UNIVERSITY OF SÃO PAULO

INSTITUTE OF CHEMISTRY OF SÃO CARLOS

POST GRADUATION PROGRAM IN CHEMISTRY

Medicinal Chemistry Group - NEQUIMED/IQSC/USP

STRUCTURE-ACTIVITY RELATIONSHIPS OF
CYSTEINE PROTEASE INHIBITORS

Fernanda dos Reis Rocho

Advisor: Prof. Dr. Carlos A. Montanari

SÃO CARLOS

2019

FERNANDA DOS REIS ROCHO

STRUCTURE-ACTIVITY RELATIONSHIPS OF

CYSTEINE PROTEASE INHIBITORS

Masters dissertation presented to the post-graduation program in chemistry of the São Carlos Institute of Chemistry, as required for the achievement of the Title of Master with an emphasis on organic and biological chemistry.

Concentration Area: Organic and Biological Chemistry

Advisor: Prof. Dr. Carlos A. Montanari

SÃO CARLOS

2019

RESUMO

Enzimas são responsáveis por regular grande parte dos processos biossintéticos de organismos vivos. Devido a isso, a identificação de compostos bioativos que atuam em enzimas é fundamental na área de química medicinal. As cisteíno proteases (CPs) apresentam diversas funções essenciais em organismos vivos, logo, membros desta classe de enzimas são considerados alvos promissores para intervenções terapêuticas. Cinco CPs que estão envolvidas em um grande número de doenças humanas foram estudadas neste trabalho, tais como as catepsinas L, S e K, alvos para o câncer e osteoporose, cruzaína e a cisteíno protease da classe B (*Lm*CPB), alvos para doenças parasitárias como doença de Chagas e leishmanioses, respectivamente. Neste trabalho foi realizada a caracterização cinética de uma série de dipeptidil nitrilas que atuam como inibidores de CPs por meio de ligação covalente reversível. Modificações estruturais foram implementadas nas posições P1, P2, e P3 para avaliar a interação com os correspondentes subsítios S1, S2 e S3 das CPs. Através da análise de pares moleculares (MMP) e relação estrutura atividade (SAR), estimamos como o efeito de não aditividade para diferentes grupos nas posições P1 e P2 pode influenciar no modo geral de interação dos inibidores. Demonstramos que, apesar do grande conhecimento sobre o subsítio S2, os subsítios S1 e S3 também podem aumentar a afinidade e seletividade para as CPs desejadas. O desenovelamento térmico da cruzaína foi avaliado neste trabalho através da técnica de calorimetria diferencial de varredura (DSC), na qual um desenovelamento irreversível foi obtido. Juntamente, reportamos que alguns dos ligantes utilizados na análise foram capazes de estabilizar a estrutura da cruzaína em mais de 13 °C. Além disso, foi observada uma correlação linear entre a afinidade (pK_i) dos ligantes e os valores de ΔTm obtidos para a mesma série de inibidores.

Palavras-chave: Dipeptidil nitrilas, cisteíno proteases, relação estrutura atividade

ABSTRACT

Enzymes regulate almost every biosynthetic process of living organisms and the identification of bioactive compounds that act on enzymes is fundamental in the medicinal chemistry field. Cysteine proteases (CPs) have several essential functions in the organisms, and members of this enzyme class are considered to be potential targets for therapeutic intervention. Five CPs that are involved in a range of human diseases were studied in this work, such as cathepsin L, S and K, targets for cancer and osteoporosis, cruzain and class B cysteine protease (*Lm*CPB), targets for parasitic diseases such as Chagas disease and leishmaniasis, respectively. Kinetic characterization for a series of dipeptidyl nitriles that acts as reversible covalent inhibitors of CPs was performed. Structural modifications at the P1, P2, and P3 positions were implemented to study the role of the corresponding A1, S2 and S3 subsites of the CPs. Through a matched molecular pair (MMP) structure-activity relationship (SAR) analysis, we estimated how the nonadditivity effect for different groups in P1 and P2 can influence the general mode of binding. We show that besides the well-known S2 pocket also S1 and S3 can enhance affinity and selectivity for the desired CPs. The thermal unfolding of cruzain was evaluated in this work through differential scanning calorimetry (DSC), in which the irreversibility of cruzain unfolding was determined. Herein, we report that some of the ligands were able to stabilize cruzain structure of more than 13 ºC. In addition, a trend was observed for a linear correlation between the affinity of the ligands (pK_i) and ΔTm.

FIGURE LIST

TABLE LIST

ABREVIATION LIST

CPs – Cysteine Protease

Cz – Cruzain

*Lm*CPB – Class B Cysteine Protease from *Leishmania*

*h*CatL – Human Cathepsin L

*h*CatS – Human Cathepsin S

*h*CatK – Human Cathepsin K

CI – Covalent Inhibitors

K_i – Inhibition Constant

K_M - Michaelis Constant

V_{max} – Maximum Rate

SAR – Structure-Activity Relationship

MMPA – Matched Molecular Pair Analysis

DSC – Differential Scanning Calorimetry

Z-Phe-Arg-AMC – Z-phenylalanine-arginine-7-amido-4-methylcoumarin

Z-Leu-Arg-AMC – Z-Leucine-Arginine-7-Amido-4-methylcoumarin

k_{cat} – Apparent first-order rate constant (Turnover Number)

SUMMARY

1. INTRODUCTION

1.1. Cysteine Protease

Proteases (EC 3.4) are a large family of enzymes that hydrolyze peptide bonds in an amino acid sequence through a nucleophilic attack (VERMA; DIXIT; PANDEY, 2016). They are one of the most abundant classes of enzymes, being involved in many biological processes, such as cell-cycle, signaling, and cell death (DEU; VERDOES; BOGYO, 2012). Due to the many processes in which these enzymes are involved, they turned into targets for various types of diseases ranging from cardiovascular disorders to cancer and Alzheimer's disease, as well as for bacterial and infectious diseases ((DEU; VERDOES; BOGYO, 2012; HASANBASIC et al., 2016; KĘDZIOR; SEREDYŃSKI; GUTOWICZ, 2016; MCKERROW, 2018; ROBINSON; DALTON, 2011).

Cysteine proteases (CPs) are found in every living organism (DRAG; SALVESEN, 2010a; SIKLOS; BENAISSA; THATCHER, 2015). The most abundant CPs are papain-like, due to a high homology with the papain structure, a plant protease isolated from *Carica papaya* (POWERS et al., 2002).

CPs usually have two domains, with the active site being in a cleft between them. The active site consists of three amino acid residues: cysteine (Cys), histidine (His) and asparagine (Asn) (Figure 1). The Cys residue will be activated by the His residue creating an ionic pair that will be stabilized by the Asn residue. The reaction with the substrate will be started through a nucleophilic attack performed by the Cys thiol.

Figure 1: Structural representation of the cysteine protease Falcipain-2 with the active site (in green) in a cleft between the two domains, with the catalytic triad Cys-His-Asn (PDB ID 3PNR).
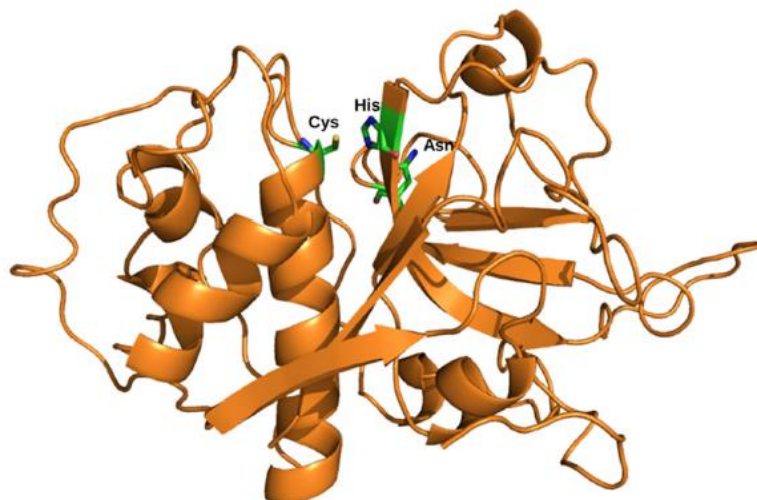
In detail, the mechanism of peptide hydrolysis (Figure 2) initiates with the deprotonation of the thiol group of the Cys by the His residue. Then, the nucleophilic Cys binds the carbonyl carbon of the substrate, forming the first tetrahedral intermediate through an acylation reaction. The intermediate is stabilized by the oxyanion hole, a flexible cavity in the CPs structures, through a hydrogen bond. Afterward, the amino-terminal fragment of the substrate is released, and the His residue is restored to the deprotonated form. The hydrolysis of the acyl-enzyme intermediate occurs and a second tetrahedral intermediate is formed and stabilized by the oxyanion hole. Finally, the carboxylic acid formed is released, reestablishing the active site of the enzyme (ARAFET; FERRER; MOLINER, 2017; POWERS et al., 2002; ZHAI; MEEK, 2018).

Figure 2: Schematic representation of the mechanism of peptide hydrolysis played by cysteine proteases.
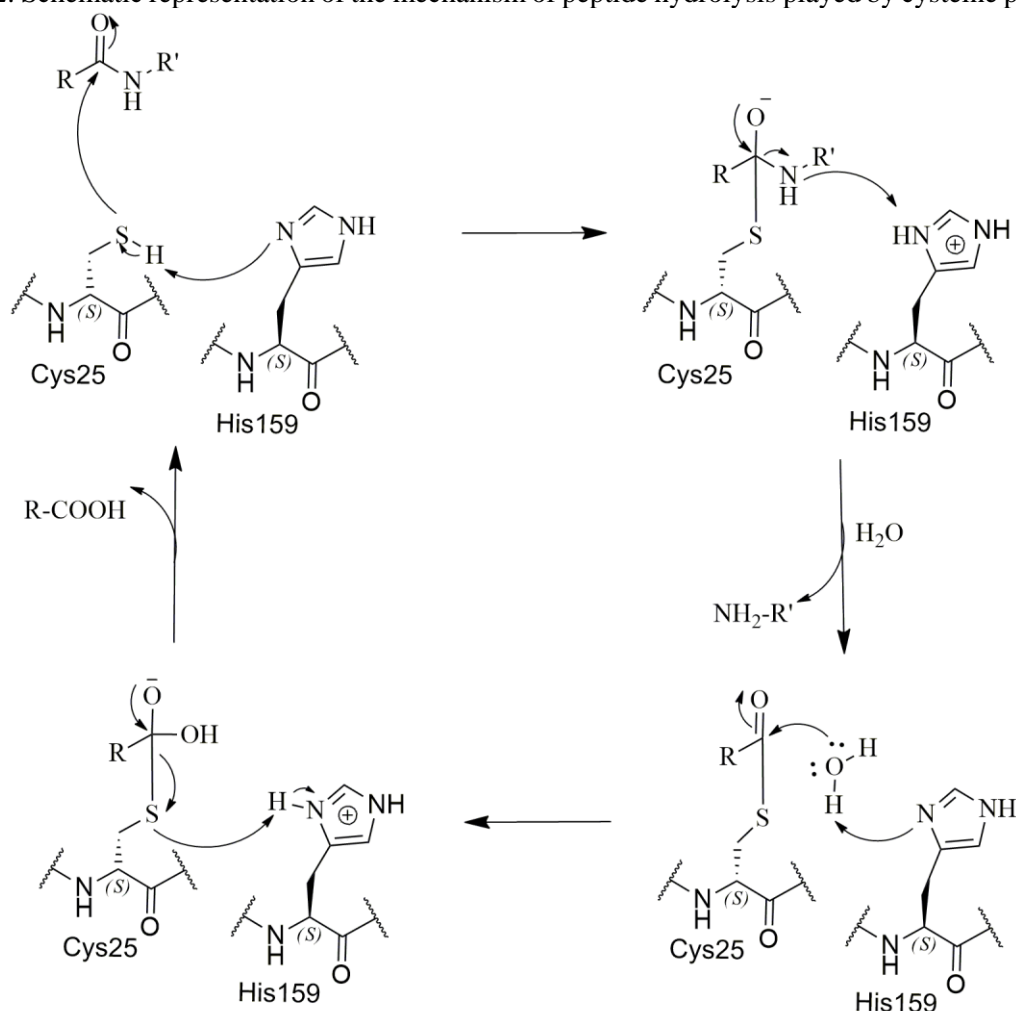
Source: Adapted from ZHAI, X.; MEEK, T. D. Catalytic Mechanism of Cruzain from *Trypanosoma cruzi* As Determined from Solvent Kinetic Isotope Effects of Steady-State and Pre-Steady-State Kinetics. **Biochemistry**, v. 57, n. 22, p. 3176–3190, 5 Jun. 2018

The proteases can recognize their substrates through the peptide sequence or motif that will bind to specific subsites pockets of the active site cleft (Figure 3) (DRAG; SALVESEN, 2010b). In that way, most of the inhibitors design for these enzymes are peptidic substances that interact with specific subsites of the enzyme.

Figure 3: Recognition subsites pockets nomenclature used for proteolytic enzymes. The numbering increase in the direction of the N-terminal residue of the enzyme.
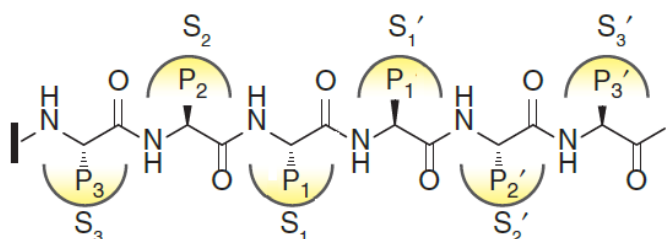
Source: Adapted from SIKLOS, M.; BENAISSA, M.; THATCHER, G. R. J. Cysteine proteases as therapeutic targets: does selectivity matter? A systematic review of calpain and cathepsin inhibitors. **Acta Pharmaceutica Sinica B**, v. 5, n. 6, p. 506–519, nov. 2015.

Five CPs were studied in this work that present functional structural similarities in which three are mammalian cathepsins: Cathepsin L (*h*CatL), Cathepsin S (*h*CatS) and Cathepsin K (*h*CatK), and two protozoa CPs: cruzain (Cz) from *Trypanosoma cruzi* and *Leishmania mexicana* CPB (*Lm*CPB). The human cathepsins are widely studied for being essential targets for cancer and immune system-related diseases (DENNEMÄRKER et al., 2010; STONE et al., 2019). As a result of that, a lot is known of their structures and substrate specificity, having a large number of inhibitors deposited in the ChEMBL data bank (GAULTON et al., 2017).

Protozoan CPs, Cz and *Lm*CPB are essential targets for Chagas disease and Leishmaniasis, respectively, which display crucial roles in the life cycle of the parasites. Cz is a well-established target for Chagas disease, and by that, much is known of its structure. Also there is a significant number of inhibitors deposited in the ChEMBL databank (GAULTON et al., 2017). While for *Lm*CPB, little is known of its structure and inhibitors, with only one crystal structure deposited in the PDB databank. However, *Lm*CPB is considered a promising target for the treatment of *Leishmania mexicana*, the cutaneous form of *Leishmania donovani* (WILLIAMS et al., 2006).

In that way, these five CPs were studied with 22 inhibitors aiming to identify new hits that display high affinity and selectivity for these enzymes.

1.1.1. *Cruzain*

Cruzain, also assign as cruzipain (the full-length native enzyme), is the major cysteine protease expressed in the parasite *Trypanossoma cruzi* (DA SILVA; DO NASCIMENTO PEREIRA; FERREIRA, 2016). Cz its present in all stages of the parasite life cycle, inducing damage to the host tissue and facilitating the parasite invasion. Therefore, its inhibition interrupts several cellular processes leading to the death of the parasite (ATKINSON; BABBITT; SAJID, 2009; DOYLE et al., 2011; ROGERS et al., 2012).

Cz is an endopeptidase enzyme that is expressed as a pro-enzyme (BARBOZA et al., 2005). It has two domains (Figure 4), one mostly formed of α-helix secondary structure and the other consisting of an extensive β-sheets interaction. The catalytic triad consists of three amino acid residues, Cys25, His159, and Asn175 (AVELAR et al., 2015; GILLMOR; CRAIK; FLETTERICK, 1997; HERRERA-MAYORGA et al., 2019).

Figure 4: (a) Three-dimensional structure representation of Cruzain bound covalently to a vinyl α-ketoester (PDB ID 1U9Q). (b) Cruzain binding site showing the S1, S2 and S3 pocket, covalently bound to a dipeptidyl nitrile (NEQUIMED/IQSC/USP PBD ID 4QH6).
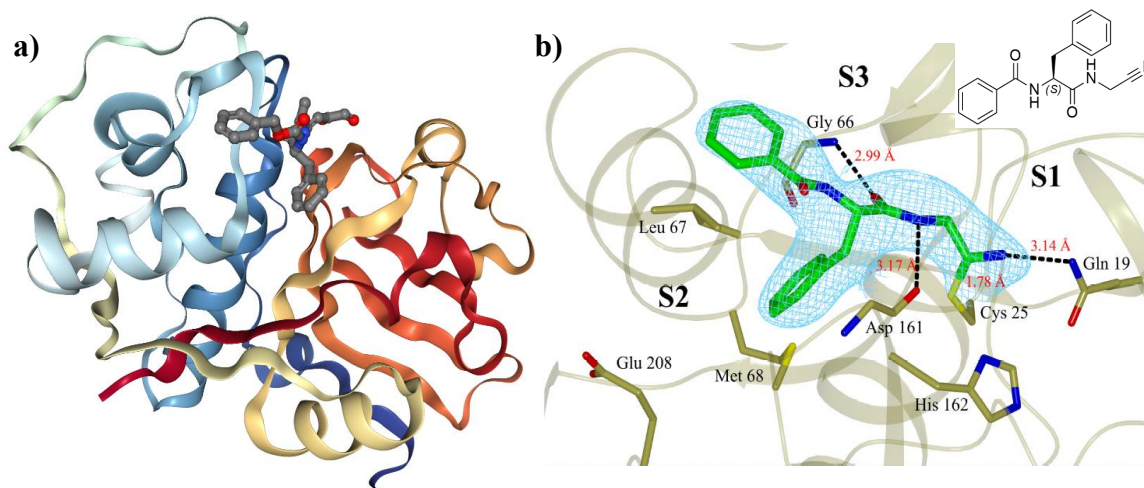
Cz is a member of the cathepsin L-family, being *h*CatL like; in that way, it has a high degree of identity with *h*CatL (ATKINSON; BABBITT; SAJID, 2009). Cz, as the other CPs, has a high preference for noncharged aliphatic or aromatic amino acids side chains, with a hydrophobic characteristic (ROBINSON; DALTON, 2011).

The vinyl sulfone K777 (Figure 5) is one of the most known Cz irreversible covalent inhibitor. It was first described by the Sandler Center for Research in Tropical Parasitic Diseases (University of California) as a promising drug for the treatment of Chagas disease.

The compound was active against a wide range of *T. cruzi* strains and reached the preclinical phase without progressing to the next phase, due to high toxicity problems, probably related to its mode of action (MCKERROW et al., 2009; SAJID et al., 2011).

A promising Cz reversible inhibitor is the BaCZ007 (Figure 5b) that reached the preclinical phase, developed by Beaulie and co-workers. BaCZ007 has a nitrile as warhead and displays a reversible covalent inhibition, different from the K777. Interestingly, the compound exhibit low toxicity, probably due to its reversibility (BEAULIEU et al., 2010).

Figure 5: (a) Covalent irreversible inhibitor K777 structure developed by the University of California and (b) Covalent reversible inhibitor BaCZ007 structure of the Merck pharmaceuticals, both cruzain inhibitors.
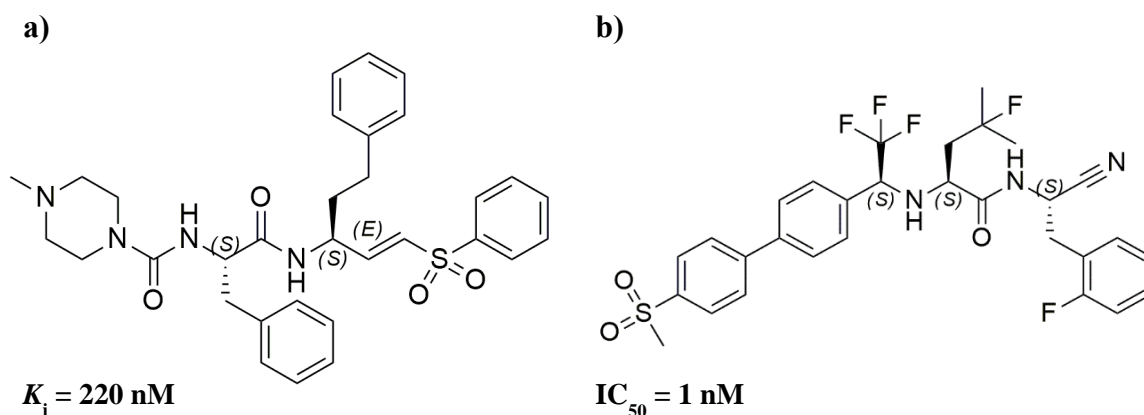
Source: Adapted from BEAULIEU, C. et al. Identification of potent and reversible cruzipain inhibitors for the treatment of Chagas disease. **Bioorganic & Medicinal Chemistry Letters**, v. 20, n. 24, p. 7444–7449, dez. 2010 and MCKERROW, J. et al. Two approaches to discovering and developing new drugs for Chagas disease. **Memórias do Instituto Oswaldo Cruz**, v. 104, n. suppl 1, p. 263–269, jul. 2009.

Nitrile class of inhibitors is attractive for research in medicinal chemistry due to their high affinity and low toxicity.

1.1.2.   LmCPB

*Lm*CPB is one of the three families of CPs genes expressed in the genus *Leishmania*, along with CPA and CPC. The class b cysteine protease (CPB) is also part of the papain family and has essential involvement in the parasite life cycle in *L. mexicana,* (MOTTRAM; BROOKS; COOMBS, 1998; SANDERSON et al., 2000). Alongside with CPA, the enzymes display an important role in autophagy and differentiation of promastigotes of the parasite to infective megacycles (WILLIAMS et al., 2006). Recent studies have shown that *Lm*CPB is responsible for cleaving host cell proteins and for causing modifications in cell signaling and

function of the immune system. This occurs in the promastigote and amastigote stages of the parasite (ABU-DAYYEH et al., 2010).

*Lm*CPB is the recombinant form of the amastigote specific isoform CPB2.8, lacking the C-terminal extension, denominated CPB2.8ΔCTE. The enzyme is a typical cathepsin-L-like endopeptidase (DE LUCA et al., 2018). *Lm*CPB is also expressed as a pro-enzyme, and it has two domains structure with a catalytic triad consisting of a Cys26, His164, and Asn184.

This enzyme has only one high resolution X-ray crystallography structure so far. The structure was obtained by the NEQUIMED/IQSC/USP group and co-workers of the University of Nottingham, in which an azanitrile inhibitor is covalently bound to the enzyme. The structure was deposited in the PDB databank with the ID code 6P4E.

*Lm*CPB has only five inhibitors reported in the ChEMBL database, due to that, new studies of this enzyme can have immense contribution in the medicinal chemistry field.

1.1.3.   *Cathepsin Family*

The cathepsin family is constituted by 11 members, in which the majority are endopeptidases being all synthesized as zymogens in the human genome. They present a two-domain structure and are also part of the papain-family (KRAMER; TURK; TURK, 2017).

The human cathepsins play chief roles in tumor progression, cardiovascular diseases, osteoporosis and arthritis, neurodegenerative diseases and obesity, making them attractive as drug targets (DENNEMÄRKER et al., 2010; OLSON; JOYCE, 2015; STONE et al., 2019).

