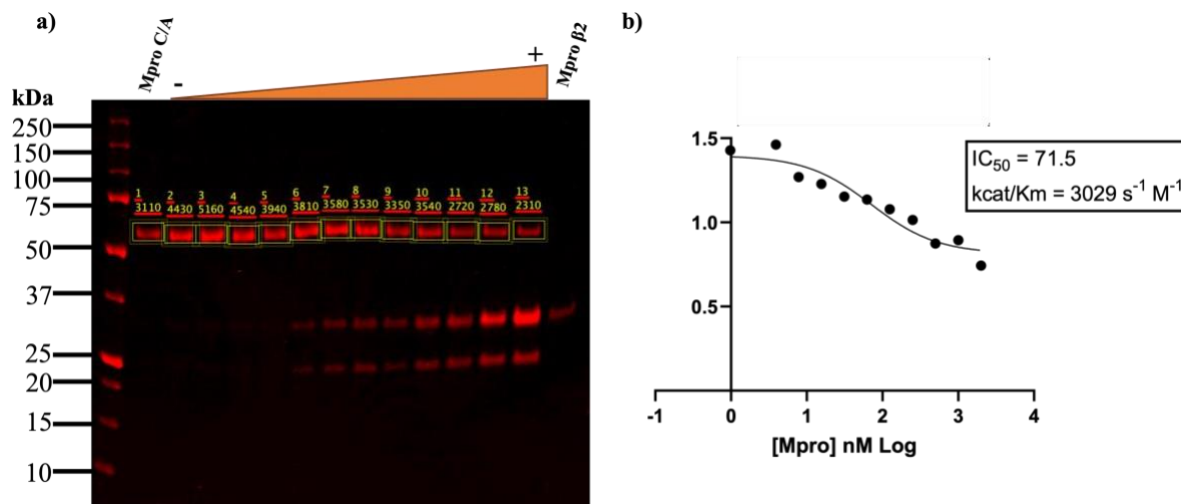

Figure S10: Protein substrate cleavage for the β M^{pro}. a) The gel was scanned and analyzed using Image Studio software, and b) the catalytic efficiency was calculated using equation (1).
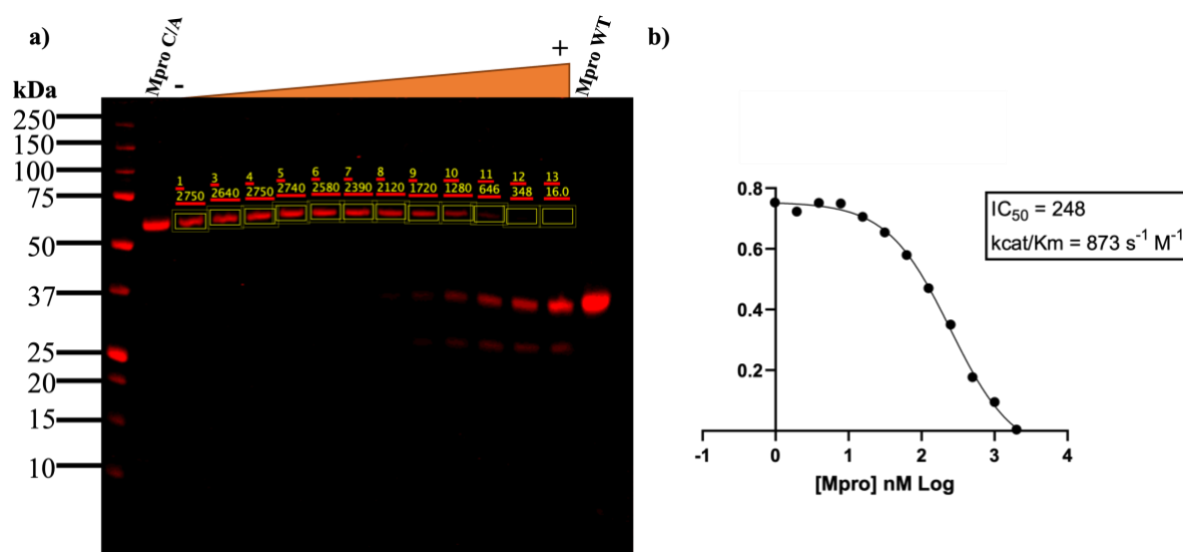
Figure S11: Protein substrate cleavage for the β M^{pro} replicate. a) The gel was scanned and analyzed using Image Studio software, and b) the catalytic efficiency was calculated using equation (1).