The *h*CatL is an endosomal/lysosomal CP one of the most expressed cathepsins, which plays a vital role in degrading intracellular proteins. Like the other cathepsins, the enzyme is synthesized as a zymogen, and it is activated in an acidic pH (POREBA et al., 2018). It is known to be involved in tumor invasion and metastasis for many types of cancers like ovarian, breast, prostate, lung, gastric and pancreatic (DENNEMÄRKER et al., 2010; QUILLES JR et al., 2019; TABISH et al., 2019),

*h*CatK is predominantly expressed in osteoclasts. It is a therapeutic target for bone diseases, such as osteoporosis (STONE et al., 2019; ZAIDI et al., 2001). The *h*CatK are also involved in atherosclerosis and cancer (MULDER et al., 2014). One of the most promising substances that act as an inhibitor of *h*CatK is Odanacatib (Figure 5) that has reached phase III clinical trials (MULLARD, 2016). Another promising drug is the MIV-711 ($K_i = 0.98$ nM) a candidate for the treatment of atherosclerosis developed by the Medivir company. The drug has

yet no chemical structure published, and it is currently on phase II of clinical trials (CONAGHAN et al., 2018; LINDSTRÖM et al., 2018).

hCatS is a lysosomal enzyme linked to the modulation of the immune response. It has 57 % of identity with hCatL and hCatK but differs with these enzymes by its preference for neutral and slightly basic pH (MCGRATH et al., 1998). hCatS is a therapeutic target for several types of diseases such as multiple sclerosis (FOTI CUZZOLA et al., 2012), psoriasis (AINSCOUGH et al., 2017) and obesity (CHEN et al., 2018). Many inhibitors of this enzyme are in clinical trials; the inhibitor VBY-891, for example, has undergone phase I studies, with positive safety feedback. The structure of this substance is not yet published.

Figure 6: The dipeptidyl nitrile Odanacatib structure, a covalent reversible inhibitor of the human cathepsin K.

$IC_{50} = 0.2$ nM
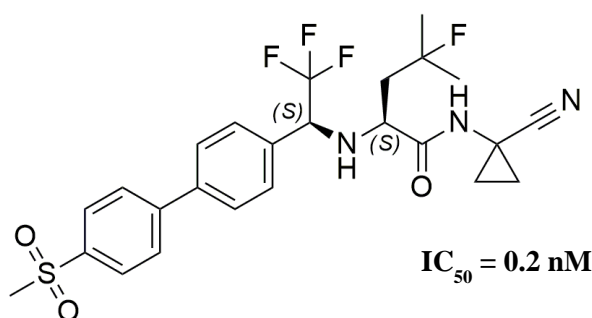
Source: Adapted from MULLARD, A. Merck & Co. drops osteoporosis drug odanacatib. **Nature Reviews Drug Discovery**, v. 15, n. 10, p. 669–669, out. 2016.

The human cathepsins L, K, and S, share a high degree of structural homology, with a similar active site region (Figure 6). The sequence forming the S1 subsite for all the three cathepsins is precisely the same, with small changes in their shapes. The S2 and S3 subsites are less conserved. The S2 subsite for hCatL and hCatS have a deep binding pocket, while hCatK has a small S2 pocket. The three CPs have a preference for hydrophobic residues.

Figure 7: Crystallographic structure representation of the cathepsins L, K, and S subsites. The PDB ID codes for the enzymes are 2Y2J, 4X6H, and 3OVX, respectively.
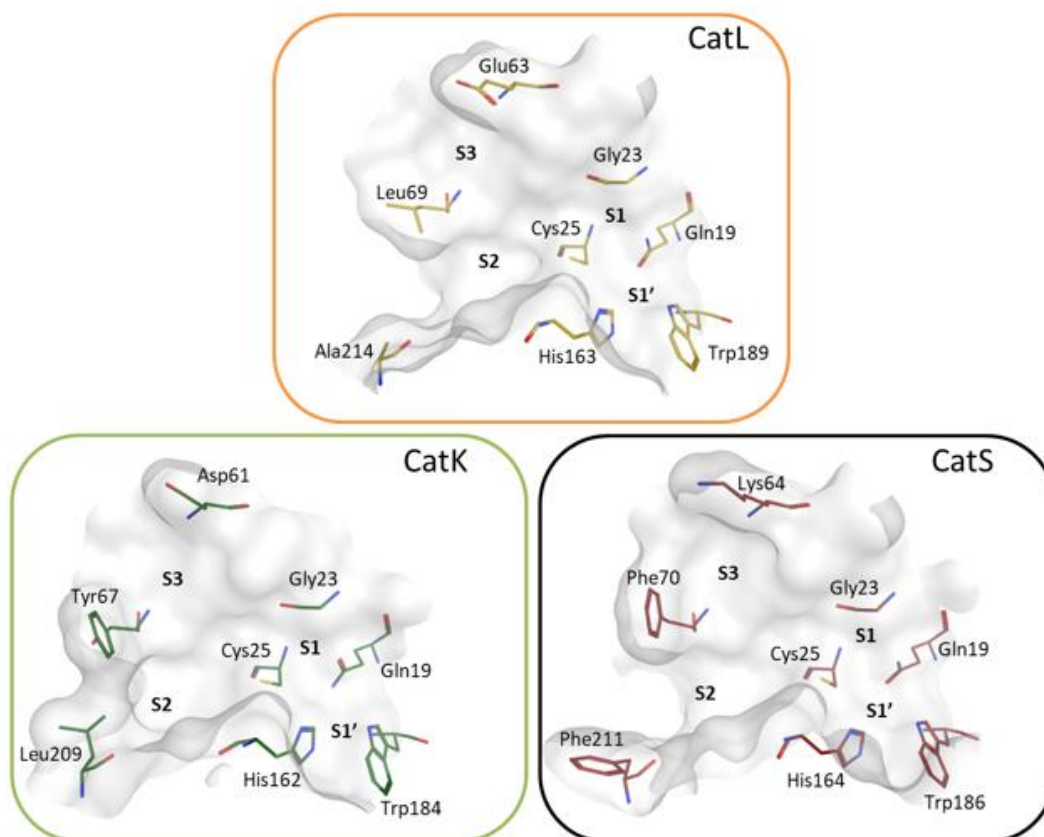
Source: Adapted from CIANNI, L. et al. Can Cysteine Protease Cross-class Inhibitors Achieve Selectivity? **Journal of Medicinal Chemistry**, 30 jul. 2019.

1.2. Covalent Reversible Inhibitors

The CPs are effectively modulated by several classes of covalent inhibitors (CI) due to the presence of the active Cys residue of the enzyme, which facilitates nucleophilic attack to a susceptible electrophile. Examples of successful drugs that act through a covalent mechanism are aspirin and omeprazole (ÁBRÁNYI-BALOGH et al., 2018; SINGH et al., 2011).

CI possesses many advantages over non-covalent compounds, such as high affinity that can be translated into lower and less frequent dosing with decreased potential for off-target effects. Covalent drugs also result in long residence times on the target, which can lead to prolonged action (ÁBRÁNYI-BALOGH et al., 2018; DE CESCO et al., 2017; SMITH et al., 2009).

Covalent inhibitors can bind to the enzyme reversibly or irreversibly, depending on the rate of the reverse reaction (Figure 8).

Figure 8: Generic mechanism in two-step played by covalent inhibitors, where E is the target enzyme, I is the inhibitor, E.I is the reversibly bound enzyme-inhibitor complex, and E-I is the formation of a covalent bond between the target and inhibitor.
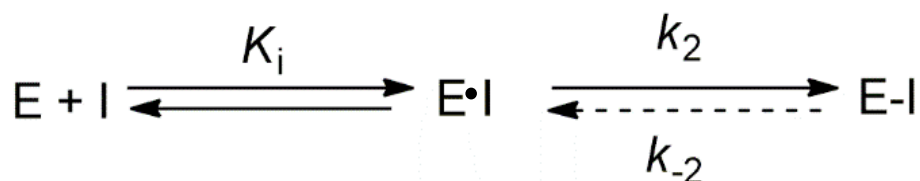
Source: Adapted from STRELOW, J. M. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. **SLAS DISCOVERY: Advancing Life Sciences R&D**, v. 22, n. 1, p. 3–20, Jan. 2017.

The K_i (inhibition constant) is the noncovalent binding constant that tells how strong the affinity of the inhibitor for the target is. Reversible covalent inhibitors have a finite value for the constants k_2 and k_{-2}, as for irreversible inhibitors, the k_{-2} will be virtually zero (BAILLIE, 2016; SINGH et al., 2011). For the therapeutic field, reversible inhibitors are more interesting than irreversible ones, due to the high toxicity that sometimes comes along with irreversible drugs.

CI has an electrophilic center responsible for binding to the enzyme known as a warhead. The warhead modulates affinity, selectivity, and the mechanism of inhibition, being an essential moiety for the recognition of the molecule by its target. A wide range of functional groups can be used as warheads (SILVA et al., 2017).

In the present work, a nitrile moiety was used as a warhead. Nitriles are of abundant interest due to its countless versatility, related to the short and polarized triple bond, which makes it possible to go into deep clefts (BERTEOTTI et al., 2014), displaying as well, good metabolic stability (FLEMING et al., 2010). The interaction of this moiety with the CPs involves the *sp* carbon of the nitrile that will form a reversible covalent bond with the Cys thiol.

A class of inhibitors called dipeptidyl nitrile was used in this work, these compounds are known to be CPs inhibitors for decades, presenting good inhibitory activity and selectivity (AVELAR et al., 2015; OTTO; SCHIRMEISTER, 1997).

A dipeptidyl nitrile scaffold (Figure 9a) was designed considering the recognition sites for proteolytic enzymes. The aim was to determine which variations in P1, P2, and P3 (Figure 9b), lead to a higher affinity with the targets.

Figure 9: (a) Dipeptidyl nitrile scaffold with the nitrile group as a warhead. (b) **Neq0570** and the mode of interaction with the binding site of the CP.
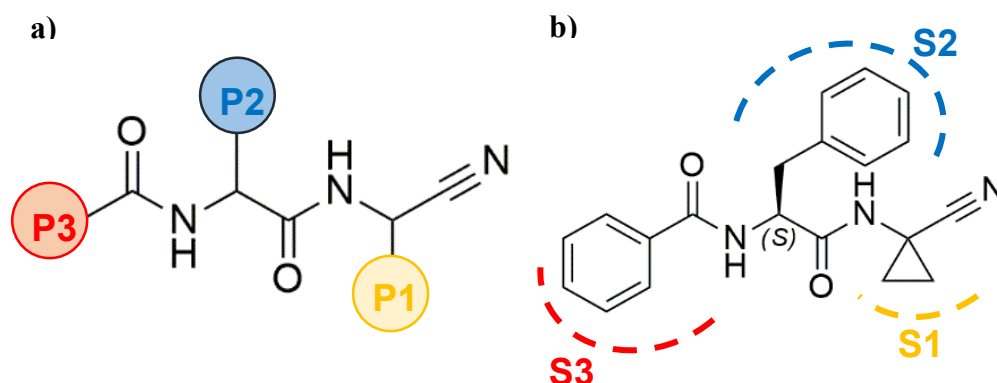
Source: Own author

The substances used in this work will be analyzed by their affinity, through the analysis of the structural-activity relationship and by comparison between the dot structure modifications in pairs of molecules.

1.3. Structure-Activity Relationship

Drug discovery is a multidisciplinary area, in which compounds need to be optimized for several reasons, which includes affinity improvement, bioavailability, safety, among other properties before they become a promising drug candidate. One way to optimize compounds is by using structure-activity relationship (SAR) analysis.

SAR is the relationship between the chemical or 3D structure of a molecule and its biological activity. SAR is a fundamental analysis in drug discovery, from primary screening to lead optimization and can be divided into two groups – those based on regression models (QSAR) and those based on physicochemical approaches (GUHA, 2013). An understanding of the SAR for a set of compounds, allows one to elucidate the chemical space of the desired molecule and use that information for further optimization regarding physicochemical properties or activity/selectivity index (GUHA, 2013; WAWER; BAJORATH, 2011).

In a research conducted by the NEQUIMED/IQSC/USP group at the University of São Paulo (USP), a SAR study was conducted for some Cz inhibitors. Several structural modifications at the P2 and P3 positions were performed in a dipeptidyl nitrile scaffold (Figure 10). The modifications led to a compound with over 100 fold higher affinity toward Cz than the prototypical compound, providing information about the enzyme preferences in the S2 and S3 pocket (AVELAR et al., 2015).

Figure 10: Structural representation of a dipeptidyl nitrile as cruzain covalent reversible inhibitor used for SAR studies, with the positions were structural modifications occurred market as blue (P2) and red (P3) in the molecule structure.

Neq0409 pK_i = 6.3
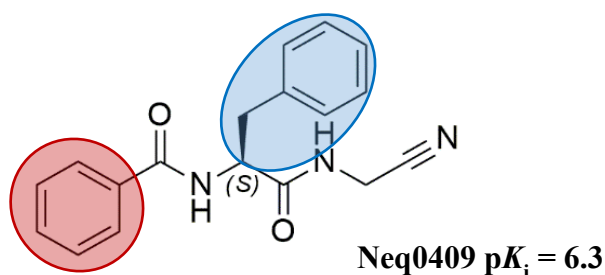
Source: Adapted from AVELAR, L. A. A. et al. Molecular Design, Synthesis and Trypanocidal Activity of Dipeptidyl Nitriles as Cruzain Inhibitors. **PLOS Neglected Tropical Diseases**, v. 9, n. 7, p. e0003916, 14 Jul. 2015

1.4. Matched Molecular Pair Analysis (MMPA)

MMPA is a concept that seeks to evaluate substances properties associated with a single located structural change, as exemplified in Figure 11 (GRIFFEN et al., 2011). The substructure shared between MMP partners is called key fragment, and the substructures that are different are called value fragment. If this change in the affinity, commonly expressed as pK_i values, corresponds to be equal or higher than 2.0 log units, matched molecular pairs (MMPs) are called activity cliffs. If no meaningful change (ΔpK_i ± 0.2 log units) is present, such pairs are called bioisosteres (FUCHS et al., 2015; KRAMER et al., 2014; STUMPFE; BAJORATH, 2012).

Figure 11: MMPA for two dipeptidyl nitrile inhibitors of Cathepsin B, where a change, shown in red, for the compound (1) from the compound (2), in orange, exhibit a difference in the activity of 1.0 log unit.

ΔpK_i = + 1.0

(1)   **pK_i = 5.9**   **(2)**   **pK_i = 6.9**
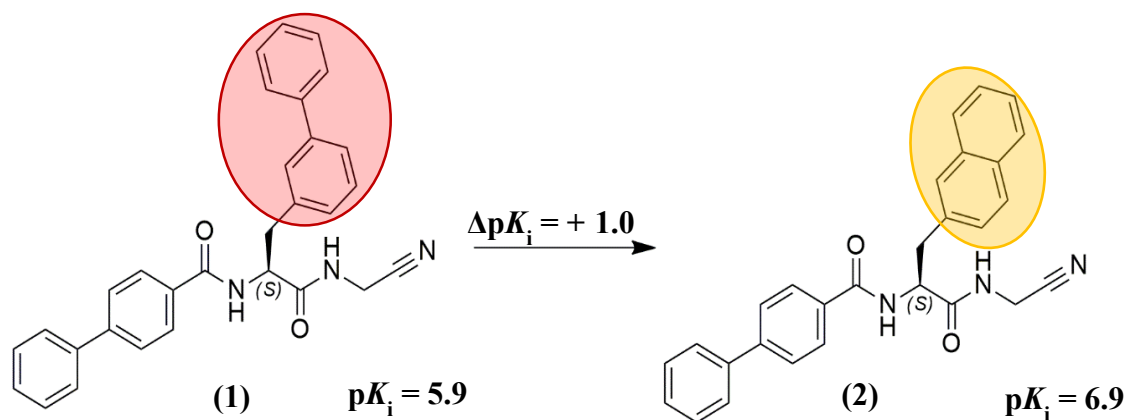
Source: Adapted from SCHMITZ, J. et al. Cathepsin B Inhibitors: Combining Dipeptide Nitriles with an Occluding Loop Recognition Element by Click Chemistry. **ACS Medicinal Chemistry Letters**, v. 7, n. 3, p. 211–216, 10 Mar. 2016.

The relationships between MMPs may be simple, as a hydrogen atom change for a chlorine atom, or more complex involving a significant change in some part of the molecule (BIRCH et al., 2009). Despite the structural changes, MMPA results also depend on the chemical environment. The chemical properties of an atom/group can have a pronounced impact on the effect of the chemical transformation (KRAMER et al., 2014; PAPADATOS et al., 2010)

A crucial role of the MMPA is its use in the optimization of molecules aiming to obtain a relationship between structure and activity (SAR), having an advantage of dealing directly with measured data, providing clear understanding and interpretation of the results (GRIFFEN et al., 2011).

In the present work, the affinity of a large number of inhibitors was determined using a fluorimeter, in which the pK_i values were obtained. We were able to use similar protocols for the five CPs making a fast and efficient determination of the constants.

Furthermore, a study using the differential scanning calorimetry (DSC) technique was performed with Cz. The technique allows one to identify the temperature in which the protein is 50% in the folded form and 50% unfolded (Tm) by heating the protein solution at constant pressure. Along with the Tm determination, the ΔCp and ΔH of the process of unfolding are also determined. Therefore, the thermal stability investigation of Cz was performed in the presence of some of the bests inhibitors in this work, using DSC, seeking to evaluate the magnitude of stabilization (ΔTm) of the Cz structure when incubated with the ligands.

2. OBJECTIVES

The main goal of this work is to evaluate the affinity of 22 covalent reversible inhibitors with a dipeptidyl nitrile scaffold against five cysteine proteases of the papain family, being two from parasites and three from mammalians. Additionally, characterize the thermal unfolding of cruzain, as well as to evaluate the thermal stabilization of its structure in the presence of different ligands by differential scanning calorimetry (DSC).

2.1. Specific Objectives

I.   Determine the inhibition constants (K_i) for a range of covalent reversible inhibitors against cruzain, *Lm*CPB, and *h*CatL using a fluorimetric technique.
II. Select some inhibitors based on their structures similarities with a pK_i equal or higher than 7.0, which are going to be tested against *h*CatS and *h*CatK.
III. Analyze the chosen compounds relating their structures and activities with the five CPs.
IV. Study the thermal unfolding of cruzain.
V.   Study the thermal unfolding of cruzain in the presence of ligands.

3. MATERIAL AND METHODS

The enzymes cruzain and *Lm*CPB were expressed in the NEQUIMED/IQSC/USP laboratories by Ms. Fabiana Rosini and Dr. Jean F. R. Ribeiro, respectively. The synthesis of the inhibitors used in this work was performed by Ms. Lorenzo Cianni, Dr. Daniela de Vita, and Dr. Anwar Shamin.

All reagents and solvents were purchased from the companies Sigma-Aldrich, Combi-Blocks, and Bachem. No prior procedures of purification were performed. The human enzymes Cathepsin-L, Cathepsin-S, and Cathepsin-K, were purchased from the companies Sigma-Aldrich and Enzo Life Sciences.

The computer software used to analyze the fluorimeter (spectroscopic technique) and the DSC data were the Microsoft Excel, Gen5™Biotek®, SygmaPlot 10, Origin 8.5 and Nano Analyze.

3.1. Enzyme Kinetics Assays

The assays were conducted with the fluorimeter equipment Biotek Synergy HT system, in which the rate of hydrolysis of the fluorogenic substrates Z-Phe-Arg-AMC (Sigma-Aldrich/ Combi-Blocks) and Z-Leu-Arg-AMC (Bachem) were monitored for five enzymes. Different protocols were used in the assays, as reported below.

3.1.1. Determination of the Michaelis constant (K_M) for Cruzain, LmCPB, and hCatL

The enzyme kinetics assays were done at 25 °C using a 96-well black microplate (Corning®) with an excitation wavelength of 360 nm and an emission wavelength of 460 nm. The enzymes were activated with an assay buffer 100 mM sodium acetate pH 5.5, 300 mM of NaCl and 5 mM of EDTA (Sigma-Aldrich), containing 7 mM DTT (Sigma-Aldrich), 0.014 % of Triton X-100 (Sigma-Aldrich), in an ice bath between 15-20 minutes, depending on the enzyme. The concentrations used for the enzymes were 0.15 nM, 7 nM and 1.9 nM for cruzain, *Lm*CPB and *h*CatL (Sigma-Aldrich/Enzo Life Sciences), respectively. The reaction rate was followed for 5 min in the fluorimeter, in which the rate of hydrolysis of the substrate Z-Phe-Arg-AMC was monitored. It was used a final concentration of DMSO (Sigma-Aldrich) of 5 %

(10 µL) in each well. The assays were performed in triplicates. It was used eight different substrate concentrations, prepared with a 0.5 dilution factor.

3.1.2. Determination of the Michaelis constant (K_M) for hCatK

Human recombinant Cat-K (Enzo Life Sciences) was assayed at 25 °C using a 96-well black microplate with an excitation wavelength of 360 nm and an emission wavelength of 460 nm. The enzymes were activated at 37 °C (SANYO CO_2 incubator) for 30 minutes in the assay buffer 100 mM sodium acetate pH 5.5, 300mM of NaCl and 5 mM of EDTA, containing 7 mM DTT, 0.014 % of Triton X-100. The concentration of the enzyme was 0.15 nM. The reaction rate was followed for 5 min in the fluorimeter, in which the rate of hydrolysis of the substrate Z-Leu-Arg-AMC was monitored. It was used a final concentration of DMSO of 5 % (10 µL) in each well. The assay was performed in triplicates. It was used eight different substrate concentrations, prepared with a 0.5 dilution factor.

3.1.3. Determination of the Michaelis constant (K_M) for hCatS

Human recombinant Cat-S (Enzo Life Sciences) was assayed at 25 °C using a 96-well black microplate with an excitation wavelength of 360 nm and an emission wavelength of 460 nm. The enzymes were activated with an assay buffer 100 mM sodium citrate pH 6.0, 50 mM of NaCl and 2 mM of EDTA, containing 7 mM DTT, 0.014 % of Triton X-100, and it was then incubated at 37 °C for 60 minutes. The concentration of the enzyme was 1.12 nM. The reaction rate was followed for 5 min in the fluorimeter, in which the rate of hydrolysis of the substrate Z-Phe-Arg-AMC was monitored. It was used a final concentration of DMSO of 5 % (10 µL) in each well. The assay was performed in triplicates. It was used eight different substrate concentrations, prepared with a 0.5 dilution factor.

3.1.4. Determination of the inhibition constant (K_i) for cruzain, LmCPB, hCatL, hCatK, and hCatS

The inhibition constants were determined through an indirect test, in which substrate and inhibitor are in the same solution of the enzyme. It was used a competitive inhibition protocol; the experiment was already done by previous members of the NEQUIMED/IQSC/USP group (CIANNI et al., 2018).

For enzymes activation steps and enzymes concentrations, it was used the same protocol described above (section 3.1). The substrates concentration used in the assay was fixed and equal to the K_M, so [S] = K_M.

The initial concentrations of the inhibitors varied from 50 to 0.01 μM, in which seven solutions were prepared in DMSO with a 0.5 dilution factor. Negative control was performed with DMSO. The experiment was done in triplicates. The final concentration of DMSO in each well was 5 %.

The inhibitors were incubated with the enzyme solution for 2 minutes before the reaction started by the addition of the substrate.

The apparent inhibition constant of the inhibitors was calculated using a non-linear adjustment of the data, where v_s is the steady-state velocity, v_0 is the velocity in the absence of inhibitor and [I] is the inhibitor concentration.

$$v_S = \frac{v_0}{1+\frac{[I]}{K_i^{app}}} \qquad \text{Equation 2}$$

The actual inhibition constant K_i was determined by the correction of K_i^{app}, using the Cheng and Prusoff (YUNG-CHI; PRUSOFF, 1973) equation. Substrates concentrations were equal to the K_M value.

$$K_i = \frac{K_i^{app}}{1+\frac{[S]}{[KM]}} \quad \rightarrow \quad K_i = \frac{K_i^{app}}{2} \qquad \text{Equation 3}$$

The assays were conducted, if necessary, until a higher concentration of 50 μM for weak inhibitors. When the pK_i value was lower than 5.0 the results were displayed as a percentage of inhibition, which is equal to the RFU s^{-1} obtained for the highest inhibitor concentration divided by the RFU s^{-1} obtained for the blank sample.

3.2. Differential Scanning Calorimetry Assay

The thermal stability assay was carried out for cruzain. The Tm, temperature in which half of the protein is in the folded and unfolded form, were determined along with the ΔH_{Tm}. The experiment was performed in a differential scanning calorimetry- Nano DSC of the TA Instruments. The instrument consists of two cells with a 600 μL of capacity, being one for the sample solution and the other for the reference.

3.2.1. Cruzain Reversibility Assay

Cruzain was thawed at room temperature for approximately 5 minutes. After that, the protein was added in a dialysis membrane (Amicon Ultra) of 15 mL and 10 kDa, along with a buffer solution of 100 mM sodium acetate, 300 mM of NaCl and 5 mM of EDTA pH 5.5 and 1 mM of β-mercaptoethanol (Sigma-Aldrich). The solution was centrifuged (Eppendorf 5804R) at 4500 rcf at 4 °C for 35 minutes. Afterward, the solution was transferred to a micro centrifuge tube and kept on ice. The Cz concentration was measured in mg mL^{-1} using a DeNovixDS-11+ spectrometer; the measurement was repeated three times.

The samples for the reversibility assay were prepared with a final volume of 1000 μL. The sample solution was prepared with a Cz concentration of 30 μM (0.7 mg mL^{-1}), acetate buffer pH 5.5, 0.014 % v/v of Triton X-100, and 5 % of DMSO. The reference cell was prepared with the acetate buffer pH 5.5, 0.014 % v/v of Triton X-100, and 5 % of DMSO. Both samples were degassed for 5 minutes before being added in the equipment cells.

The experiment was performed with sequential scans at 3 atm and with a rate of 2 °C/min. First, the sample was heated from 0 to 90 degrees, then it was cooled to 0 °C and again heated to 90 °C. The experiment was performed in duplicates.

3.2.2. Thermal stability of cruzain with different ligands

The dialysis of the protein was made in the same way described before. The samples were prepared with a final volume of 1000 μL. The sample solution was prepared with a Cz concentration of 30 μM (0.7 mg mL^{-1}), acetate buffer pH 5.5, 0.014 % v/v of Triton X-100, 5 % of DMSO and with a ligand concentration of 120 μM. The reference cell was prepared with the acetate buffer pH 5.5, 0.014 % v/v of Triton X-100, and 5 % of DMSO. Both samples were degassed for 5 minutes before being added in the equipment cells.

The blank was made for all the compounds, where the protein was removed from the sample cell. The experiment was performed in duplicates.

The experiment was performed with a single scan at 3 atm and with a rate of 2 °C/min, where the samples were heated, ranging from 0 °C to 90 °C.

4. RESULTS AND DISCUSSION

4.1. Kinetic Characterization of Dipeptidyl Nitriles

In order to evaluate the activity of the five CPs, the K_M and V_{max} were determined through the Michaelis-Menten equation. In the assays, two different substrates were used, the Z-Phe-Arg-AMC and Z-Leu-Arg-AMC (Figure 12). Both substrates are fluorogenic due to the presence of a coumarin moiety. Through a hydrolysis reaction, the substrate is cleaved by the enzyme which releases the 7-amino-4-methylcoumarin fluorescent group.

Figure 12: Structures of the substrates Z-Phe-Arg-AMC (a) and Z-Leu-Arg-AMC (b) with their structural differences highlighted in red at the P2 position. The guanidine group (pK_a = 12.5) of the arginine, present in both substrates, is protonated under the assay conditions (pH = 5.5).
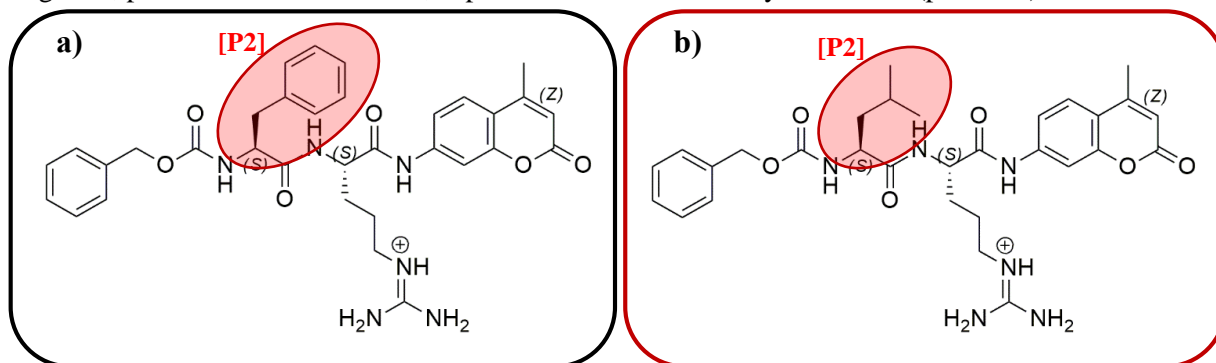
Source: Own author

For Cz, *Lm*CPB, *h*CatL, and *h*CatS the substrate Z-Phe-Arg-AMC was used, while Z-Leu-Arg-AMC was used for *h*CatK. The main reason for the use of the two substrates lays in the differences in the S2 pocket of these CPs. Cz, *Lm*CPB, *h*CatL, and *h*CatS have a deep mainly hydrophobic S2 pocket. So, the Z-Phe-Arg-AMC bearing Phenylalanine in P2 has moderate to high affinity for these CPs. On the other hand, *h*CatK has a smaller S2 pocket and it privileges Leucine or Isoleucine as amino acid in P2.

The protocols also varied in the time and temperature of activation, for Cz, *Lm*CPB, and *h*CatL the activation occurred in an ice bath in an interval of 10 to 20 minutes. *h*CatK and *h*CatS were activated at 37 °C taking about 30 to 60 minutes (FRIZLER et al., 2011).

Two different pHs were used in the assays, one being pH 5.5 and the other 6.0. Because the activation of Cz, *Lm*CPB, *h*CatL and *h*CatK occurs in slightly acidic media, the pH chosen for the assays was 5.5. *h*CatS have an optimum pH between 6.0 and 7.5, so the pH 6.0 was chosen for the assays (TURK et al., 2012).

The Michaelis-Menten plot was obtained for the five CPs having the rate of the reaction (μmol g^{-1} s^{-1}) in the y-axis and concentration (μM) in the x-axis, as shown in Figure 13 for Cz.

Figure 13: Velocity plot as a function of Z-Phe-Arg-AMC concentration for cruzain fit in the Michaelis-Menten equation.

$K_M = 2.18 \pm 0.17 (\mu M)^a$

$V_{max} = 7.90 \pm 0.18^a$

pseudo-$R^2 = 0.98$

p-value = 0.13

Sy.x = 0.33
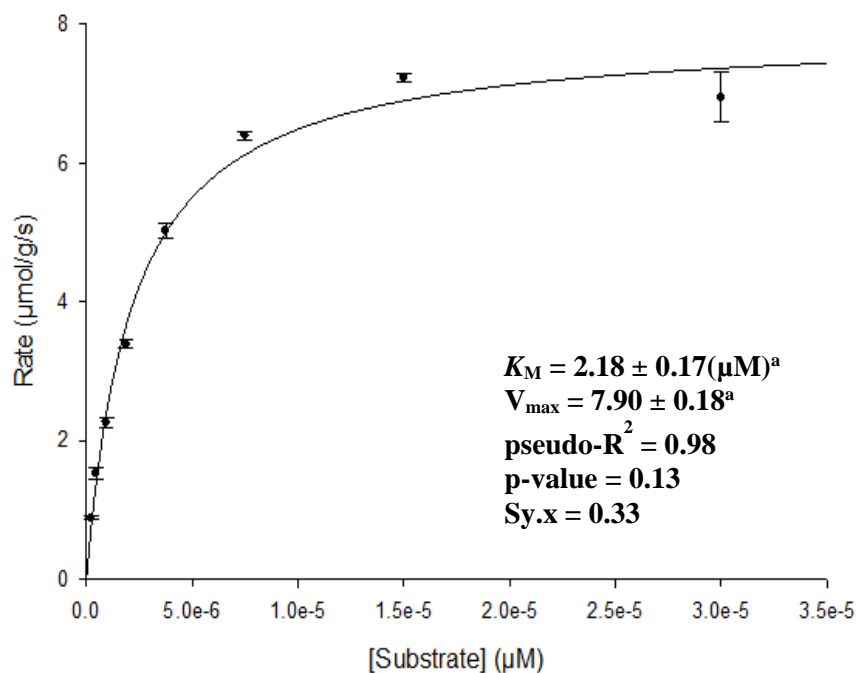
[a] Obtained using $v = \dfrac{V_{max}\,[S]}{K_M + [S]}$

The K_M and V_{max} obtained for all the CPs are shown in Table 1, every assay was done in triplicate and at least twice. The plots can be seen in Appendix A.

Table 1: K_M and V_{max} results for the five CPs obtained through the Michaelis-Menten equation. The experimental K_M results showed at BRENDA[a] website for Cz is 1 to 3 μM; LmCPB 5 to 17 μM; hCatL = 0.7 to 6 μM; hCatS = 14 to 42 μM; hCatK is 1.5 to 4.6 μM.

Enzyme	K_M (μM)	V_{max} (μmol g^{-1} s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
Cruzain [b]	1.8 ± 0.2	7.4 ± 0.6	0.2 ± 0.01	11.1x10^4
LmCPB [b]	6.7 ± 0.9	81 ± 8.5	2.1 ± 0.22	31.3x10^4
hCatL [c]	3.3 ± 0.2	11 ± 0.8	0.3 ± 0.02	9.1x10^4
hCatS [d]	37 ± 0.5	4.6 ± 1.5	0.1 ± 0.04	0.3x10^4
hCatK [d]	1.7 ± 0.07	2.2 ± 0.7	0.06 ± 0.01	3.5x10^4

[a] https://www.brenda-enzymes.org/
[b] Results obtained through the average of four measurement
[c] Results obtained through the average of three measurement
[d] Results obtained through the average of two measurement

From the K_M and V_{max}, it is possible to calculate the k_{cat} and the k_{cat}/K_M ratio. The less efficient enzyme in the assay conditions is hCatS. This enzyme has a high K_M value, which indicates a lower affinity for the substrate. The turnover number (k_{cat}) and the ratio k_{cat}/K_M for hCatS were the lowest among the enzymes, indicating the low efficiency in hydrolyzing the substrate. Indeed, hCatS has a preference for branched hydrophobic residues in the S2 subsite (CHOE et al., 2006). As for the other CPs, k_{cat} and k_{cat}/K_M exhibit satisfying results with high conversion rates of the substrates.

Having obtained the K_M values for the five enzymes, the inhibition assays were performed with a wide range of covalent reversible inhibitors. In the inhibition assays, the compounds were incubated with the enzyme in which the substrate solution is subsequently added. The substrate will try to displace the inhibitor from the active site of the enzyme by a displacement mechanism. As soon as the inhibitor is displaced, the substrate will be hydrolyzed by the enzyme, and the fluorescence signal will be observed. The higher the inhibitor affinity for the enzyme, the lower the observed fluorescence signal. The fluorescence signal has units of velocity (RFU s^{-1}) and was plotted against different inhibitors concentrations (mol L^{-1}) as shown in Figure 14 for the compound **Neq0570**. The inhibition constant K_i was obtained from equation 2 through a nonlinear fitting and was then converted to pK_i (- $\log_{10} = K_i$). The higher the pK_i, the greater the affinity of the compound for the target.

Figure 14: Inhibition curve obtained for the prototypic compound **Neq0570** against Cruzain in the presence of the substrate Z-Phe-Arg-AMC.

$K_i^{app} = 0.404 \pm 0.006$ (μM)

$K_i = 0.202 \pm 0.003$ (μM)

p$K_i = 6.7 \pm 0.01$

$R^2 = 0.99$
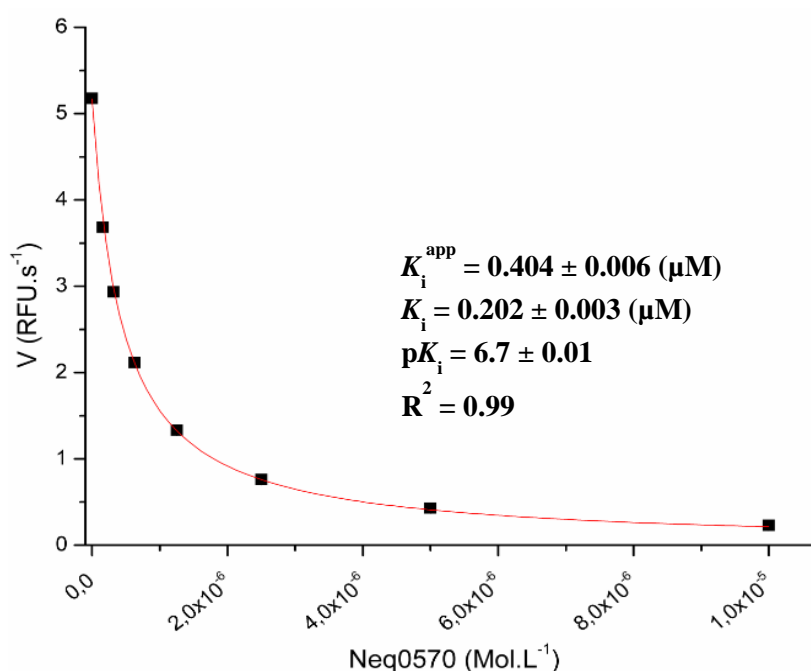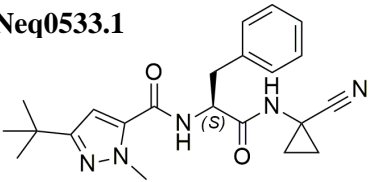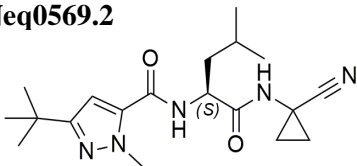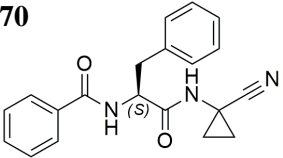
Twenty-two compounds were selected for the inhibition assays with the five CP in which their pK_i values and structures are presented in Table 2. Compounds bearing a pyrazole group in the P3 position were chosen to evaluate the influence of this group in the bimolecular recognition process. The trifluoro ethylamine group in P3/P2 was also chosen, as they are derivatives of Odanacatib and BaCz007. Compounds with a sulfonamide in P3 were selected to evaluate if this class of compounds is an acceptable bioisostere to replace the amide bond in P3/P2. Lastly, pairs of compounds with different structural groups in P1 were selected to map the S1 and S1' subsites, since little is known about these interactions.
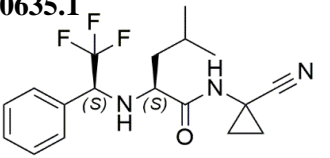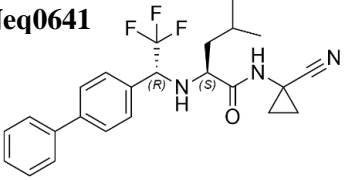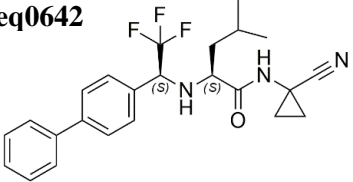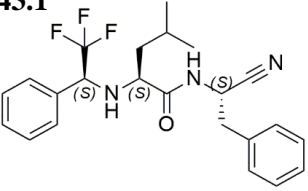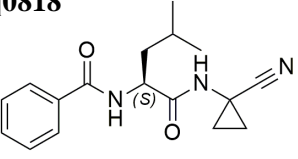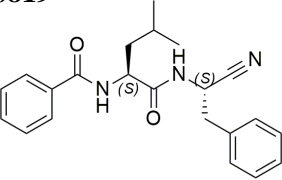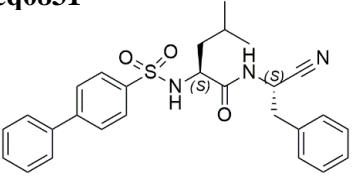
In addition to the 22 compounds selected, a total of 130 compounds were tested against Cz, 110 against *Lm*CPB and 82 against *h*Cat-L. These inhibition studies were perform in collaboration with other members of the NEQUIMED/IQSC/USP group.
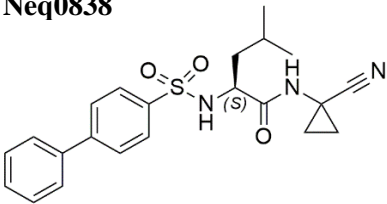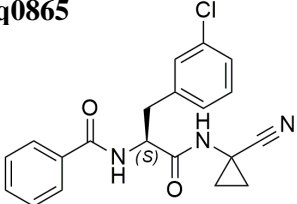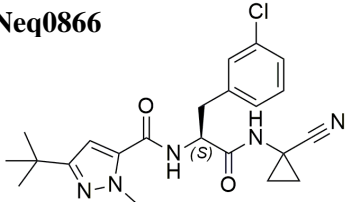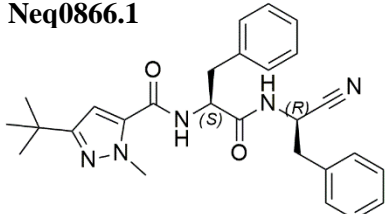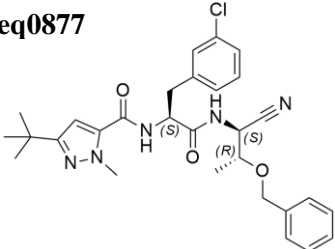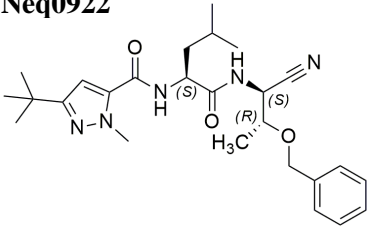
4.2. MMP/SAR analysis of a series of dipeptidyl nitriles
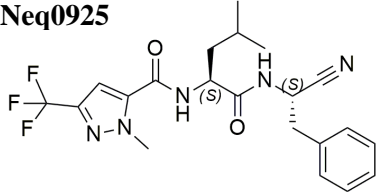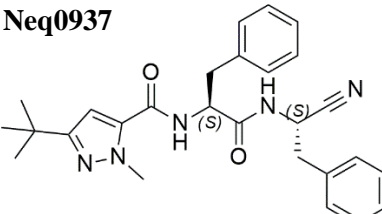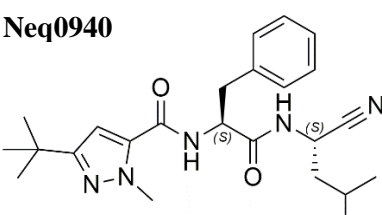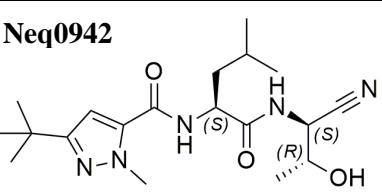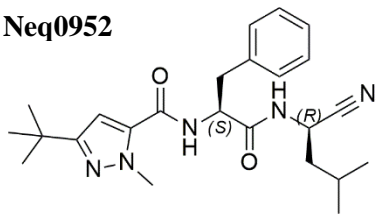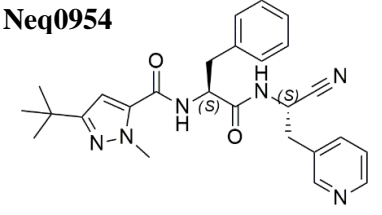
The chosen compounds are dipeptidyl nitrile that acts as reversible covalent inhibitors and displays a competitive mechanism of inhibition (AVELAR et al., 2015).

Table 2: Inhibitors chosen for the SAR and bio calorimetric studies with the pK_i values for the five CPs, as well as their molecular structures. Protozoa CPs are black colored and mammalian blue-colored.

Structure	pK_i (Cz)	pK_i (*Lm*CPB)	pK_i (*h*CatL)	pK_i (*h*CatK)	pK_i (*h*CatS)
Neq0533.1	7.3 ± 0.01	7.1 ± 0.01	8.2 ± 0.01[a]	6.5 ± 0.03[a]	6.8 ± 0.07[a]
Neq0569.2	7.8 ± 0.01	7.3 ± 0.04	7.6 ± 0.01[a]	8.3 ± 0.2[a]	7.4 ± 0.03[a]
Neq0570	6.7 ± 0.01	6.6 ± 0.01	7.4 ± 0.03	6.4 ± 0.02	7.3 ± 0.01

Compound	Structure					
Neq0635.1		7.3 ± 0.02	6.9 ± 0.02	**7.4 ± 0.05**	**6.4 ± 0.02**	**7.6 ± 0.02**
Neq0641		7.3 ± 0.01	7.3 ± 0.03	**6.6 ± 0.03**	**7.1 ± 0.01**	**7.2 ± 0.05**
Neq0642		8.1 ± 0.01	7.7 ± 0.05	**6.1 ± 0.03**	**7.9 ± 0.03**	**6.9 ± 0.03**
Neq0643.1		8.7 ± 0.03	8.6 ± 0.02	**8.8 ± 0.03**	**8.6 ± 0.01**	**8.7 ± 0.06**
Neq0818		7.3 ± 0.01	6.8 ± 0.05	**7.7 ± 0.02**	**7.2 ± 0.04**	**8.0 ± 0.03**
Neq0819		7.4 ± 0.03	7.1 ± 0.02	**7.4 ±0.05**	**7.1 ± 0.02**	**7.4 ± 0.03**
Neq0831		7.0 ± 0.07	7.0 ± 0.06	**6.4 ± 0.02**	**6.8 ± 0.03**	**7.4 ± 0.03**

Compound	Structure					
Neq0838		6.8 ± 0.02	6.8 ± 0.02	**6.5 ± 0.01**	**6.6 ± 0.07**	**7.0 ± 0.02**
Neq0865		6.9 ± 0.01	6.2 ± 0.02	**7.3 ± 0.03[a]**	**6.2 ± 0.07[a]**	**7.7 ± 0.1[a]**
Neq0866		7.4 ± 0.04	6.7 ± 0.01	**8.6 ± 0.05[a]**	**6.4 ± 0.07[a]**	**6.7 ± 0.1[a]**
Neq0866.1		6.3 ± 0.03	6.1 ± 0.03	**6.9 ± 0.01[a]**	**5.6 ± 0.04[a]**	**5.9 ± 0.05 [a]**
Neq0877		7.2 ± 0.05	6.8 ± 0.04	**8.1 ± 0.04[a]**	**6.0 ± 0.03[a]**	**6.5 ± 0.05 [a]**
Neq0922		7.9 ± 0.01	7.7 ± 0.01	**7.2 ± 0.03[a]**	**7.8 ± 0.02[a]**	**7.3 ± 0.01 [a]**

Neq0925		7.4 ± 0.03	7.3 ± 0.01	**7.6 ± 0.07**	**7.0 ± 0.03**	**7.4 ± 0.02**
Neq0937		7.7 ± 0.02	7.8 ± 0.01	**8.5 ± 0.03[a]**	**6.3 ± 0.05[a]**	**7.3 ± 0.04 [a]**
Neq0940		7.5 ± 0.06	7.4 ± 0.02	**8.3 ± 0.02[a]**	**6.7 ±0.04 [a]**	**6.9 ± 0.08 [a]**
Neq0942		7.1 ± 0.04	6.7 ± 0.01	**7.0 ± 0.03[a]**	**7.8 ± 0.02[a]**	**5.0 ± 0.1 [a]**
Neq0952		6.5 ± 0.02	6.4 ± 0.03	**7.0 ± 0.01[a]**	**6.0 ± 0.04[a]**	**6.0 ± 0.05 [a]**
Neq0954		7.5 ± 0.02	7.2 ± 0.01	**8.3 ± 0.01[a]**	**6.5 ± 0.04[a]**	**7.2 ± 0.06 [a]**

[a] Kinetics assays performed at the University of Bonn (Germany). The same protocol was used for the determination of the pK_is.

The MMP/SAR analysis considered the differences in affinity related to a structural change, defined as ΔpK_i. The ΔpK_i is calculated through the differences in affinity between two molecules (e.g., ΔpK_i = pK_i [**Neq0569.2**] − pK_i [**Neq0533.1**]). ΔpK_i values lower than 0.2 log

units were considered not significant for affinity. For selectivity, a ΔpK_i equal or higher than 1.0 log units was considered significant. Therefore, 8 compounds with different substituents at the P1, P2, and P3 positions and different stereochemistry were evaluated in Figure 15.