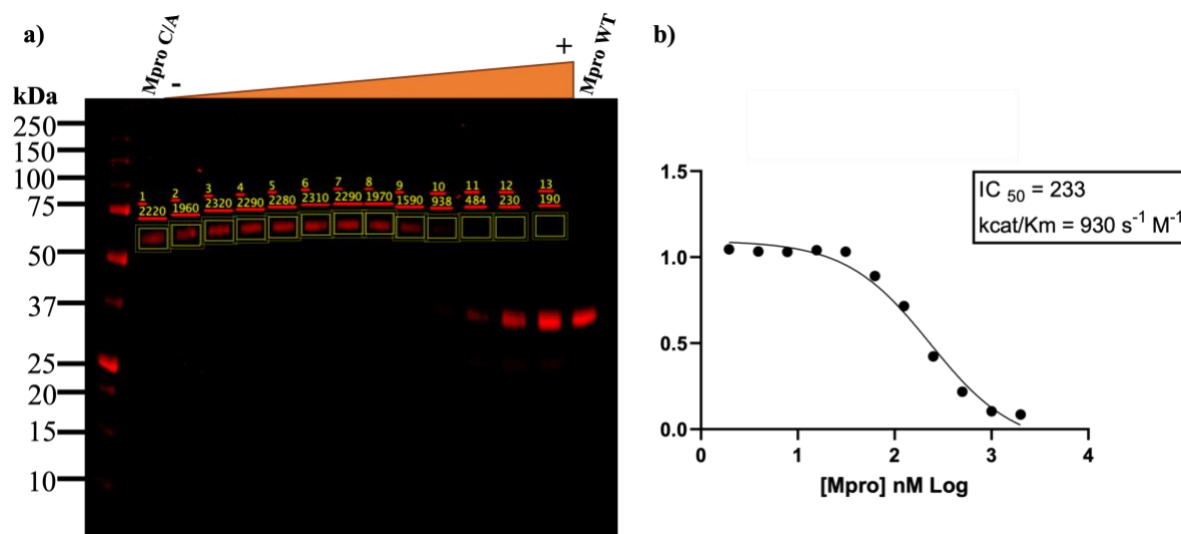
Figure S12: Protein substrate cleavage for the β M^{pro} carrying two mutations. a) The gel was scanned and analyzed using Image Studio software, and b) the catalytic efficiency was calculated using equation (1).

Figure S13: Protein substrate cleavage for the β M^{pro} carrying two mutations replicate. a) The gel was scanned and analyzed using Image Studio software, and b) the catalytic efficiency was calculated using equation (1).

Figure S14: Protein substrate cleavage for the o M^{pro} replicate. a) The gel was scanned and analyzed using Image Studio software, and b) the catalytic efficiency was calculated using equation (1).
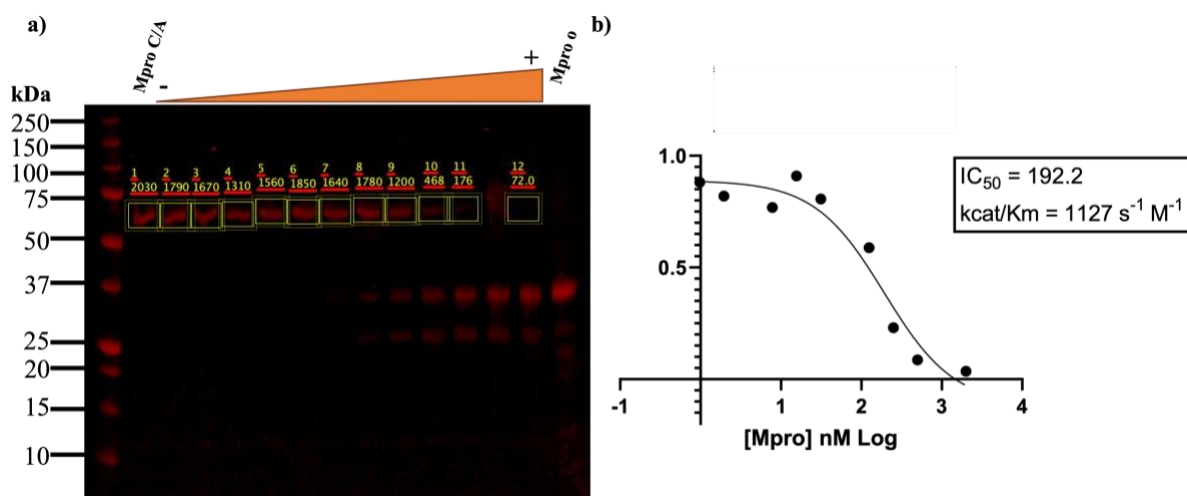
## ITC data for Neq1184 against WT SARS-CoV-2 M^{pro}

Figure S15: Isothermal titration calorimetry experiment replicate for Neq1183 against SARS-CoV-2.