Figure 15: MMPA of **Neq0533.1** with seven compounds with ΔpK_i values corresponding to structural modifications targeting five enzymes. Positive modifications are shown in green, negative in red, and when no significant difference was observed in grey. Modifications in P1 are highlighted as orange, P2 blue and P3 red.
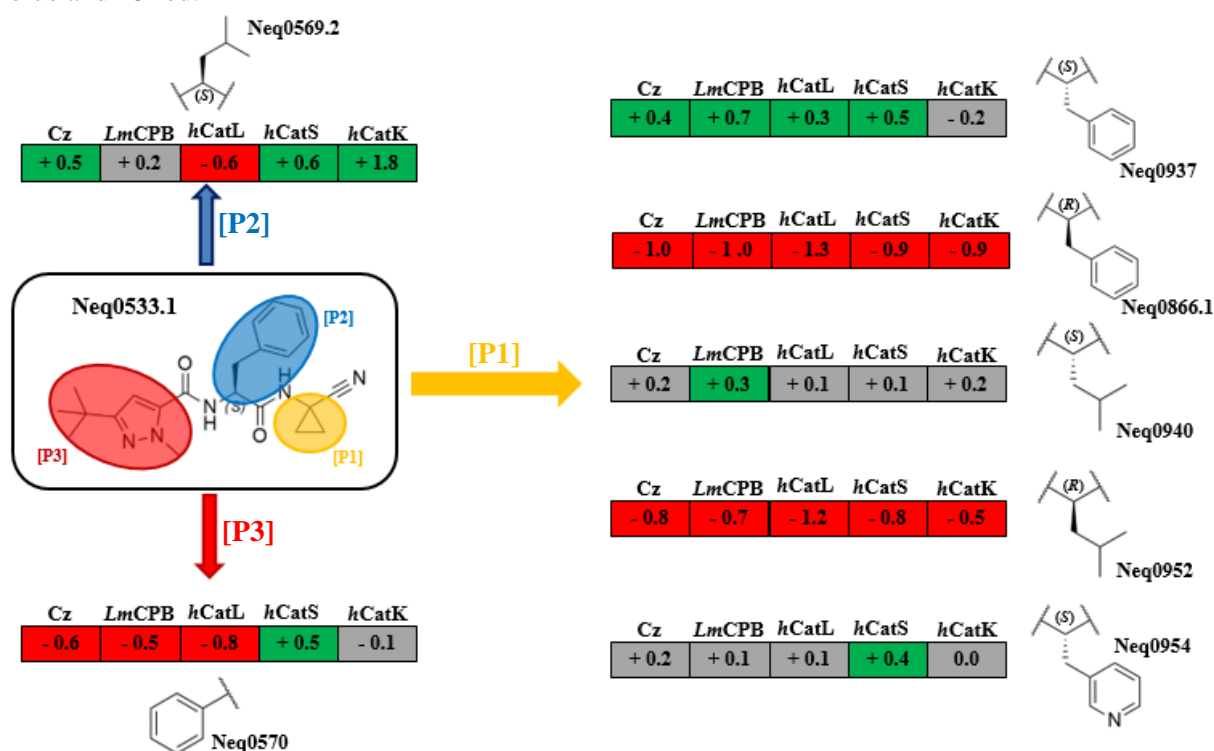
The structural modification at the P3 position changing the 3-(*tert*-butyl)-1-methyl-1*H*-pyrazole (**Neq0533.1**) for the phenyl group (**Neq0570**) led to a loss in affinity for almost all the CPs, except for *h*CatS. Changing the 3-(*tert*-butyl)-1-methyl-1*H*-pyrazole for the phenyl brought a significant decrease for *h*CatL affinity due to the methyl in the 3-(*tert*-butyl)-1-methyl-1*H*-pyrazole group, which is able to perform van der Waals interactions with the Leu69 residue in the S3 pocket of *h*CatL (ASAAD et al., 2009). As for *h*CatS, the Phe70 residue in the S3 pocket can possibly perform π-π stacking interactions with the phenyl ring of the inhibitor without having enough space to accommodate the bulkier 3-*tert*-butyl-N-methyl-pyrazole group (JADHAV et al., 2014).

Modification in P2 changing a Phe (**Neq0533.1**) for a Leu (**Neq0569.2**) led to a gain in the affinity for four of the CPs, with the highest change observed for *h*CatK (+1.8 log units).

*h*CatL was affected as expected due to the preference of this enzyme for aromatic residues in the P2 position.

Structural modifications in P1 take into account five compounds. Not surprisingly, the *(S)* configuration of the moiety in P1 was favored rather than the *(R)* configuration, as also observed in a recent study for amino acids in P2 (AVELAR et al., 2015). The substitution of the cyclopropane to the (*S*)-benzyl (**Neq0937**) raised the affinity for all CPs. However, when the *(R)* configuration (**Neq0866.1**) was analyzed, a considerable loss in affinity occurred. The exchange of the cyclopropane for the *iso*-propyl (**Neq0940**) did not lead to any particular improvement.

The modification from the cyclopropane to the pyridine group (**Neq0954**), maintains the affinity for all the CPs. Despite no significant change, pyridine is a fascinating group in the medicinal chemistry field. The pyridine raises the solubility of the inhibitor without modifying the main hydrophobic interaction with the target protein.

Other structural modifications at the P1 position were investigated among pairs of compounds, as shown in Figure 16.

Figure 16: MMPA with ΔpK_i values corresponding to structural modifications in the P1 position targeting five enzymes. Positive modifications are shown in green, negative in red, and when no significant difference was observed in grey. Modifications in P1 are marked as orange.

Cz	LmCPB	hCatL	hCatS	hCatK
+ 0.8	+ 1.0	+ 0.2	+ 2.3	0.0

Cz	LmCPB	hCatL	hCatS	hCatK
+ 0.2	- 0.1	+ 0.5	+ 0.2	+ 0.4

Cz	LmCPB	hCatL	hCatS	hCatK
+ 0.2	+ 0.2	- 0.1	+ 0.4	+ 0.2

Cz	LmCPB	hCatL	hCatS	hCatK
+ 1.4	+ 1.7	+ 1.4	+ 1.1	+ 2.2
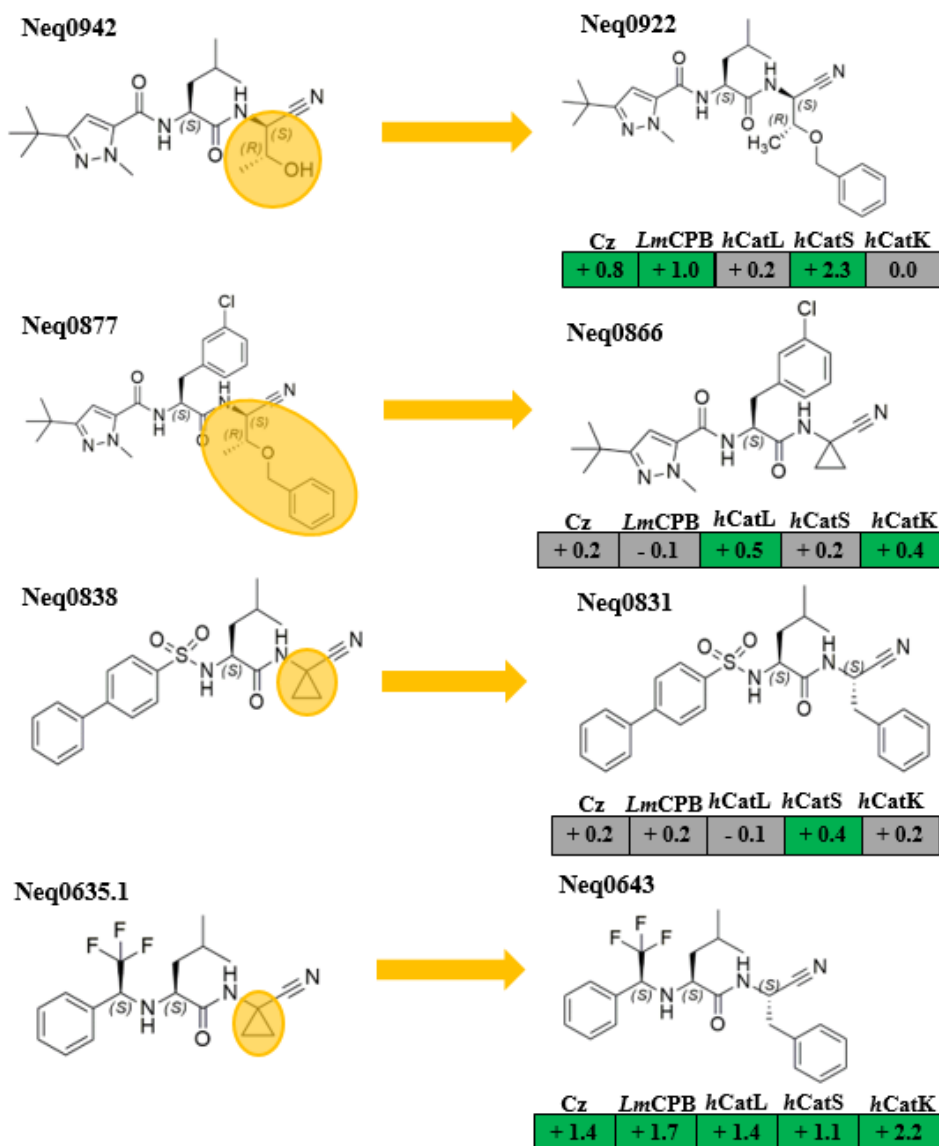
The modification in P1 for **Neq0942** to **Neq0922** brought significant changes in affinity for Cz, *Lm*CPB, and *h*CatS, especially for *h*CatS, for the latter we observed an activity cliff with an improvement in affinity over one hundredfold. *h*CatS, Cz, and *Lm*CPB may not tolerate hydrogen bond donors as the hydroxyl group in the threonine, while showing a preference for hydrogen bond acceptors and large hydrophobic moiety as the protected benzyl threonine (Thr-O-Bn) in **Neq0922**. It is important to point out that the Thr-O-Bn moiety is interacting with the S1' subsite of the CPs and not the S1. The Thr-O-Bn is a large and flexible group and thus have a different mode of interaction leading to the S1' pocket.

For **Neq0877** and **Neq0866,** no significant change in affinity occurred for the five CPs when the Thr-O-Bn was changed to the smallest cyclopropane moiety. Probably the loss of S1'-P1 interaction is counterbalanced with the higher electrophilic character of the carbon on the nitrile due to its interactions with the cyclopropane. In a medicinal chemist point of view, the cyclopropane group contains other important features as metabolic stability, small size, and absence of stereogenic centers (TALELE, 2016).

The effects described above were detected as well for the sulfonamides **Neq0838** and **Neq0831**, in which the substitution of the cyclopropane group to a benzyl group led to no meaningful differences in affinity.

The greatest improvement in affinity was observed when the non-peptidic pairs **Neq0635.1** and **Neq0643.1** were analyzed. The modification of a cyclopropane to a benzyl group led to over one log unit increase in affinity. An activity cliff was observed for *h*CatK, with over one hundredfold of gain in the compounds' affinity. For this reason, the inhibitors will be further analyzed for their chemical properties through a SAR cycle (Figure 20).

Some modifications in the P2 and P3 positions were also analyzed between pairs, as depicted in Figure 17.

Figure 17: MMPA with ΔpK_i values corresponding to structural modifications in P2 and P3 targeting five enzymes. Positive modifications are shown in green, negative in red, and when no significant difference was observed in grey. Modifications in P2 are marked as blue and P3 as red.
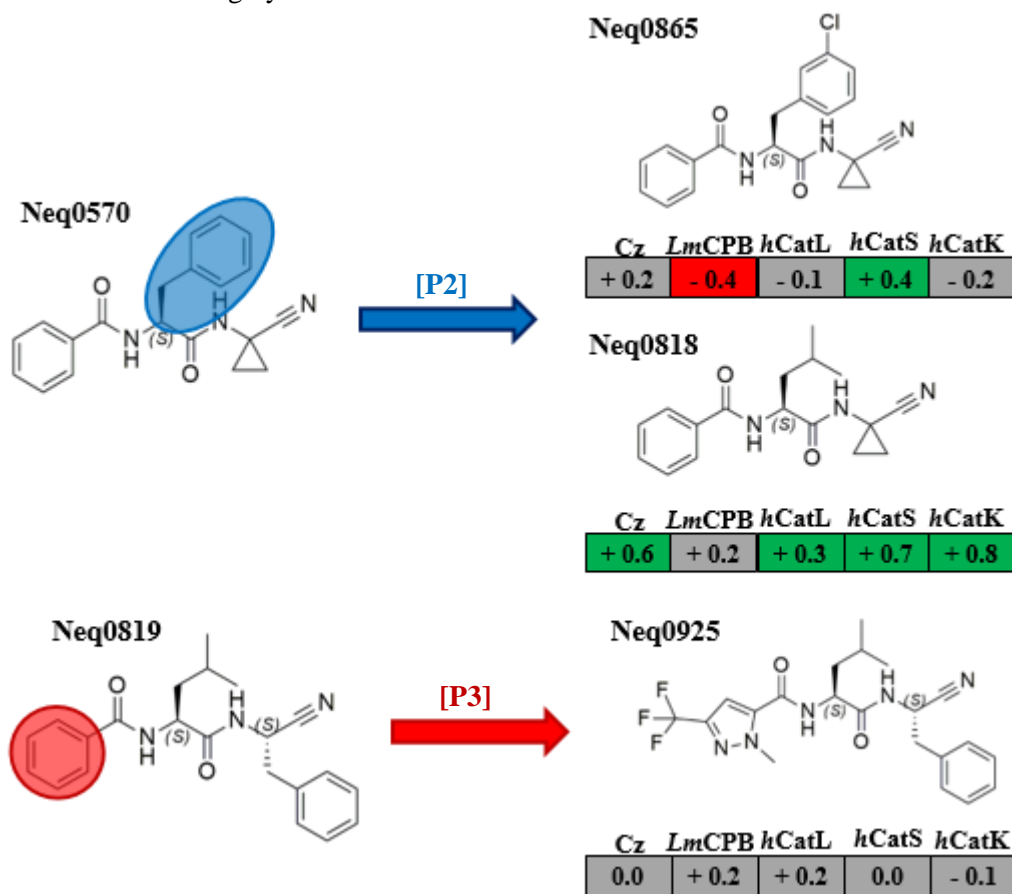
The transformation at P2 changing the phenyl ring for the 3-chlorophenyl leads to no significant improvement. When changing the *(S)*-Phe for the *(S)*-Leu group a slight increase in affinity occurred for all the five CPs. *h*CatK was the enzyme in which the modification affects more positively due to the preference for a small group in that position.

When taking into account the P3 modifications, the use of the 1-methyl-3-(trifluoromethyl)-1*H*-pyrazole group (**Neq0925**) has brought no significant effect in affinity. This implies that the substitution of the *tert*-butyl for the CF_3 in the 1-methyl-1*H*-pyrazole structure influenced negatively the affinity. It was possible to assume that the trifluoromethyl moiety, due to its high polarity, forced the orientation of the pyrazole in the P3 pocket blocking hydrophobic interactions with the Leu69. It is important to point out that in all cases when the 3-(*tert*-butyl)-1-methyl-1*H*-pyrazole was used a gain in the inhibitors' affinity was observed.

Stereochemistry differences in pairs, for the P1 and P3 position were evaluated as shown in Figure 18.

Figure 18: Analysis of different stereochemistry in MMP with ΔpK_i values corresponding to structural modifications targeting five enzymes. Positive modifications are shown in green, negative in red. Modifications in P1 are marked as orange and P2/P3 as blue.

Cz	LmCPB	hCatL	hCatS	hCatK
+ 1.0	+ 1.0	+ 1.3	+ 0.9	+ 0.7

Cz	LmCPB	hCatL	hCatS	hCatK
+ 0.8	+ 0.4	- 0.5	- 0.3	+ 0.8

Cz	LmCPB	hCatL	hCatS	hCatK
+ 1.4	+ 1.7	+ 1.6	+ 1.4	+ 0.7
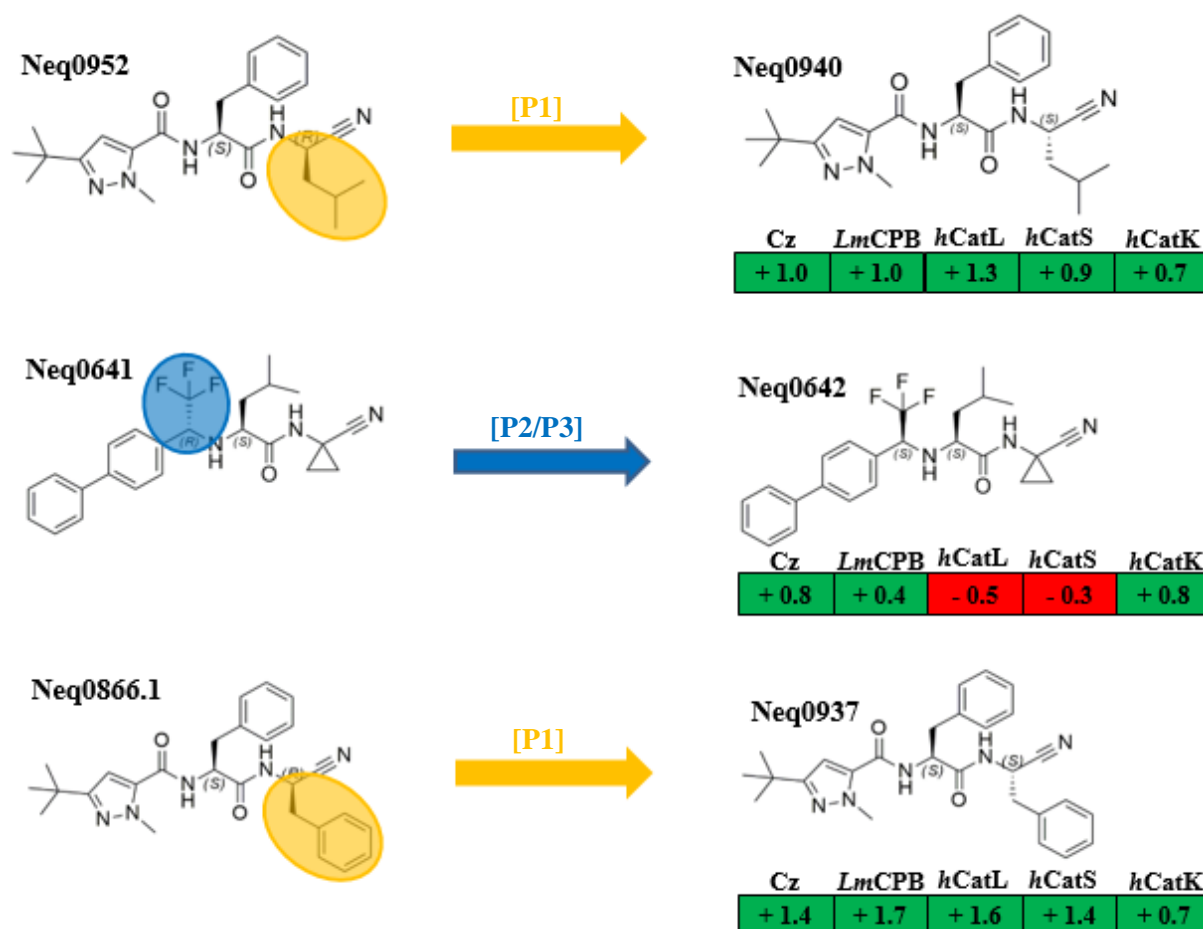
As discussed earlier, the transformations at P1 lead to a gain of almost 1.0 logarithmic unit or more, for the 5 enzymes, thereby showing a strong preference for the S configuration. When analyzing the P2/P3 position, specifically for the trifluoro ethylamine group configuration, the enzymes did not suffer significant changes in affinity regard of the inversion of configuration. These results give the opportunity to use both configurations without greater loss in affinity for CPs. The CF_3 group is an interesting group in medicinal chemistry because it can mimic the C=O of the peptide bond and bring metabolic stability to the molecule. The sp_3 hybridization of the atoms in the CF_3 group allows better orientation of the molecule in the enzyme's active site, improving interactions with the target and thus improving affinity (SANI; VOLONTERIO; ZANDA, 2007).

A double transformation cycle was made aiming to identify structural changes that leads to a gain in the compound's affinity by analyzing the additivity in the SAR (Figure 19). Additivity can be evaluated when comparing a ΔpK_i value related to double structural

transformations of a pair with the sum of the ΔpK_i values resulted from the pairs of individual transformations. The additivity analysis considers all the processes that occur when a structural modification is made in a molecule, such as changes in the types of molecular interactions and their physicochemical characteristics.

The effects of the structural transformations can be additive, superadditive or subadditive. A superadditive effect occurs when the ΔpK_i related to more than one structural transformation is higher than the sum of the pairs ΔpK_i. A subadditive effect occurs when the sum of the affinities of individual structural transformations are greater than the affinity of the two structural transformations. The additive effect occurs when the individual structural transformations are equal to the affinity observed for the two structural transformations.

Figure 19: Nonadditivity in SAR analysis of four compounds where the ΔpK_i values shown are related to *hCatK* inhibition. The green arrow represents positive changes in affinity and the grey arrow no significant change. Blue circles indicate a structural modification in P2 and brown circle in P3.
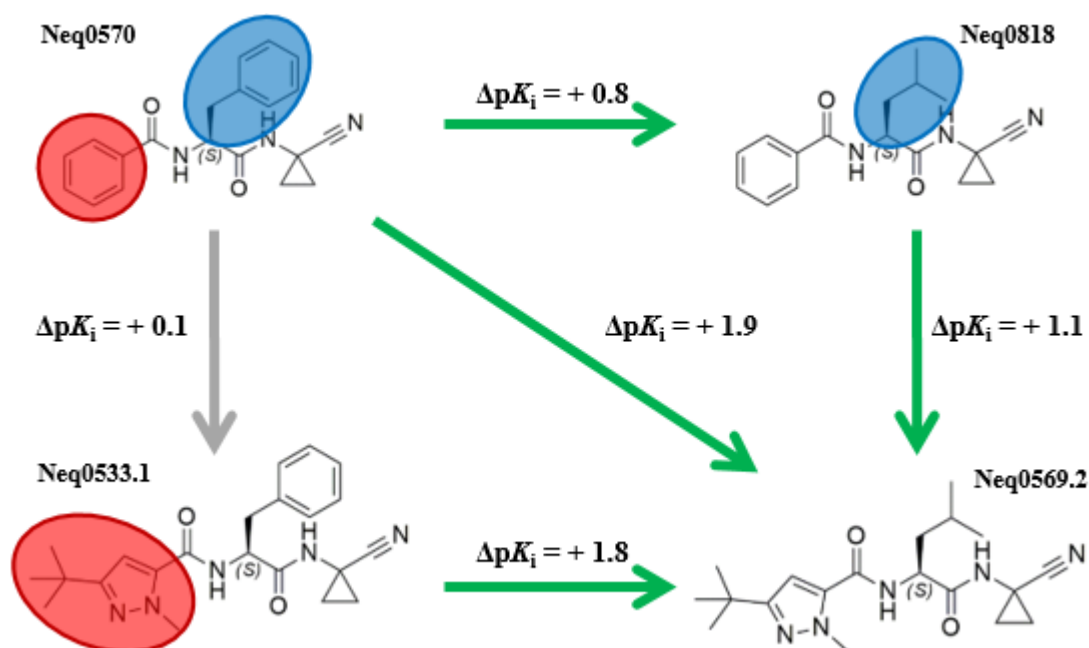
Table 3 shows the ΔpK_i for the five CPs related to the figure 19 cycle.

Table 3: ΔpK_i for the four CPs from the SAR analysis of Figure 19, the notation [X → Y] indicates a structural modification. Diagonal transformation is highlighted in orange

Structural transformation	Cz	*Lm*CPB	*h*CatL	*h*CatK	*h*CatS
[Neq0570 → Neq0818]	+ 0.6	+ 0.2	+ 0.3	+ 0.8	+ 0.7
[Neq0570 → Neq0533.1]	+ 0.6	+ 0.5	+ 0.8	+ 0.1	- 0.5
[Neq0533.1 → Neq0569.2]	+ 0.5	+ 0.2	- 0.6	+ 1.8	+ 0.6
[Neq0818 → Neq0569.2]	+ 0.5	+ 0.5	- 0.1	+ 1.1	- 0.6

[Neq0570 → Neq0569.2]	+ 1.1	+ 0.7	+ 0.2	+ 1.9	+ 0.1

Strong superadditivity was observed for hCatK in which the sum of the ΔpK_i of the transformations [**Neq0570 → Neq0818**] and [**Neq0570 → Neq0533.1**] was equal to + 0.9 and the double structural modification to a ΔpK_i of + 1.9. The obtained results demonstrated the greater influence that substituents at the P2 position have for hCatK. The structural transformations between [**Neq0570 → Neq0818**] and [**Neq0533.1 → Neq0569.2**] displayed the largest gain in affinity. This gain is related to the structural change of a Phe for the Leu group in P2 which fits better in the S2 pocket of the enzyme. In addition, the structural changes in the P3 position demonstrate that the 3-(*tert*-butyl)-1-methyl-1*H*-pyrazole has a positive effect in affinity for hCatK when combined with the Leu in P2. The 3-*(tert*-butyl)-1-methyl-1*H*-pyrazole can increase the number of interactions between the inhibitor and the target. The S3 pocket of hCatK has a Tyr67 and an Asn61 residue which can form H-bond interactions with the nitrogen atom on the pyrazole ring.

Cz ΔpK_i values for the individual structural transformations [**Neq0570 → Neq0818**] = + 0.6 and [**Neq0570 → Neq0533.1**] = + 0.6 shown in Figure 19 sum up +1.2. The double structural transformations that correspond to [**Neq0570 → Neq0569.2**] has a ΔpK_i of +1.1. As a difference in ΔpK_i of ± 0.2 is considered not meaningful, the results showed that the structural transformations on Cz have an additive effect. The same effect was observed for LmCPB since both ΔpK_i were equal to + 0.7.

hCatL and hCatS displayed effects of subadditivity for the structural transformations reported in Figure 19.

Alongside with the cycle shown in Figure 19, another double transformation cycle was made using some of the compounds in Table 2 (Figure 20). We seek to evaluate if there are additivity effects in the structural transformations between the pairs that lead to **Neq0643.1**.

Figure 20: Nonadditivity in SAR analysis of six compounds where the ΔpK_i values shown are related to cruzain inhibition. The green arrow represents positive changes in affinity, red arrows a negative change in affinity and grey arrow no change in affinity. Orange circles indicate a structural modification in P1, blue circles modifications in P2 and brown circle in P3.
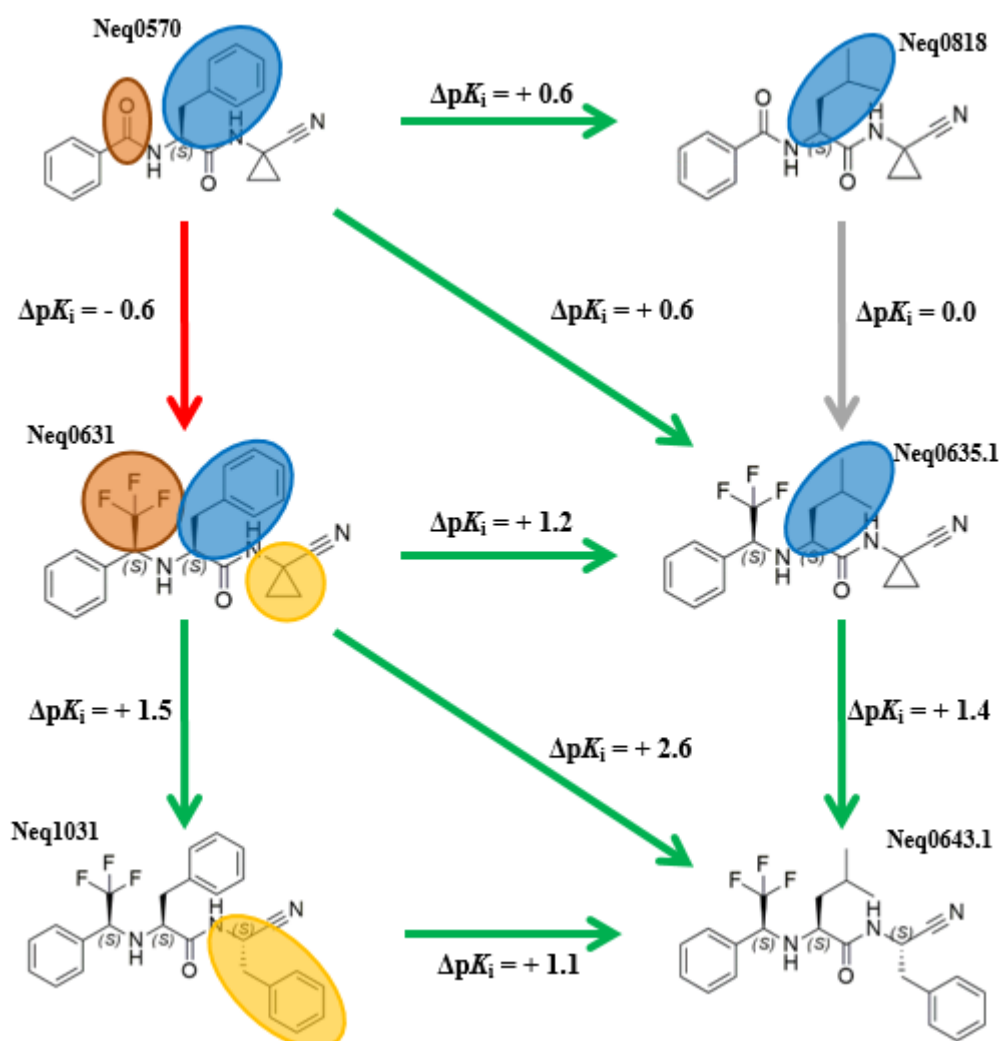
In the Table below, the ΔpK_i for *Lm*CPB, *h*CatL, *h*CatS, and *h*CatK of Figure 20 are presented.

Table 4: ΔpK_i for the four CPs from the SAR analysis of Figure 20, the notation [X → Y] indicates a structural modification. Diagonal transformations are highlighted in orange.

Structural transformation	Cz	*Lm*CPB	*h*CatL	*h*CatS	*h*CatK
[Neq0570 → Neq0818]	+ 0.6	+ 0.2	+ 0.3	+ 0.7	+ 0.8
[Neq0570 → Neq0631]	- 0.6	- 0.8	- 0.6	- 0.1	- 1.0
[Neq0570 → Neq0635.1]	+ 0.6	+ 0.3	0.0	+ 0.3	0.0
[Neq0818 → Neq0635.1]	0.0	+ 0.1	- 0.3	- 0.4	- 0.8
[Neq0631 → Neq0635.1]	+ 1.2	+ 1.1	+ 0.6	+ 0.4	+ 1.6
[Neq0631 → Neq1031]	+ 1.5	+ 1.9	+ 1.7	+ 1.0	+ 1.5
[Neq0631 → Neq0643.1]	+ 2.6	+ 2.8	+ 2.0	+ 1.5	+ 3.2
[Neq0635.1 → Neq0643.1]	+ 1.4	+ 1.7	+ 1.4	+ 1.1	+ 2.2

[Neq1031 → Neq0643.1]	+ 1.1	+ 0.9	+ 0.2	+ 0.5	+ 1.7

The upper part of the cycle consists of a flat SAR because there were no significant gains in affinity. At the same time, all transformations observed were superadditive. The structural transformation related to the ΔpK_i of [**Neq0570 → Neq0635.1**] were below 1.0 log units. The individual transformation of [**Neq0570 → Neq0631**] brought a decrease in the compounds affinities for the five CPs. This loss of almost 1.0 log unit is due to the change in the P3 position, changing the C=O bond to the CF_3 group. In addition, we see that the loss related to the CF_3 group only occurs when the Phe and cyclopropane are at the P2 and P1 position, respectively. The structural modification of [**Neq0631 → Neq0635.1**] with the change of the Phe in the P2 position for the Leu and maintaining the CF_3 group has brought an expressive increase in affinity of almost 100 times. Therefore, the combination of the CF_3 group with the Leu at the P2 position has a positive effect on affinity.

For the lower part of the cycle, which leads to **Neq0643.1** from **Neq0631**, the structural transformations for the five CPs were all additive as the differences in ΔpK_i were between ± 0.1/0.2 log units, considered insignificant.

Opposed to the first cycle, a massive improvement in affinity was obtained in the second cycle. One could justify these phenomena by considering that when the CF_3 is implemented, the mode of binding (MoB) of the inhibitors change, allowing a better accommodation of P1, P2 and P3 moieties in the subsites pockets. Indeed, the CF_3 moiety raises the affinity of the inhibitors of about one hundredfold. Modifications at P1 and P2 displayed high gain in affinity in an additive and subadditive way. These structural modifications lead to the most potent compound of the work, **Neq0643.1**. Table 5 shows the physical-chemical properties for the compounds.

Table 5: Physical-chemical properties of the compounds presented in Figure 20.

Compounds	MW (g Mol^{-1})[a]	H-Bond Donor [a]	H-Bond Acceptor [a]	Log P$_{(o/w)}$ [a]
Neq0570	319.4	2	3	2.2
Neq0631	387.4	2	6	3.6
Neq0818	299.4	2	3	2.0
Neq0635.1	353.4	2	6	3.4
Neq1031	451.5	2	6	4.8
Neq0643.1	417.5	2	6	4.5

[a] Results obtained in the swissadme website, access in 10/07/2019.

As we can see in the Table above, the molecular weight of **Neq0643.1** is bigger than **Neq0570** and **Neq0631**. Compared with **Neq0570**, the number of atoms that can accept

hydrogen bond was folded, being one of the possible reasons for the high affinity observed for the five CPs. The Log $P_{(o/w)}$ was also higher, making the molecule more lipophilic, a very desirable characteristic for bioactive compounds. Nonetheless, these properties will need to be better balanced with cell-based studies.

The five CPs have high affinity for the **Neq0643.1**, with pK_i varying from 8.6 to 8.8. These results are very optimistic for a bioactive compound due to the high capability to inhibit all the CPs. At the same time, has brought a surplus characteristic, that is the lack of selectivity of the molecule.

4.3. Selectivity study amid the five CPs

Selectivity can be a very difficult task to accomplish for CPs inhibition (CIANNI et al., 2019). In this way, an analysis was conducted to evaluate the selectivity of 22 compounds (Table 2), considering their affinities for the five CPs. The results obtained for the comparison between *Lm*CPB x *h*CatL are shown in Figure 21. The rest of the analysis is shown in the Appendix B.

Figure 21: Selectivity analysis obtained for *Lm*CPB over *h*CatL. The compounds in the blue and orange circles have differences in pK_i greater or equal to 1.0 log units.
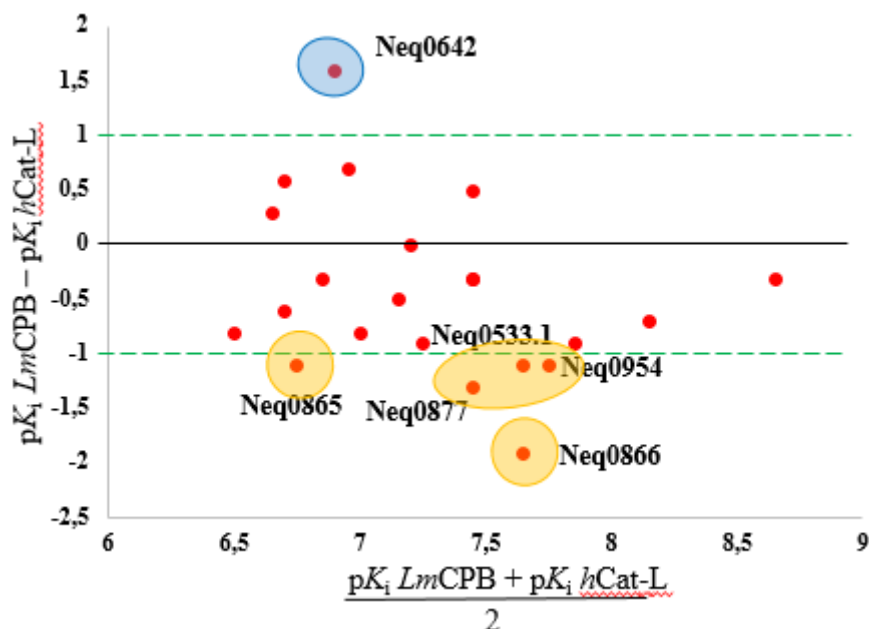
The y-axis shows the differences between the pK_i values for two enzymes and the x-axis the average of the pK_is. Inhibitors for which CPs have high affinity (x-axis) are placed in

the right part of the graph. The highest selectivity is observed for inhibitors that are farthest from zero on the y-axis.

As shown in Figure 21, the compounds are mostly distributed in the negative part of the graph. This indicates a tendency in selectivity for hCatL. Above the + 1.0 line, only **Neq0642** exhibits a selectivity of more than 1.0 log unit for LmCPB. As for hCatL, five compounds displayed high selectivity for this enzyme with over 1.0 log unit of differences in pK_i.

Cz was analyzed (Appendix B) with the cathepsins in which **Neq0642** and **Neq0942** displayed a difference in pK_i of more than 2.0 log units over hCatL and hCatS respectively. Analyzing the selectivity of Cz with hCatK, **Neq0866**, **Neq0877**, **Neq0937** and **Neq0954**, had greater selectivity for Cz.

For hCatL and hCatS, the compounds showed to be more selective for hCatL, with 7 compounds with differences in pK_i greater than 1.0 log unit. As for hCatL and hCatK, the differences are even greater, having ten compounds with over 1.0 log units of differences in pK_i, in which **Neq0866**, **Neq0877,** and **Neq0937** displayed more than 2.0 log unit of selectivity over hCatK.

Finally, the observed selectivity between hCatS and hCatK places most inhibitors close to the 0 axis. Nevertheless, most of them have more preference for hCatS. Both enzymes had two compounds displaying selectivity beyond 1.0 log unit.

In general, mammalian CPs have more affinity for the inhibitors than the protozoan CPs. Cysteine cathepsins showed high affinity for 15 of the 22 inhibitors studied. The hCatL showed better affinity/selectivity for eleven inhibitors in which the most relevant were **Neq0533.1**, **Neq0866.1,** and **Neq0954**, which exhibit similar structures with the 3-(*tert*-butyl)-1-methyl-1*H*-pyrazole in P3 and the Phe residue at P2. The key modifications sit on P1 position, in which these inhibitors carry different moieties.

On the other hand, the most selective inhibitor among the cathepsins was **Neq0942**, displaying a preference for hCatK and hCatL over hCatS of almost 3.0 log units. The most important feature in **Neq0942** structure is the threonine residue at P1, bringing a crucial element to attain selectivity among the mammalian CPs. Nevertheless, **Neq0642** displayed more selectivity for the protozoan CPs, with over 2.0 log units of difference in affinity. This inhibitor has a CF_3 moiety in P3/P2 and a biphenyl at P3, thus, could be explored to achieve selectivity over mammalian CPs.

With the relationships settled, flanked by compound structures and the enzymes, a biocalorimetric study was conducted for Cz with the ligands **Neq0533.1**, **Neq0569.2**, **Neq0570**,

Neq0635.1, **Neq0643.1**, **Neq0818**, **Neq0865**, **Neq0866**, **Neq0922**, **Neq0925**, **Neq0937**, **Neq0940**, **Neq0942**, and **Neq0954**, in which the structures are shown in the table 2.

4.4. Differential scanning calorimetry study

The DSC technique was used to evaluate the thermal stability of Cz with some of the compounds, which are identified in Table 7.

First, a reversibility study of the thermal unfolding of Cz was performed, seeking to identify if the process of unfolding is reversible or irreversible. Therefore, sequential scans were performed in which firstly the native protein solution was heated to 90 °C, cooled and then, heated again. The result obtained is shown in Figure 22.

Figure 22: Reversibility assay performed for Cz. The blue line corresponds to the first heating scan, ranging from 0 to 90 °C, the green line corresponds to the cooling scan and the red line to the second heating process. The peak of the unfolding process is depicted with Tm = 66.4 °C. The scans were done at 3 atm and with a rate of 2 °C per minute.
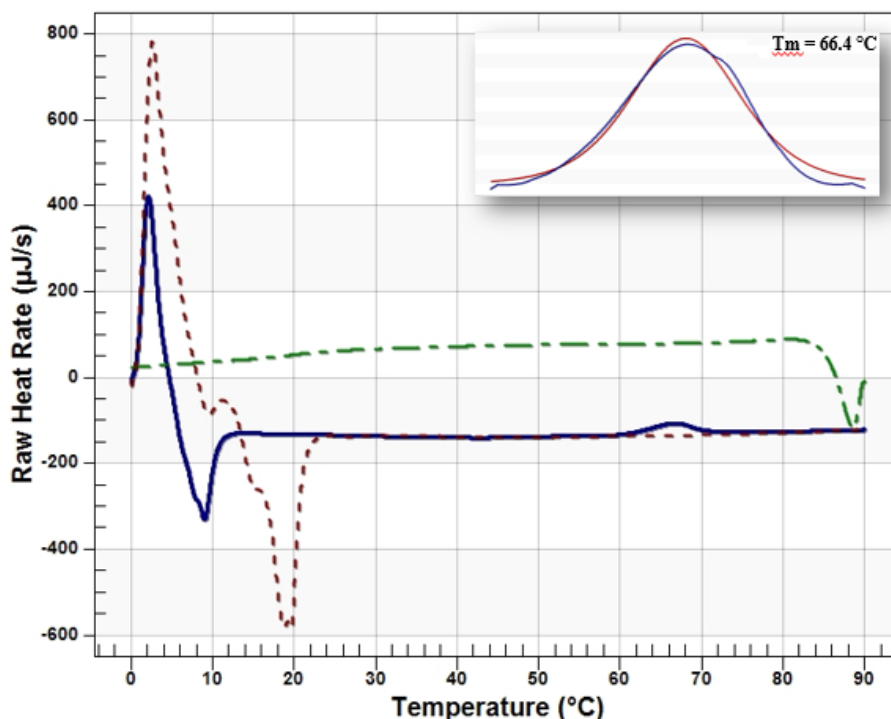
By the sequential scans, it was determined that Cz has an irreversible unfolding due to the non-appearance of a peak in the second heating scan. This result is in agreement with previous studies made for papain and other CPs of this family, in which an irreversible

unfolding was also noticed (SOLÍS-MENDIOLA; ROJO-DOMÍNGUEZ; HERNÁNDEZ-ARANA, 1993).

Having established the irreversibility unfolding of the protein, the assay of the native protein was done to determine its temperature of melting (Tm) (Figure 23).

Figure 23: (a) Raw data before the baseline subtraction where the integration of the area of the peak was performed. (b) Cruzain unfolding peak obtained after the correction of the blank (buffer-buffer), the blue line corresponds to the raw data obtained and the red line is the adjustment of the curve using a Two-State Scaled model. The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min.
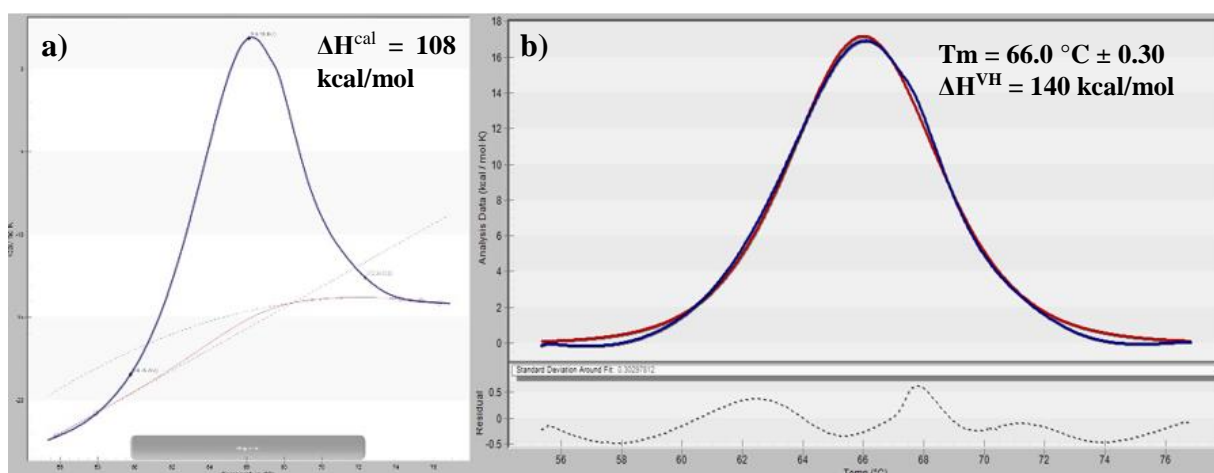
In the figure above, an endothermic peak was observed being characteristic for proteins thermal unfolding. The presence of one peak is an indicative of the protein structure cooperativity. Cz has a two-domain structure and because only one peak appeared in the scan, both domains unfold cooperatively. At Figure 23(a) the peak begins in a lower stage and finishes in a higher one due to the heat capacity (Cp) change of the thermal unfolding of the protein. A positive ΔCp was observed for Cz with a mean value of approximately + 6.8 kcal mol^{-1} K^{-1}. A positive ΔCp indicates that the Cp of the unfolding state is greater than the folded state, that is mainly due to the exposure of hydrophobic amino acids of the protein structure to the solvent molecules (PRABHU; SHARP, 2005).

Calorimetric enthalpy was obtained by integrating the area under the peak before the baseline subtraction (Figure 23a) and the value corresponds to the enthalpy at the Tm when there is an equilibrium between the folded and unfolded forms of the protein. The ΔH^{cal} consists of endothermic and exothermic contributions correlated to the bonds and interactions disrupted by the heating process. The ΔH^{VH} was obtained when the fit of the raw data was done using a

two-state scaled model (Figure 23b), this fit takes into account the equilibrium constant (K_{eq}) considering that both forms of the protein are in perfect equilibrium.

In addition, the ratio $\Delta H^{cal}/ \Delta H^{VH}$ of the enthalpies was obtained aiming to identify if the unfolding process of Cz follows a two-state model, without the formation of intermediates. If ΔH^{cal} is equal to ΔH^{VH} the unfolding process follows a two-state model, but if ΔH^{cal} were greater or lower than ΔH^{VH} the thermal unfolding may occur through the presence of an intermediate (BRUYLANTS; WOUTERS; MICHAUX, 2005).

The Tm of the unfolding process, the calorimetric enthalpy (ΔH^{cal}) and the Van't Hoff enthalpy (ΔH^{VH}), are shown in Table 6, as long as the $\Delta H^{cal}/ \Delta H^{VH}$ ratio.

Table 6: Experimental results of Cz thermal denaturation. All values are an average of four independent experiments.

Tm (°C)	ΔH^{cal} (kcal/mol)	ΔH^{VH} (kcal/mol)	$\Delta H^{cal}/ \Delta H^{VH}$
66.4 ± 0.1	100 ± 5	113 ± 12	0.9

As shown in Table 6, the ratio $\Delta H^{cal}/ \Delta H^{VH}$ is very close to 1.0 and the standard deviation of the analysis indicates that both enthalpies can be equal, and the process of unfolding can be considered of two states with an equilibrium constant equal to 1. So, as the enthalpy values are close, a thermal unfolding without the presence of intermediates was considered for Cz.

Afterward, the analysis proceeded to the investigation of the thermal unfolding of Cz in the presence of different ligands (Table 2). The parameters obtained are shown in the Table below.

Table 7: Experimental results obtained for fourteen compounds when incubated with Cz and analyzed by DSC. The ΔTm was obtained by subtracting the Tm of the incubated protein by the Tm of the native protein. The conditions for the assays were the same used for native cruzain assays with the only difference being the presence of the ligands. The mean values of the two measurements are presented.