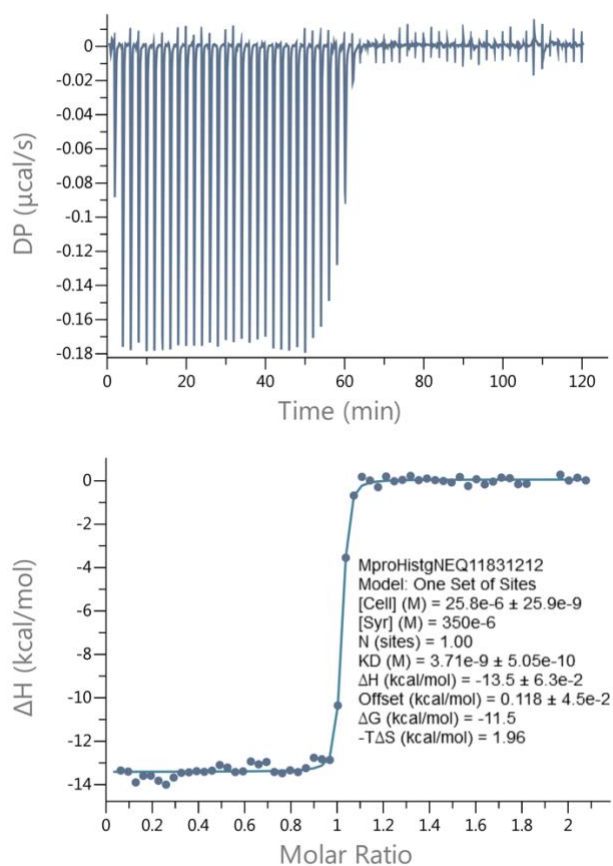
Figure S16: a) Isothermal titration calorimetry experiment for Neq1184 against SARS-CoV-2 and its thermodynamic profile (b).
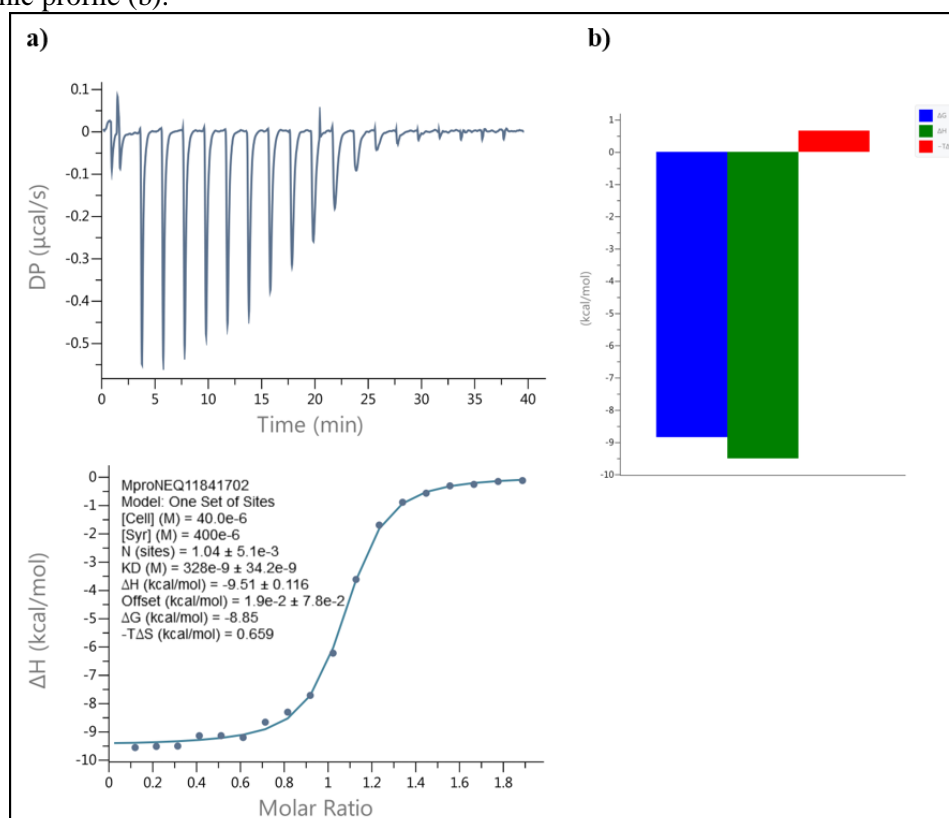
Figure S17: Isothermal titration calorimetry experiment replicate for Neq1184 against SARS-CoV-2.