Ligand	Tm (°C)	ΔTm (°C)	ΔH^{Cal} (kcal/mol)	ΔH^{VH} (kcal/mol)	$\Delta H^{cal}/ \Delta H^{VH}$
Neq0533.1	76.7	10.5	105 ± 6	189 ± 13	0.5
Neq0569.2	79.3	13.1	117 ± 2	196 ± 7	0.6
Neq0570	72.5	6.1	115 ± 2	133 ± 2	0.8
Neq0635.1	76.4	10.2	107 ± 6	193 ± 12	0.5
Neq0643.1	79.7	13.5	127 ± 15	187 ± 2	0.7
Neq0818	76.9	10.7	128 ± 28	191 ± 12	0.7
Neq0865	75.0	8.8	68 ± 2	190 ± 22	0.4
Neq0866	76.2	10.0	93 ± 2	185 ± 9	0.5
Neq0922	77.1	10.9	125 ± 13	178 ± 11	0.7
Neq0925	76.7	10.5	92 ± 2	190 ± 4	0.5
Neq0937	76.2	10.0	75 ± 2	198 ± 12	0.4
Neq0940	76.8	10.6	88 ± 2	202 ± 12	0.4

Neq0942	75.8	9.60	77 ± 5	212 ± 15	0.4
Neq0954	76.6	10.4	74 ± 3	188 ± 9	0.4

In Table 7 all $\Delta H^{cal} / \Delta H^{VH}$ ratios were lower than 1.0, indicating a shift in the equilibrium of the process. The Tm values are higher than that obtained for the native protein, showing that the ligands are interacting with the folded protein, therefore, stabilizing Cz structure. The positive shift indicates that the equilibrium constant was displaced to the folded form due to the stabilization of the Cz structure. Because of that, when Cz is bound to the ligands the thermal unfolding probably occurred through the presence of an intermediate and the two domains may unfold separately but not independently as only one peak was observed. Figure 24 shows the Gaussian obtained for Cz bound to **Neq0569.2**, one that best stabilizes Cz.

Figure 24: Cruzain unfolding peak when incubated with **Neq0569.2** obtained after the correction of the blank (buffer-buffer), the blue line corresponds to the raw data obtained and the red line is the adjustment of the curve using a Two-State Scaled model. The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min.
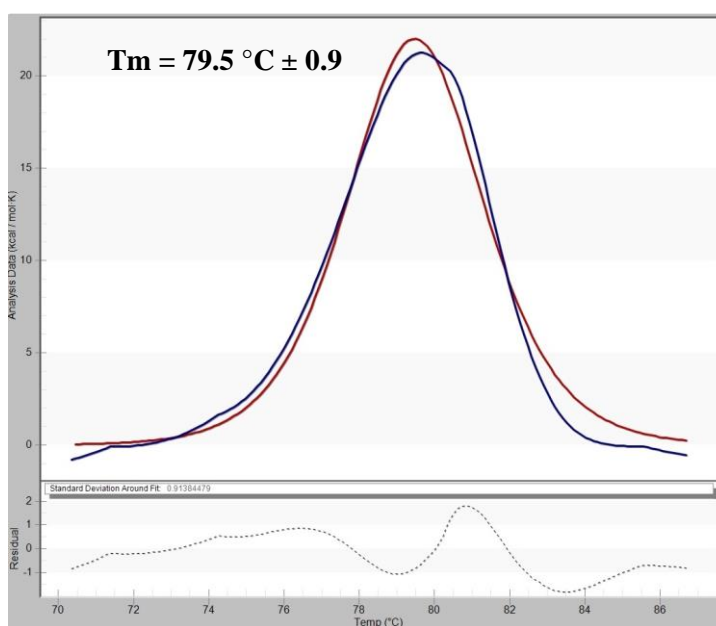
The ΔCp for the unfolding of Cz in the presence of the ligands varied between positive and negative values. A negative ΔCp characterizes polar solvation (PRABHU; SHARP, 2005) that was noticed for compounds that have more atoms capable of accepting hydrogen bonds, being the case for ligands with the CF_3 and pyrazole groups.

A positive temperature shift, like the one obtained in Figure 24, is believed to be proportional to the affinity of a ligand. A plot of pK_i versus ΔTm (Figure 25) was made to identify a possible correlation among the two parameters.

Figure 25: Scatter plot correlating affinity (pK_i) and thermal stability (ΔTm) of Cz with fourteen inhibitors.
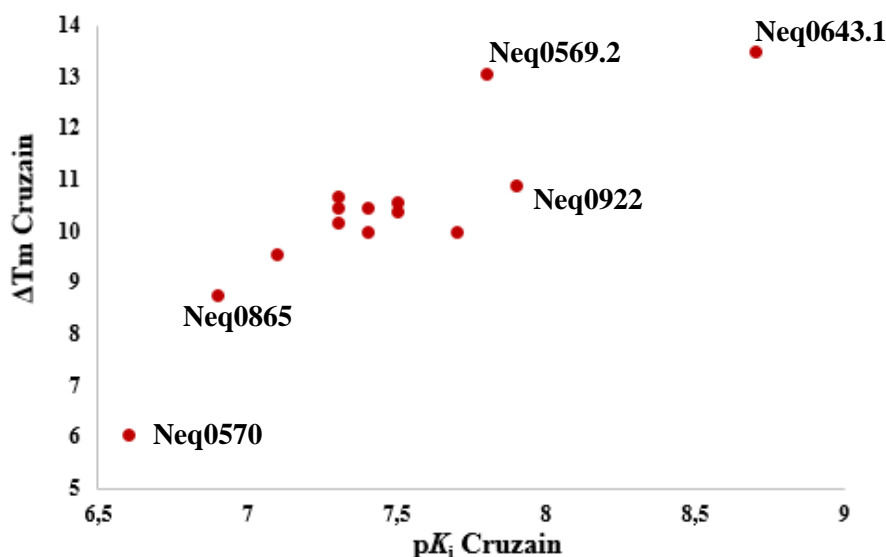
The scattered points suggest a linear trend that, however, should be further investigated. **Neq0570** and **Neq0865** which have the lowest pK_i values, 6.7 and 6.9 respectively, were the only ones that present a ΔTm lower than 10 °C. This may be due to the less quantity of groups/atoms that are H-bond acceptors in these compounds, making the inhibitors to perform fewer interactions with the protein when compared with the rest of the ligands.

Most of the ligands have pK_i between 7.0 and 8.0, because of that, the ΔTm did not suffer significant changes being around 9.5 to 11 °C.

The most potent ligand of the series against Cz, **Neq0643.1** (pK_i = 8.7) was able to thermally stabilized Cz structure in 13.5 °C. This molecule has 6 H-bond acceptors groups/atoms in its structure, different from the other molecules that have between 3 and 4, being **Neq0635.1** the only exception (it has also 6 H-bond acceptors). This may be one of the reasons for its high affinity and capability to stabilize Cz.

Neq0922 (pK_i = 7.9), that has a similar affinity for Cz as **Neq0569.2**, stabilized Cz structure two degrees lower than **Neq0569.2**. Both ligands have the same moieties at the P3 and P2 positions, but differ in the P1. The ligands must have a different pose when bound to the enzyme, as **Neq0922** have a moiety in P1 that goes to the S1' pocket. Furthermore, **Neq0569.2** and **Neq0643.1** have almost 1.0 log unit of differences in pK_i and their structures lack similarity, however, they have almost the same ΔTm. A molecular docking study could be performed to identify the possible reasons for these results, evaluating the pose that these molecules have when in the protein's active site.

5. CONCLUSION

In this work, we have identified covalent reversible inhibitors with high affinity for protozoan and mammalian CPs which are involved in different disease states. A high efficient protocol was implemented to test several reversible inhibitors for a panel of CPs. Structural modifications in P1, P2 and P3 positions in a dipeptidyl nitrile scaffold were performed and analyzed through an MMP/SAR analysis.

The P1 position exhibits a preference for the S stereochemistry. A great loss in the affinity of the inhibitors was observed when the R configuration was employed, independently of the functional group in that position. Overall, the cyclopropane at P1 was more effective than the *(S)*-benzyl, improving the ability of inhibition for the five CPs. The threonine residue at P1 displayed less gain in affinity than the Thr-O-Bn. An improvement in affinity was observed when using the Thr-O-Bn moiety for Cz, *Lm*CPB, and *h*CatS, while no change was noticed for *h*CatL and *h*CatK. The Thr-O-Bn moiety in P1 is an attractive group for targeting *h*CatS selectively since an increase of one hundredfold in affinity was observed for **Neq0922**. The *(S)*-benzyl at P1 showed to be an interesting group when combined with the trifluoro ethylamine moiety at the P3/P2 position. When both moieties were used, a gain of more than 1.0 log unit was observed for the five CPs.

Switching Phe to a Leu on P2, when combined with a phenyl group in P3, an improvement in affinity for three CPs is observed, with a minor effect on *Lm*CPB and *h*CatL. When Leu is on P2 and 3-(*tert*-butyl)-1-methyl-1*H*-pyrazole sitting on P3, a high increase in some CPs affinities is observed, especially for *h*CatK. Nonetheless, there is a decrease in the pK_i value for *h*CatL. Therefore, these combinations can be explored to get high affinity and selective inhibitors for *h*CatK and *h*CatL.

In summary, the 3-(*tert*-butyl)-1-methyl-1*H*-pyrazole is an attractive group at the P3 position for four CPs. For *h*CatS, this moiety at P3 brought no gain in affinity being the phenyl group more compelling for this enzyme. The stereochemistry of the CF_3 group in P3/P2 brought an interesting result, in which both configurations were tolerable. A slight preference for the S configuration was noticed for Cz, *Lm*CPB, and *h*CatK, while the R were preferred for *h*CatL and *h*CatS. Thanks to our structural analysis, we have identified that P3/P2 modifications strongly influence the possible mode of interaction of inhibitors. There is better accommodation on the site and thus a greater capacity for inhibition.

Neq0643.1 is the most potent inhibitor of this work with a pK_i ranging from 8.6 and 8.8 for the five CPs, being a very interesting inhibitor for these enzymes. It is hard to say the reasons for the lack of selectivity observed for **Neq0643.1**.

In general, compounds displayed low selectivity amid the five CPs. The most important inhibitors were **Neq0642**, **Neq0533.1**, **Neq0866.1**, **Neq0954**, and **Neq0942** achieving a difference in affinity of more than 2.0 log units. Additionally, **Neq0642** was the best selective inhibitor for the protozoan CPs. For the mammalian CPs, the structural features of **Neq0533.1**, **Neq0866.1**, and **Neq0954** can be explored to achieve selectivity among the cathepsins, focusing on hCatL.

An irreversible unfolding was observed for Cz along with a two-state unfolding process. Ligands that displayed the highest thermal stabilization of Cz were **Neq0569.2** and **Neq0643.1** with over 13 °C. Other ligands have shown similar ΔTm, ranging from 6.0 to 11 °C. Moreover, a might linear trend was observed between pK_i and ΔTm values. However, further studies should be carried out to better understand these correlations at the molecular level.

Overall, this work has provided information on mammalian and protozoan CP preferences for dipeptidyl nitrile class inhibitors and how to obtain selectivity among them. Hence, with MMP/SAR analysis, we were able to discover and improve the knowledge of cross-class inhibitors for CPs within a number of structural transformations. Our efforts now offer the opportunity to plan new chemical entities (NCEs) directly, without resorting to expensive random testing campaigns.

6. REFERENCES

ÁBRÁNYI-BALOGH, P.; PETRI, L.; IMRE, T.; SCARPINO, A.; HRAST, M.; MITROVIC, A.; FONOVIC, U. P.; NEMETH, K.; BARRETEAU, H.; ROPER, D.I.; HORVATI, K.; FRENCZY, G. G.; KOS, J.; ILAS, J.; GOBEC, S.; KESERU, G. M. A road map for prioritizing warheads for cysteine targeting covalent inhibitors. **European Journal of Medicinal Chemistry**, Issy les Moulineaux Cedex v. 160, p. 94–107, dez. 2018.

ABU-DAYYEH, I.; HASSANI, K.; WESTRA, E. R.; MOTTRAM, J.; OLIVIER, M. Comparative study of the ability of Leishmania mexicana Promastigotes and Amastigotes to alter macrophage signaling and functions. **Infection and Immunity**, Washington, v. 78, n. 6, p. 2438–2445, 1 jun. 2010.

AINSCOUGH, J. S.; MACLEOD, T.; 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**, Washington, v. 114, n. 13, p. E2748–E2757, 28 mar. 2017.

ARAFET, K.; FERRER, S.; MOLINER, V. Computational study of the catalytic mechanism of the cruzain cysteine protease. **ACS Catalysis**, Washington, v. 7, n. 2, p. 1207–1215, 3 fev. 2017.

ASAAD, N.; BETTEL, P. A.; COULSON, M. D.; DAWSON, J. E.; FORD, S.J.; GREHARDT, S.; GRIST, M.; HAMLIN, G. A.; JAMES, M. J.; JONES, E. V.; KAROUTCHI, G. I.; KENNY, P. W.; MORLEY, A. D.; OLDHAM, K.; RANKINE, N.; RYAN, D.; WELLS, S. L.; WOOD, L.; AUGUSTIN, M.; KRAPP. S.; SIMADER, H.; STEINBACHER, S. Dipeptidyl nitrile inhibitors of Cathepsin L. **Bioorganic & Medicinal Chemistry Letters**, Oxford, v. 19, n. 15, p. 4280–4283, ago. 2009.

ATKINSON, H. J.; BABBITT, P. C.; SAJID, M. The global cysteine peptidase landscape in parasites. **Trends in Parasitology**, Oxford, v. 25, n. 12, p. 573–581, dez. 2009.

AVELAR, L. A. A.; CAMILO, C. D.; ALBUQUERQUE, S.; FERNANDES, W. B.; GONÇALEZ, C.; KENNY, P. W.; LEITÃO, A.; MCKERROW, J. H.; MONTANARI, C.A., OROZCO, E. V. M.; RIBEIRO, J. F. R.; ROCHA, J. R.; ROSINI, F.; SAIDEL, M. E. Molecular design, synthesis and trypanocidal activity of dipeptidyl nitriles as cruzain inhibitors. **PLOS Neglected Tropical Diseases**, San Francisco, v. 9, n. 7, p. e0003916, 14 jul. 2015.

BAILLIE, T. A. Targeted covalent inhibitors for drug design. **Angewandte Chemie International Edition**, Weinheim, v. 55, n. 43, p. 13408–13421, 17 out. 2016.

BARBOZA, M.; DUSCHAK, V. G.; FUKUYAMA, Y.; NONAMI, H.; ERRA-BALSELLS, R.; CAZZULO, J. J.; COUTO, A. S. Structural analysis of the N-glycans of the major cysteine proteinase of Trypanosoma cruzi: identification of sulfated high-mannose type oligosaccharides. **FEBS Journal**, Chichester, v. 272, n. 15, p. 3803–3815, 22 jul. 2005.

BEAULIEU, C.; ISABEL, E.; FORTIER, A.; MASSÉ, F.; MELLON, C.; MÉTHOT, N.; NDAO, M.; NICOLL-GRIFFITH, D.; LEE, D.; PARK, H.; BLACK, W. C. Identification of potent and reversible cruzipain inhibitors for the treatment of Chagas disease. **Bioorganic & Medicinal Chemistry Letters**, Oxford, v. 20, n. 24, p. 7444–7449, dez. 2010.

BERTEOTTI, A.; VACONDIO, F.; LODOLA, A.; BASSI, M.; SILVA, C.; CAVALLI, A. Predicting the reactivity of nitrile-carrying compounds with cysteine: a combined computational and experimental study. **ACS Medicinal Chemistry Letters**, Washington, v. 5, n. 5, p. 501–505, 8 maio 2014.

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

BRUYLANTS, G.; WOUTERS, J.; MICHAUX, C. Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design. **Current Medicinal Chemistry**, Sharjah, v. 12, n. 17, p. 2011–2020, 1 ago. 2005.

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, 14 dez. 2018.

CHOE, Y.; LEONETTI, F.; GREENBAUM, D. C.; LECAILLE, F.; BOGYO, M.; BROMME, D.; ELLMAN, J. A.; CRAIK, C. S. Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. **Journal of Biological Chemistry**, Bethesda, v. 281, n. 18, p. 12824–12832, 5 maio 2006.

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

CIANNI, L.; FELDMANN, C. W.; GILBERG, E.; GUTSCHOW, M.; JULIANO, L.; LEITÃO, A.; BAJORATH, J.; MONTANARI, C. A. Can cysteine protease cross-class inhibitors achieve selectivity? **Journal of Medicinal Chemistry**, Washington, [2019?]. No prelo.

CONAGHAN, P. G.; OWES, M. A.; KINGSBURY, S. R.; BRETT, A.; GUILLARD, G.; TUNBLAS, K.; RIZOSKA, B.; LARSSON, T.; HOLMGREN, A.; MANNINEN, A.; GOHLIN, K.; HEBER, W.; GRAHAM, P.; JANSSON, A.; WADELL, C.; BETHELL, R.; OHD, J. Six months' treatment with MIV-711, a novel Cathepsin K inhibitor induces osteoarthritis structure modification: results from a randomized double-blind placebo-controlled phase IIA trial. **Osteoarthritis and Cartilage**, London, v. 26, p. S25–S26, abr. 2018.

DA SILVA, E. B.; DO NASCIMENTO PEREIRA, G. A.; FERREIRA, R. S. Trypanosomal cysteine peptidases: target validation and drug design strategies. *In*: MÜLLER, S.; CERDAN, R.; RADULESCU, O. (ed.). **Comprehensive analysis of parasite biology:** from metabolism to drug discovery. Weinheim: Wiley-VCH, 2016. p. 121–145.

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 Cedex, v. 138, p. 96–114, set. 2017.

DE LUCA, L.; FERRO, S.; BUEMI, M. R.; MONFORTE, A. M.; GITTO, R.; SCHIRMEISTER, T.; MAES, L. RESCIFINA, A.; MICALE, N. Discovery of benzimidazole-based *Leishmania mexicana* cysteine protease CPB2.8ΔCTE inhibitors as potential therapeutics for leishmaniasis. **Chemical Biology & Drug Design**, Hoboken, v. 92, n. 3, p. 1585–1596, set. 2018.

DENNEMÄRKER, J.; LOHMULLER, 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, mar. 2010.

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, jan. 2012.

DOYLE, P. S.; ZHOU, Y. M.; HSIEH, I.; GREENBAUM, D. C.; MCKERROW, J. H.; ENGEL, J. C. The Trypanosoma cruzi protease cruzain mediates immune evasion. **PLoS Pathogens**, San Franscico, v. 7, n. 9, p. e1002139, 1 set. 2011.

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

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

FLEMING, F. F.; YAO, L.; RAVIKUMAR, P. C.; FUNK, L.; SHOOK, B. C. Nitrile-containing pharmaceuticals: efficacious roles of the nitrile pharmacophore. **Journal of Medicinal Chemistry**, Washington, v. 53, n. 22, p. 7902–7917, 25 nov. 2010.

FOTI CUZZOLA, V.; PALLELA, 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, dez. 2012.

FRIZLER, M.; LOHR, F.; FURTMANN, N.; KLAS, J.; GUTSCHOW, M. Structural optimization of azadipeptide nitriles strongly increases association rates and allows the development of selective cathepsin inhibitors. **Journal of Medicinal Chemistry**, Washington, v. 54, n. 1, p. 396–400, 13 jan. 2011.

FUCHS, J. E.; WELLENZOHN, B.; WESKAMP, N.; LIED, K. R. Matched peptides: tuning matched molecular pair analysis for biopharmaceutical applications. **Journal of Chemical Information and Modeling**, Washington, v. 55, n. 11, p. 2315–2323, 23 nov. 2015.

GAULTON, A.; HERSEY, A.; NOWOTKA, M.; BENTO, P.; CHAMBERS, J.; MENDES, D.; MUTOWO, P.; ATKINSON, F.; BELLIS, L. J.; CIBRIAN-UHALTE, E.; DAVIES, M.; DEDMAN, N.; KARLSSON, A.; MAGARINOS, M. P.; OVERINGTON, J. P.; PAPADATOS, G.; SMIT, I.; LEACH, A. R. The ChEMBL database in 2017. **Nucleic Acids Research**, Oxford, v. 45, n. D1, p. D945–D954, 4 jan. 2017.

GILLMOR, S. A.; CRAIK, C. S.; FLETTERICK, R. J. Structural determinants of specificity in the cysteine protease cruzain. **Protein Science**, Hoboken, v. 6, n. 8, p. 1603–1611, ago. 1997.

GRIFFEN, E.; LEACH, A. G.; ROBB, G. R.; WARNER, D. J. Matched molecular pairs as a medicinal chemistry tool: miniperspective. **Journal of Medicinal Chemistry**, Washington, v. 54, n. 22, p. 7739–7750, 24 nov. 2011.

GUHA, R. On Exploring structure–activity relationships. *In*: KORTAGERE, S. (ed.). **In silico models for drug discovery**. Totowa: Humana Press, 2013. v. 993, p. 81–94.

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.

HERRERA-MAYORGA, V.; LARA-RAMIREZ, E. E.; CHACON,VARGAS, K. F.; AGUIRRE-ALVARADO, C.; RODRIGUEZ-PAEZ, L.; ALCANTARA-FARFAN, V,; CORDERO-MARTINEZ, J.; NOGUEDA-TORRES, B.; REYES-ESPINOSA, F.; BOCANERA-GARCIA, V.; RIVERA, G. Structure-based virtual screening and in vitro evaluation of new Trypanosoma cruzi cruzain inhibitors. **International Journal of Molecular Sciences**, Basel, v. 20, n. 7, p. 1742, 9 abr. 2019.

JADHAV, P. K.; SCHIFFLER, M. A.; GAVARDINAS, K.; KIM, E. J.; MATTHEWS, D. P.; STASZAK, M. A.; COFFEY, D. S.; SHAW, B. W.; CASSIDY, K. C.; BRIER, R. A.; ZHANG, Y.; CHRISTIE, R. M.; MATLER, W. F.; QING, K.; DURBIN, J. D.; WANG, Y.; DENG, G. G. Discovery of cathepsin S inhibitor LY3000328 for the treatment of abdominal aortic aneurysm. **ACS Medicinal Chemistry Letters**, Washington, v. 5, n. 10, p. 1138–1142, 9 out. 2014.

KĘDZIOR, M.; SEREDYŃSKI, R.; GUTOWICZ, J. Microbial inhibitors of cysteine proteases. **Medical Microbiology and Immunology**, Heidelberg, v. 205, n. 4, p. 275–296, ago. 2016.

KRAMER, C.; FUCHS, J. E.; WHITEBREAD, S.; GEDECK, P.; LIEDL, K. R. Matched molecular pair analysis: significance and the impact of experimental uncertainty. **Journal of Medicinal Chemistry**, Washington, v. 57, n. 9, p. 3786–3802, 8 maio 2014.

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

LINDSTRÖM, E.; RIZOSKA, B.; HENDERSON, I.; TERELIUS, Y.; JERLING, M.; EDENIUS, C.; GRABOWASKA, U. Nonclinical and clinical pharmacological characterization of the potent and selective cathepsin K inhibitor MIV-711. **Journal of Translational Medicine**, London, v. 16, n. 1, dez. 2018.

MCGRATH, M. E.; PALMER, J. T.; BROMME, D.; SOMOZA, J. R. Crystal structure of human cathepsin S. **Protein Science**, Hoboken, v. 7, n. 6, p. 1294–1302, jun. 1998.

MCKERROW, J. H.; DOYLE, P. S.; ENGEL, J. C.; PODUST, L. M.; ROBERTSON, S. A.; FERREIRA, R.; SAXTON, T.; ARKIN, M.; KERR, I. D.; BRINEN, L. S.; CRAIK, C. S. Two approaches to discovering and developing new drugs for Chagas disease. **Memórias do Instituto Oswaldo Cruz**, Rio de Janeiro, v. 104, n. suppl 1, p. 263–269, jul. 2009.

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, 23 ago. 2018.

MOTTRAM, J. C.; BROOKS, D. R.; COOMBS, G. H. Roles of cysteine proteinases of trypanosomes and Leishmania in host-parasite interactions. **Current Opinion in Microbiology**, Oxford, v. 1, n. 4, p. 455–460, ago. 1998.

MULDER, W. J. M.; JAFFER, F. A.; FAYAD, Z. A.; NAHRENDORF, M. Imaging and nanomedicine in inflammatory atherosclerosis. **Science Translational Medicine**, Washington, v. 6, n. 239, p. 239sr1-239sr1, 4 jun. 2014.

MULLARD, A. Merck & Co. drops osteoporosis drug odanacatib. **Nature Reviews Drug Discovery**, London, v. 15, n. 10, p. 669–669, out. 2016.

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, dez. 2015.

OTTO, H.-H.; SCHIRMEISTER, T. Cysteine proteases and their inhibitors. **Chemical Reviews**, Washington, v. 97, n. 1, p. 133–172, fev. 1997.

PAPADATOS, G.; ALKAROURI, M.; GILLET, V. J.; WILLETT, P.; KADIRKAMANATHAN, V.; LUSCOMBE, C. N.; BRAVI, G.; RICHMOND, N. J.; PICKETT, S. D.; HUSSAIN, J.; JOHN M. PRITCHARD, J. M.; COOPER, A. W. J.; SIMON J. F. MACDONALD, S. J. F. lead optimization using matched molecular pairs: inclusion of contextual information for enhanced prediction of herg inhibition, solubility, and lipophilicity. **Journal of Chemical Information and Modeling**, Washington, v. 50, n. 10, p. 1872–1886, 25 out. 2010.

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**, Cambridge, v. 9, n. 8, p. 2113–2129, 2018.

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

PRABHU, N. V.; SHARP, K. A. Heat capacity in proteins. **Annual Review of Physical Chemistry**, Palo Alto, v. 56, n. 1, p. 521–548, 5 maio 2005.

QUILLES JUNIOR, J. C.; BERNARDI, 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, 8 maio 2019.

ROBINSON, M. W.; DALTON, J. P. (eds.). **Cysteine proteases of pathogenic organisms**. New York : Springer Science; Austin: Landes Bioscience, 2011.

ROGERS, K. E.; KERANEN, H.; DURRANT, J. D.; RATMAN, J.; DOAK, A.; ARKIN, M. R.; MCCAMMON, J. A. Novel cruzain inhibitors for the treatment of Chagas' Disease. **Chemical Biology & Drug Design**, Hoboken, v. 80, n. 3, p. 398–405, set. 2012.

SAJID, M.; ROBERTSON, S. A.; BRINEN, L. S.; MCKERROW, J. H. Cruzain. *In*: ROBINSON, M. W.; DALTON, J. P. (eds.). **Cysteine proteases of pathogenic organisms**. Boston: Springer, 2011. v. 712, p. 100–115.

SANDERSON, S. J.; POLLOCK, K. G.; HILLEY, J. D.; MELDAL, M.; HILAIRE, P. S.; JULIANO, M. A.; JULIANO, L.; MOTTRAM, J. C.; COOMBS, G. H. Expression and characterization of a recombinant cysteine proteinase of Leishmania mexicana. **Biochemical Journal**, London, v. 347, n. 2, p. 383, 15 abr. 2000.

SANI, M.; VOLONTERIO, A.; ZANDA, M. The Trifluoroethylamine Function as Peptide Bond Replacement. **ChemMedChem**, Weinheim, v. 2, n. 12, p. 1693–1700, 10 dez. 2007.

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, nov. 2015.

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.; 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, nov. 2017.

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, abr. 2011.

SMITH, A. J. T.; ZHANG, X.; LEACH, A.G.; HOUK, K. N. Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins. **Journal of Medicinal Chemistry**, Washington, v. 52, n. 2, p. 225–233, 22 jan. 2009.

SOLÍS-MENDIOLA, S.; ROJO-DOMÍNGUEZ, A.; HERNÁNDEZ-ARANA, A. Cooperativity in the unfolding transitions of cysteine proteinases. Calorimetric study of the heat denaturation of chymopapain and papain. **Biochimica et Biophysica Acta (BBA) -** Protein Structure and Molecular Enzymology, Amsterdam, v. 1203, n. 1, p. 121–125, nov. 1993.

STONE, J. A.; MCCREA, J. B.; WITTER, R.; STOCH, S. S. Clinical and translational pharmacology of the cathepsin K inhibitor odanacatib studied for osteoporosis. **British Journal of Clinical Pharmacology**, Chichester, v. 85, n. 6, p. 1072–1083, jun. 2019.

STUMPFE, D.; BAJORATH, J. Exploring activity cliffs in medicinal chemistry: miniperspective. **Journal of Medicinal Chemistry**, Washington, v. 55, n. 7, p. 2932–2942, 12 abr. 2012.

TABISH, T.; PRANJOL, M. Z.; HORSELL, D. W.; RAHAT, A. A. M.; WHATMORE, J. L.; WINYARD, P. G.; 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, 6 mar. 2019.

TALELE, T. T. The "Cyclopropane Fragment" is a versatile player that frequently appears in preclinical/clinical drug molecules. **Journal of Medicinal Chemistry**, Washington, v. 59, n. 19, p. 8712–8756, 13 out. 2016.

TURK, V.; STOKA, V.; VASILJEVA, O.; RENKO, M.; SUN, T.; TURK, B.; TURK, D. Cysteine cathepsins: from structure, function and regulation to new frontiers. **Biochimica et Biophysica Acta (BBA)** - Proteins and Proteomics, Amsterdam, v. 1824, n. 1, p. 68–88, jan. 2012.

VERMA, S.; DIXIT, R.; PANDEY, K. C. Cysteine proteases: modes of activation and future prospects as pharmacological targets. **Frontiers in Pharmacology**, Lausanne, v. 7, p. 1-12, 25 abr. 2016.

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, 28 abr. 2011.

WILLIAMS, R. A.; TETLEY, L.; MOTTRAM, J. C.; COOMBS, G. H. Cysteine peptidases CPA and CPB are vital for autophagy and differentiation in Leishmania mexicana. **Molecular Microbiology**, Chichester, v. 61, n. 3, p. 655–674, ago. 2006.

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, dez. 1973.

ZAIDI, M.; TROEN, B.; MOONGA, B. S.; ABE, E. Cathepsin K, osteoclastic resorption, and osteoporosis therapy. **Journal of Bone and Mineral Research**, Hoboken, v. 16, n. 10, p. 1747–1749, 1 out. 2001.

ZHAI, X.; MEEK, T. D. Catalytic mechanism of cruzain from *Trypanosoma cruzi* as determined from solvent kinetic isotope effects of steady-state and pre-steady-state kinetics. **Biochemistry**, Washington, v. 57, n. 22, p. 3176–3190, 5 jun. 2018.

7. APPENDIX

APPENDIX A- K_M ASSAYS

- Cruzain assay replicates:

Figure 1: Cruzain velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.
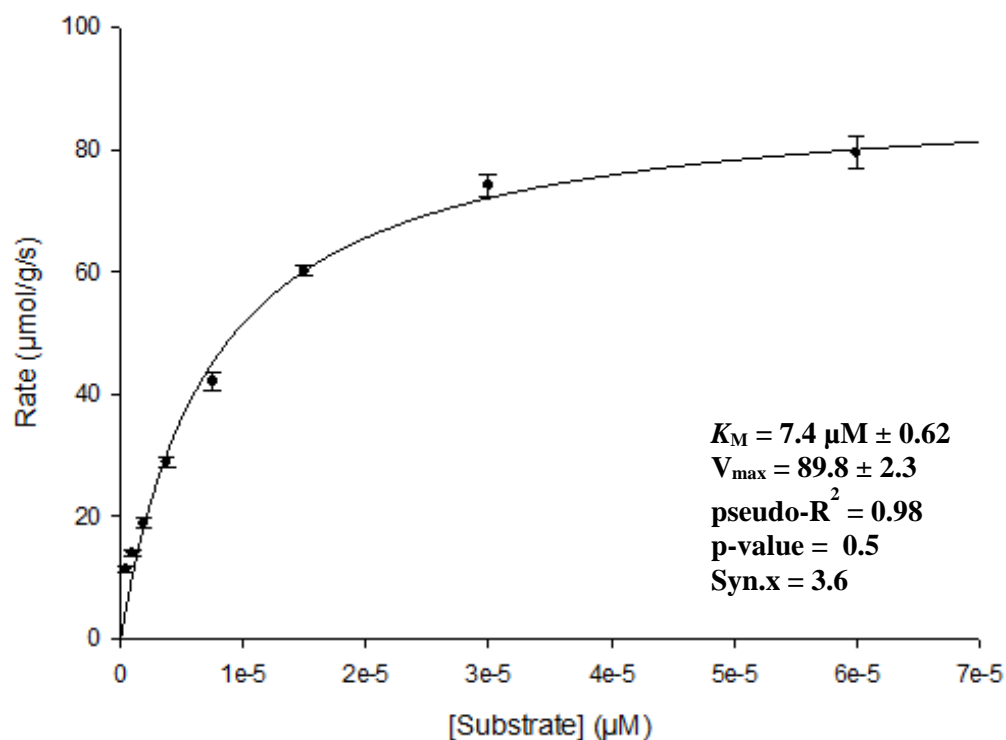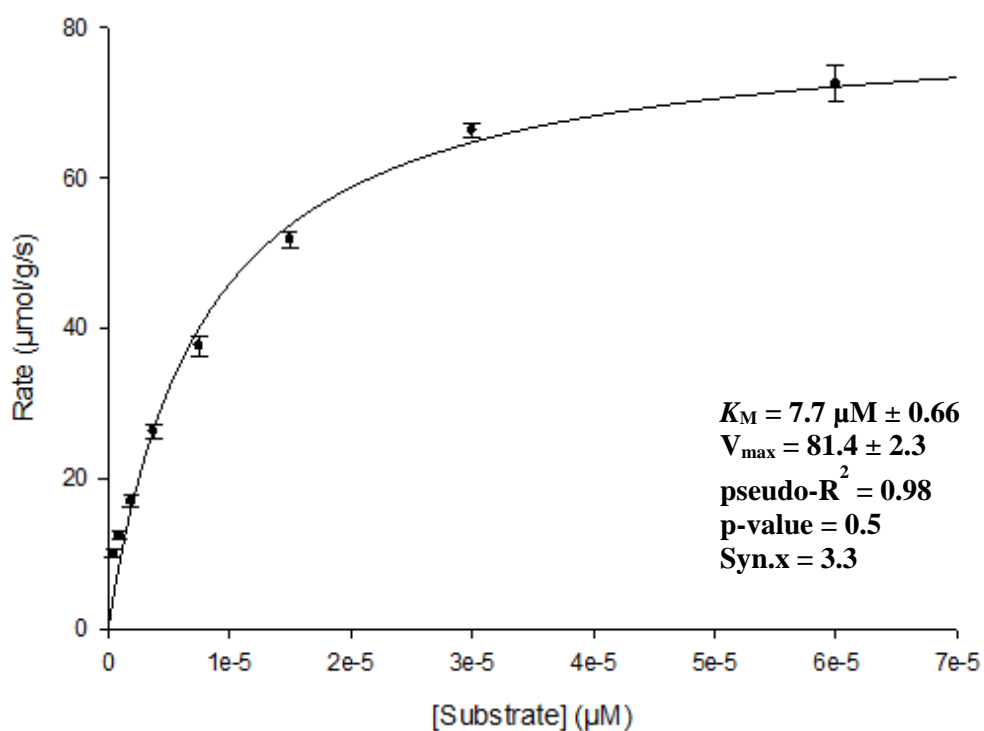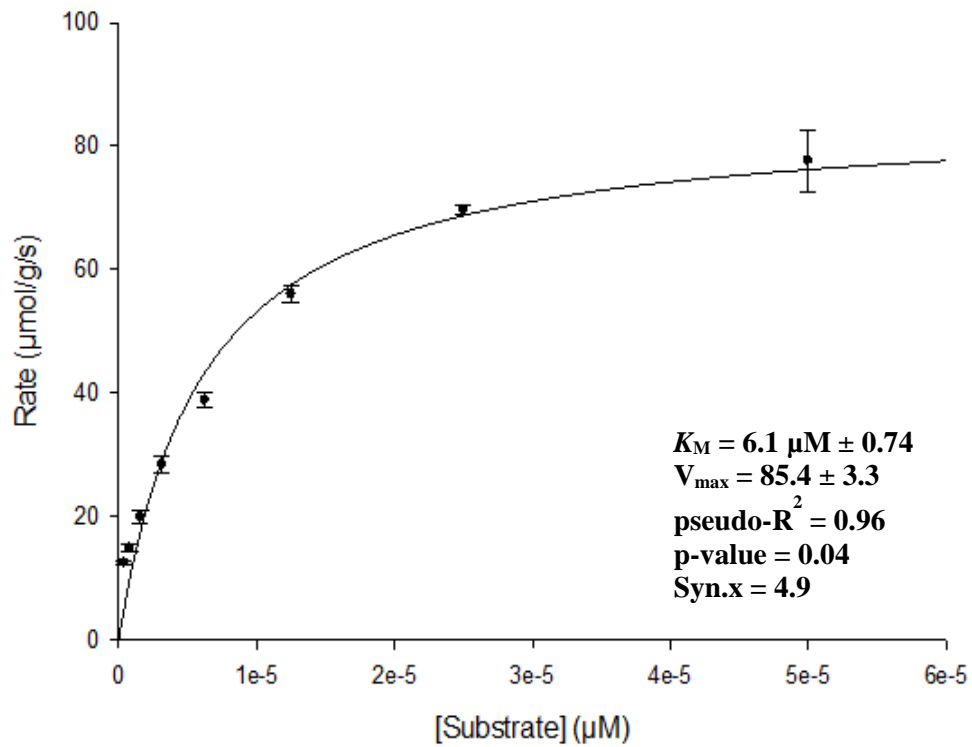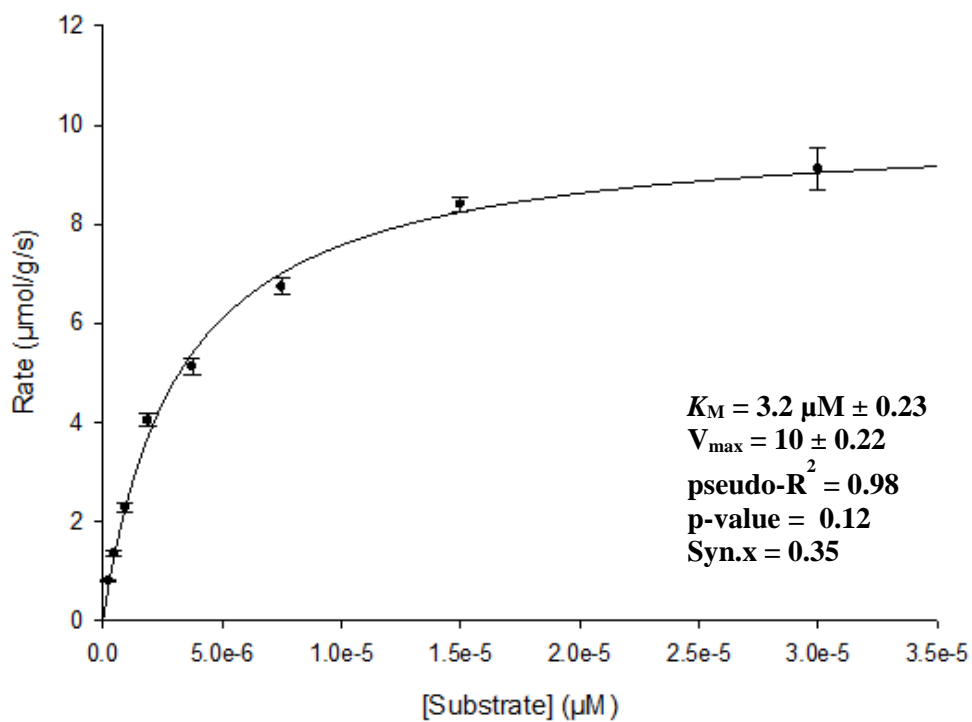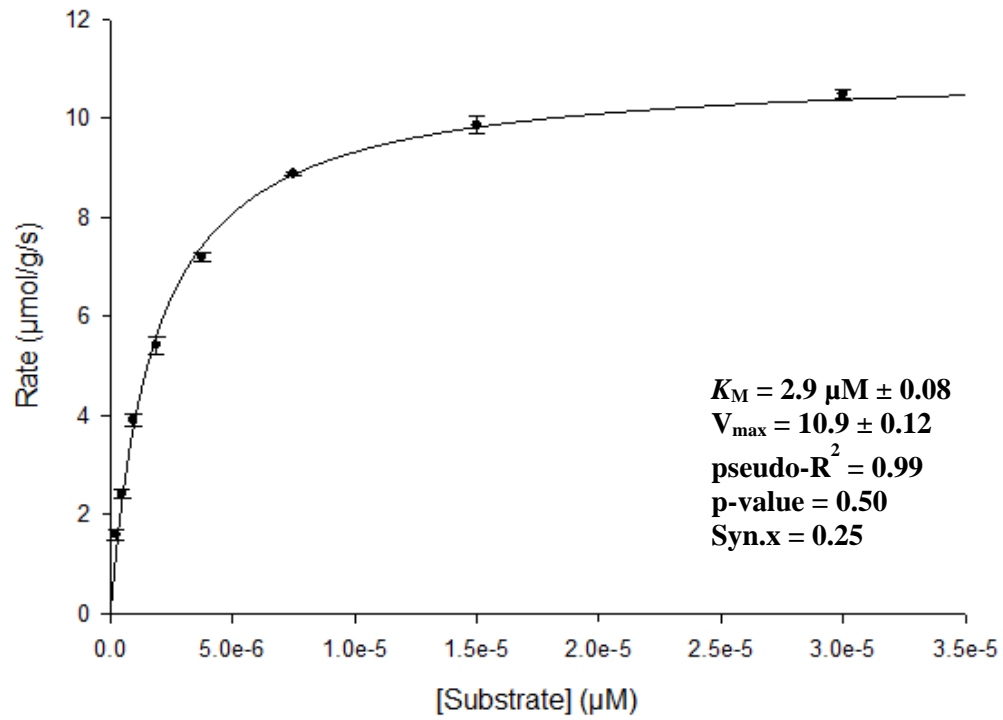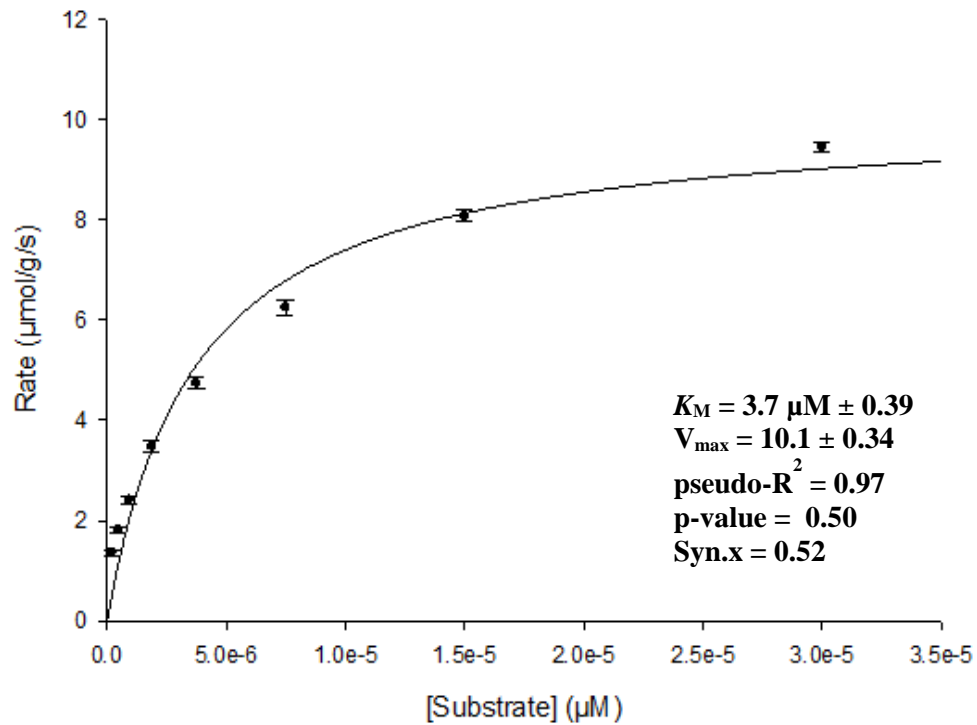

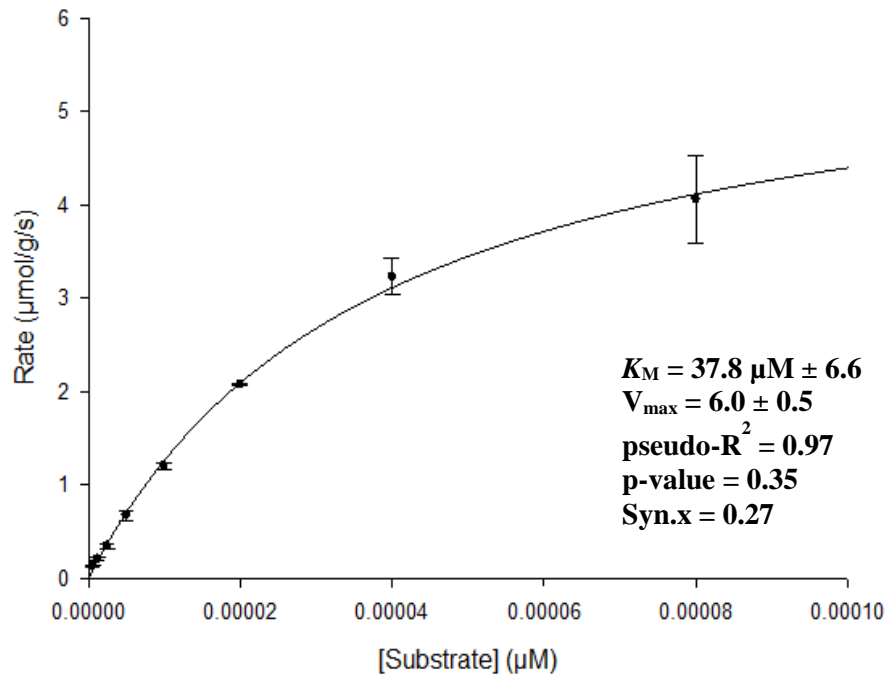K_M = 1.55 µM ± 0.013
V_{max} = 7.74 ± 0.17
pseudo-R^2 = 0.97
p-value = 0.50
Syn.x = 0.36
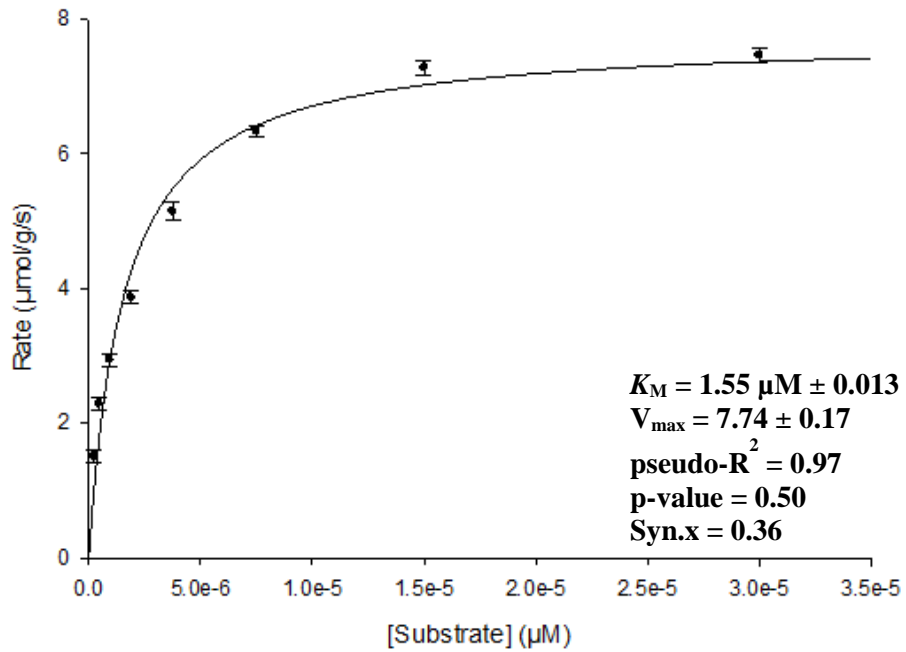
Figure 2: Cruzain velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