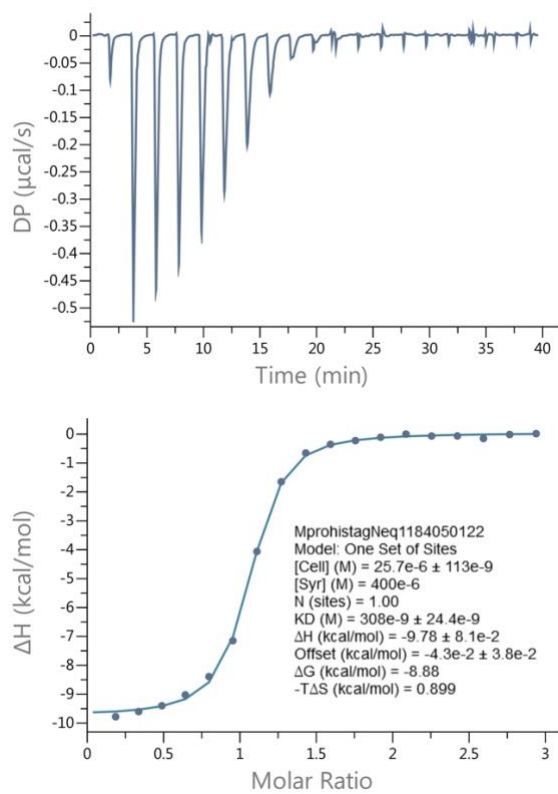
Kinetic Parameters for the irreversible binding of Neq1236

Figure S18: (A) progress curve for the inhibition of M^{pro} by Neq1236. k_{obs} was determined by using equation 2 for each inhibitor concentration. (B) The inhibitor's potency was obtained by linear regression of k_{obs} as a function of inhibitor concentration
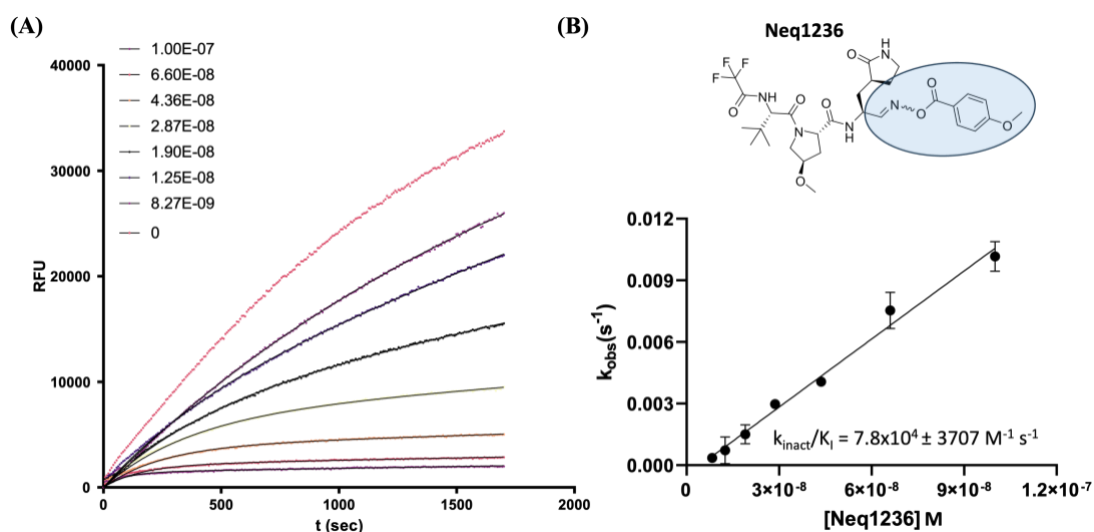
Kinetic Parameters for the human cathepsin B and S

Figure S19: Michaelis-Menten plot for the human cathepsin B with the substrate Z-FR-AMC.
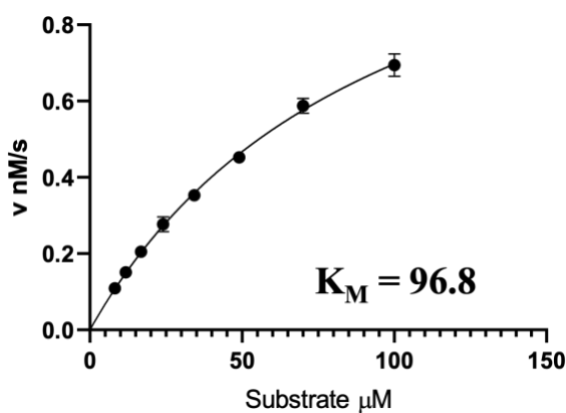
Figure S20: Michaelis-Menten plot for the human cathepsin S with the substrate Z-FR-AMC.