K_M = 1.63 µM ± 0.16
V_{max} = 7.56 ± 0.19
pseudo-R^2 = 0.97
p-value = 0.11
Syn.x = 0.39
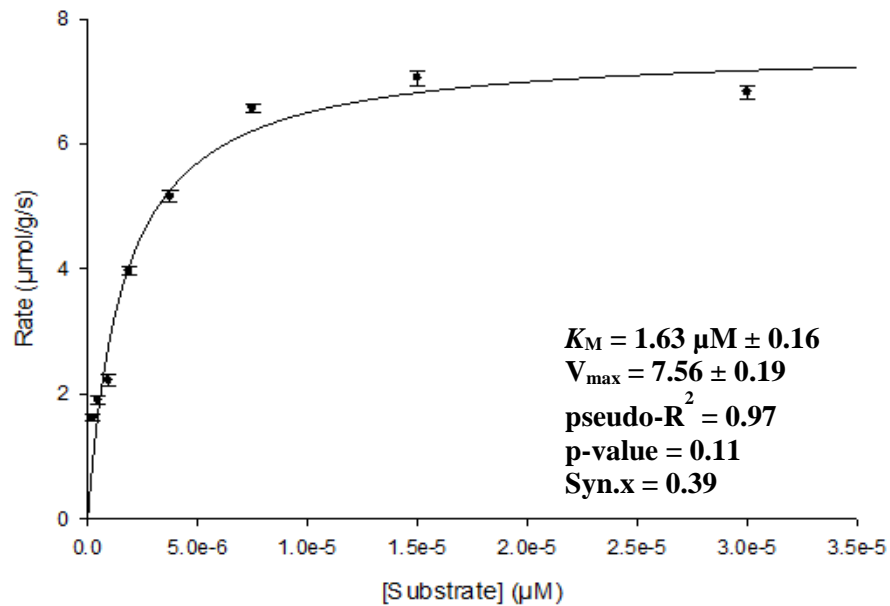
Figure 3: Cruzain velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

K_M = 1.98 μM ± 0.16
V_{max} = 6.26 ± 0.14
pseudo-R^2 = 0.98
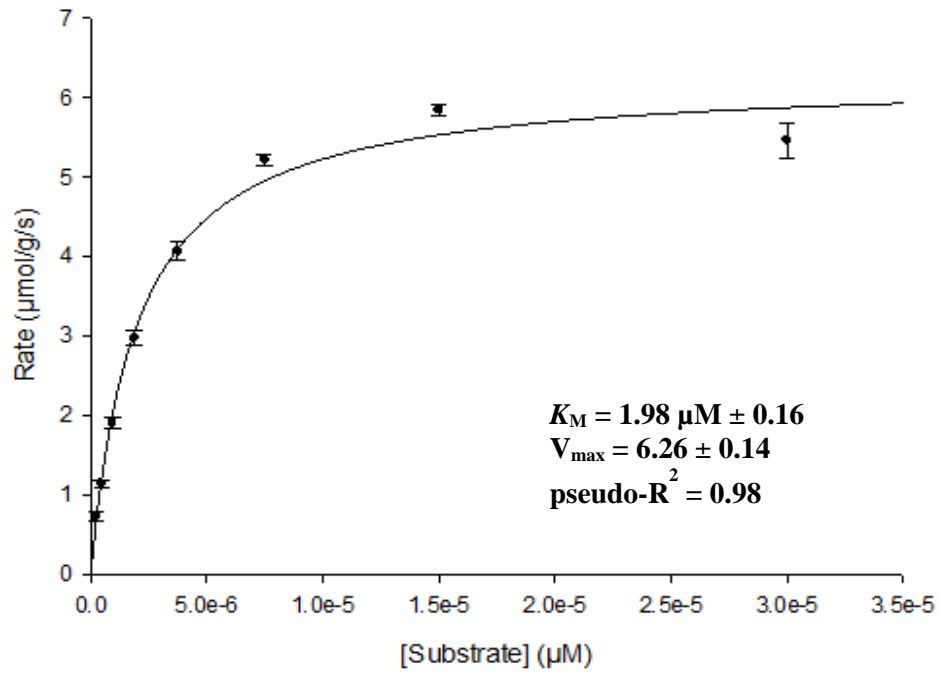
- *Lm*CPB assay replicates:

Figure 4: *Lm*CPB velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

K_M = 5.50 μM ± 0.74
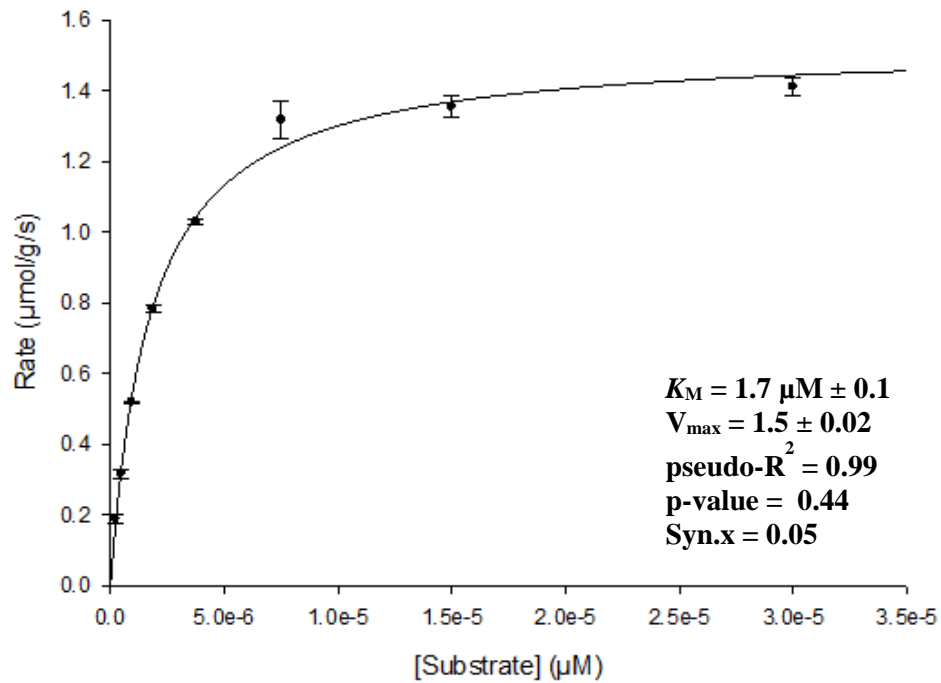V_{max} = 67.3 ± 3.6
pseudo-R^2 = 0.98
p-value = 0.16
Syn.x = 2.6
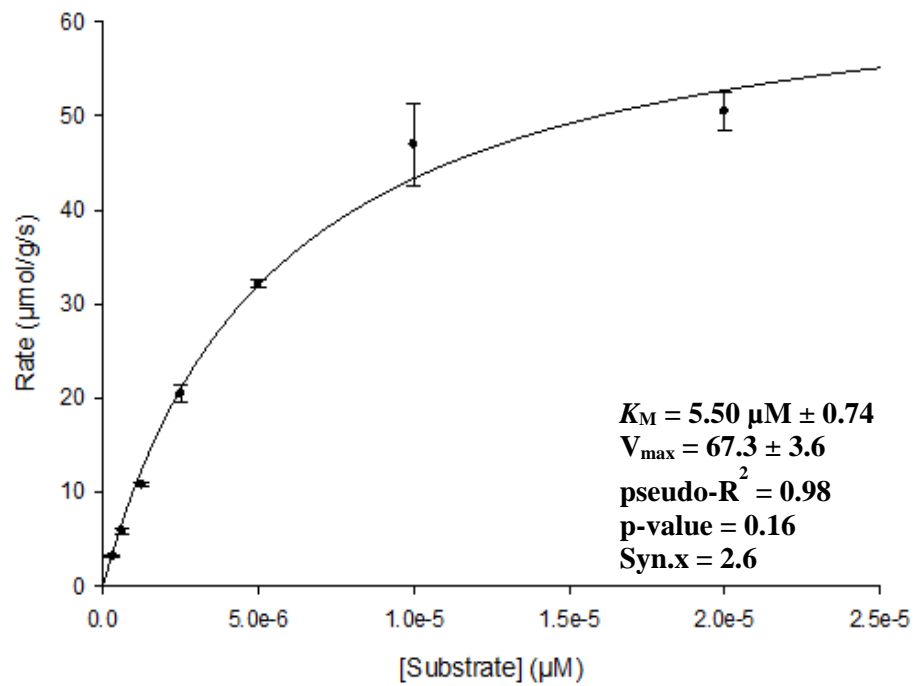
Figure 5: *Lm*CPB velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

$K_M = 7.4\ \mu M \pm 0.62$
$V_{max} = 89.8 \pm 2.3$
pseudo-$R^2 = 0.98$
p-value = 0.5
Syn.x = 3.6

Figure 6: *Lm*CPB velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

$K_M = 7.7\ \mu M \pm 0.66$
$V_{max} = 81.4 \pm 2.3$
pseudo-$R^2 = 0.98$
p-value = 0.5
Syn.x = 3.3

Figure 7: *Lm*CPB velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

K_M = 6.1 µM ± 0.74
V_{max} = 85.4 ± 3.3
pseudo-R^2 = 0.96
p-value = 0.04
Syn.x = 4.9

- *h*CatL replicates:

Figure 8: *h*CatL velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

K_M = 3.2 µM ± 0.23
V_{max} = 10 ± 0.22
pseudo-R^2 = 0.98
p-value = 0.12
Syn.x = 0.35

Figure 9: *h*CatL velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

K_M = 2.9 µM ± 0.08
V_{max} = 10.9 ± 0.12
pseudo-R^2 = 0.99
p-value = 0.50
Syn.x = 0.25

Figure 10: *h*CatL velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

K_M = 3.7 µM ± 0.39
V_{max} = 10.1 ± 0.34
pseudo-R^2 = 0.97
p-value = 0.50
Syn.x = 0.52

- *h*CatS replicates:

Figure 11: *h*CatS velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

$K_M = 37.8\ \mu M \pm 6.6$
$V_{max} = 6.0 \pm 0.5$
pseudo-$R^2 = 0.97$
p-value = 0.35
Syn.x = 0.27

Figure 12: *h*CatS velocity plot as a function of Z-Phe-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

$K_M = 36.8\ \mu M \pm 6.6$
$V_{max} = 3.0 \pm 0.2$
pseudo-$R^2 = 0.96$
p-value = 0.12
Syn.x = 0.14
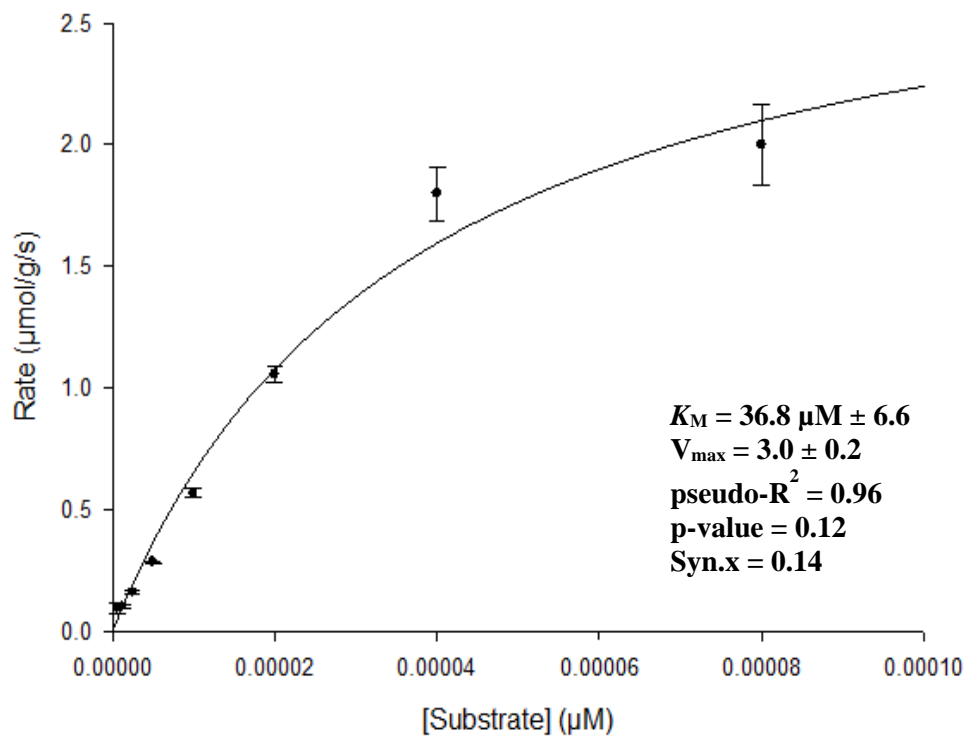
- *h*CatK replicates:

Figure 13: *h*CatK velocity plot as a function of Z-Leu-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

$K_M = 1.7\ \mu M \pm 0.1$
$V_{max} = 1.5 \pm 0.02$
pseudo-$R^2 = 0.99$
p-value = 0.44
Syn.x = 0.05

Figure 14: *h*CatK velocity plot as a function of Z-Leu-Arg-AMC concentration adjusted in the Michaelis-Menten equation.

$K_M = 1.6\ \mu M \pm 0.2$
$V_{max} = 2.9 \pm 0.1$
pseudo-$R^2 = 0.92$
p-value = 0.50
Syn.x = 0.23
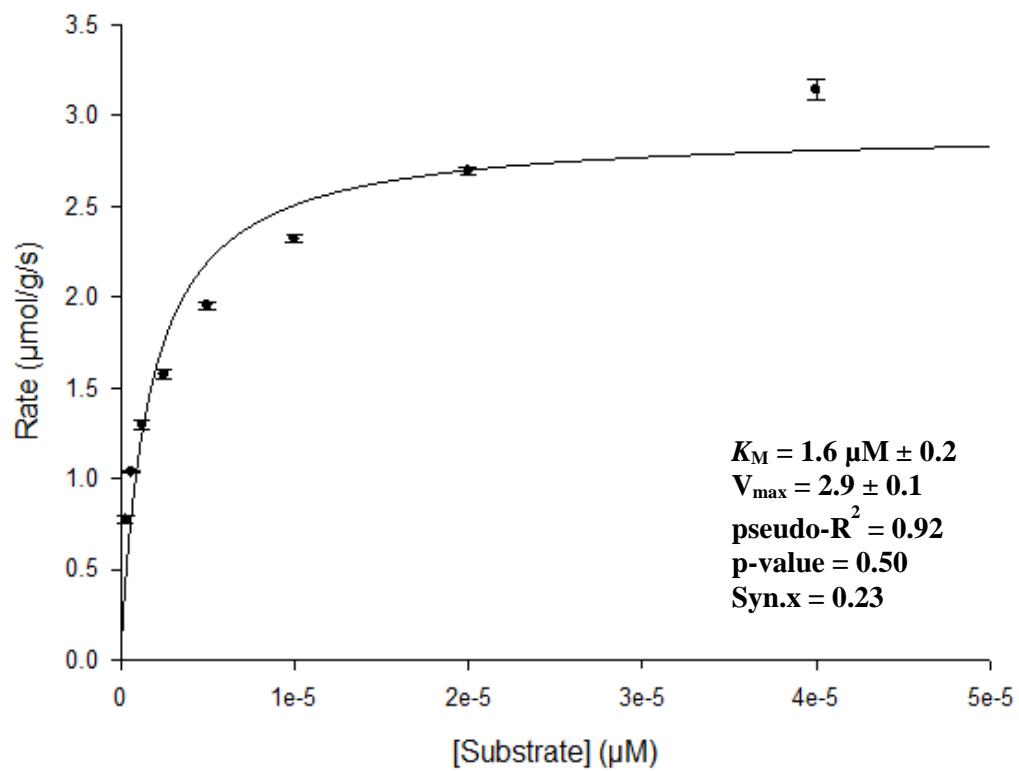
APPENDIX B- SCATTER PLOTS BETWEEN THE FIVE CPS

- Scatter plot of Cz against *Lm*CPB, *h*CatL, *h*CatS, and *h*CatK.

Figure 15: Scatter plot obtained when analyzed Cz with *Lm*CPB.
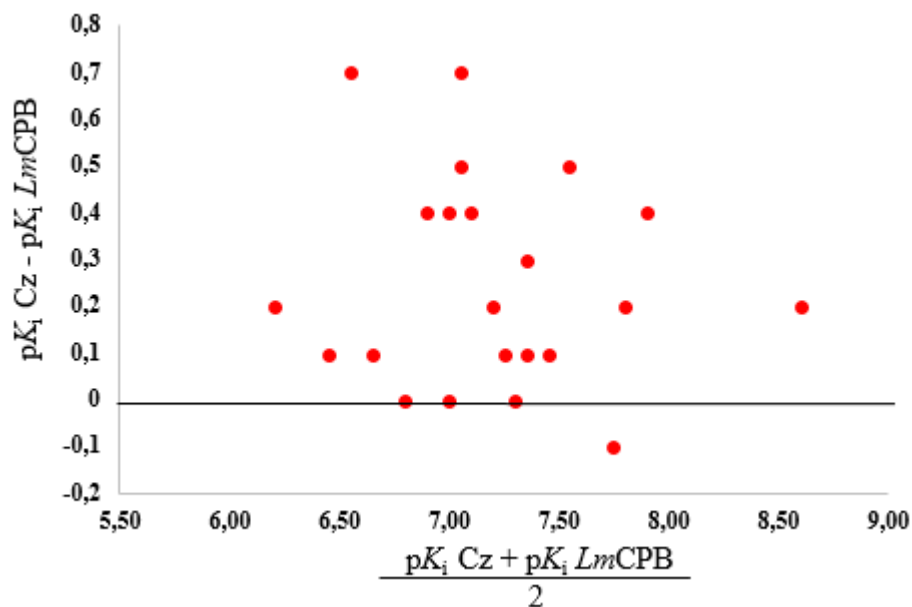
Figure 16: Scatter plot obtained when analyzed Cz with *h*CatL.
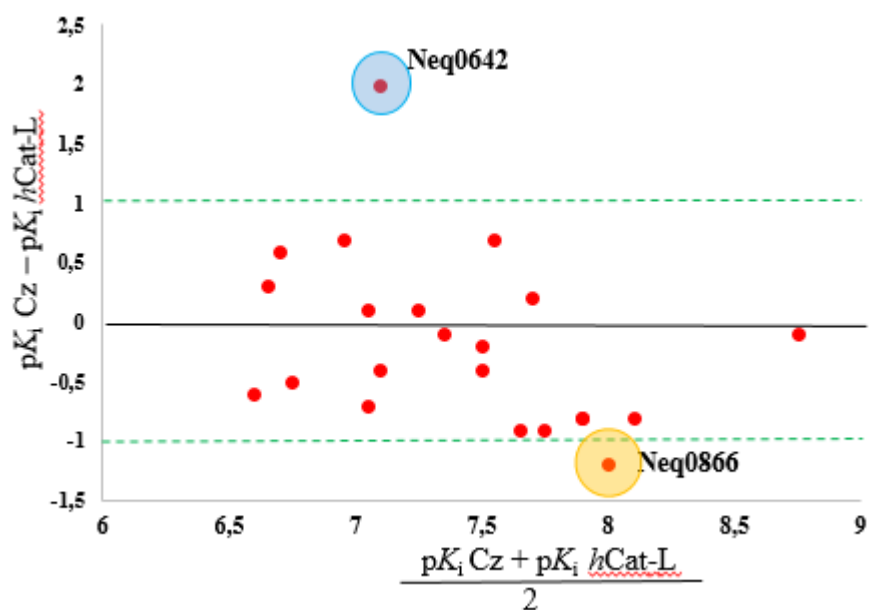
Figure 17: Scatter plot obtained when analyzed Cz with hCatS.
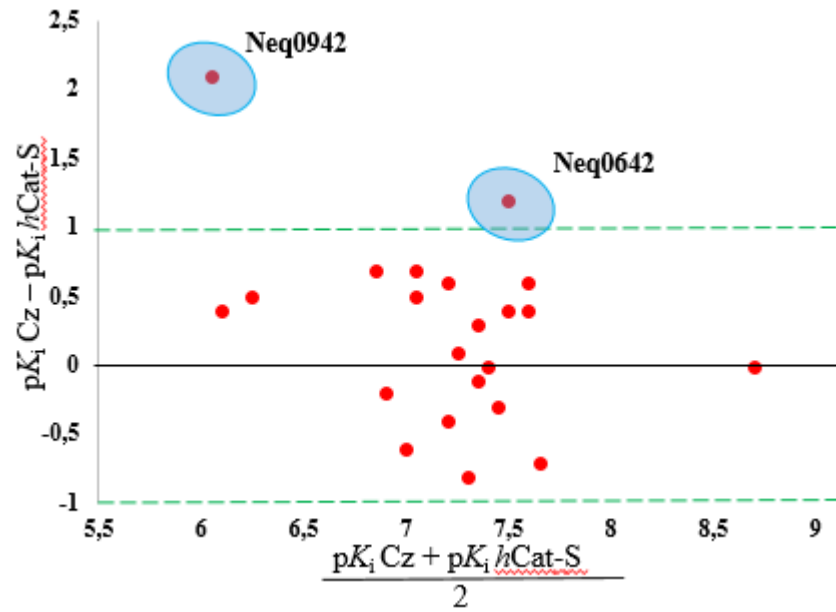
Figure 18: Scatter plot obtained when analyzed Cz with hCatK.
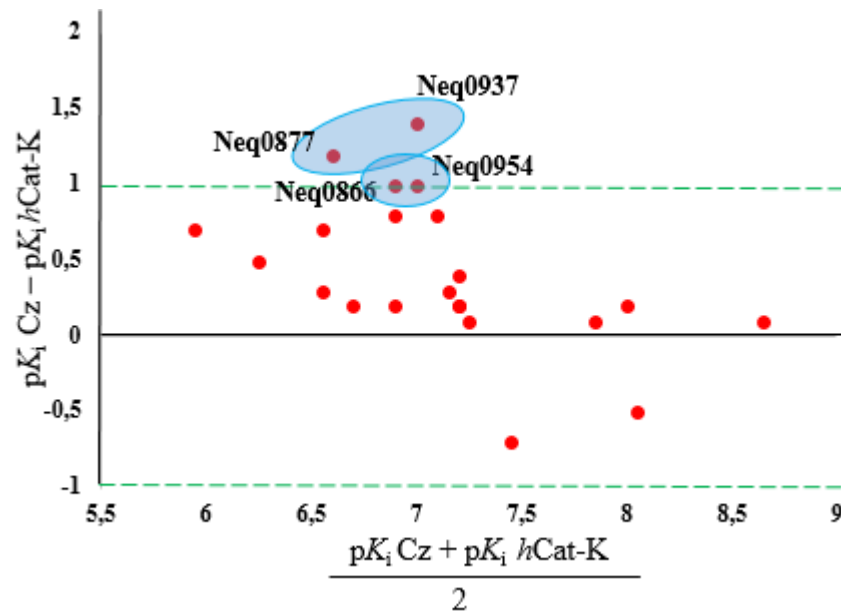
- Scatter plot of *Lm*CPB against *h*CatS and *h*CatK.

Figure 19: Scatter plot obtained when analyzed *Lm*CPB with *h*CatS.
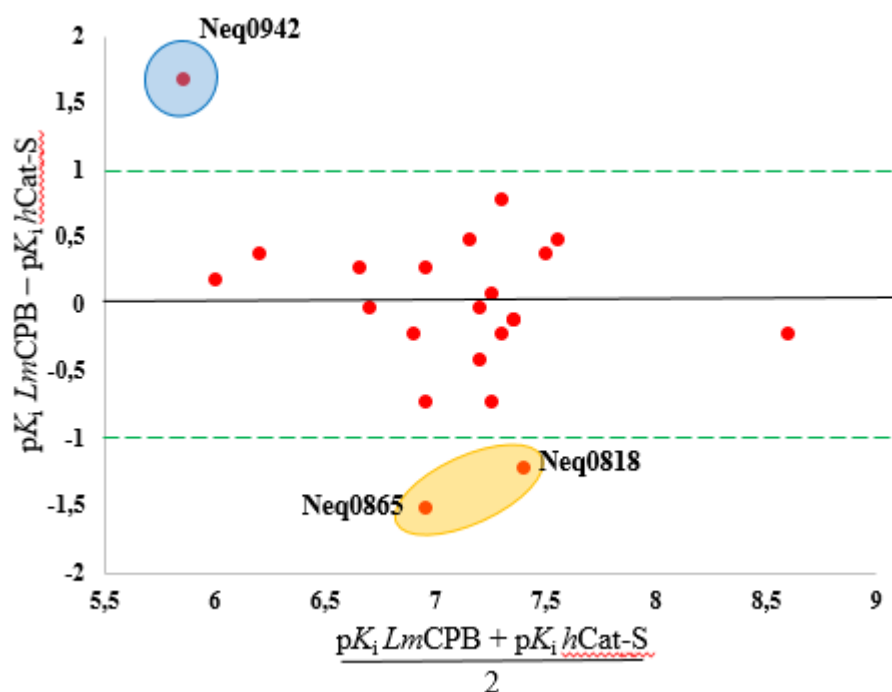
Figure 20: Scatter plot obtained when analyzed *Lm*CPB with *h*CatK.
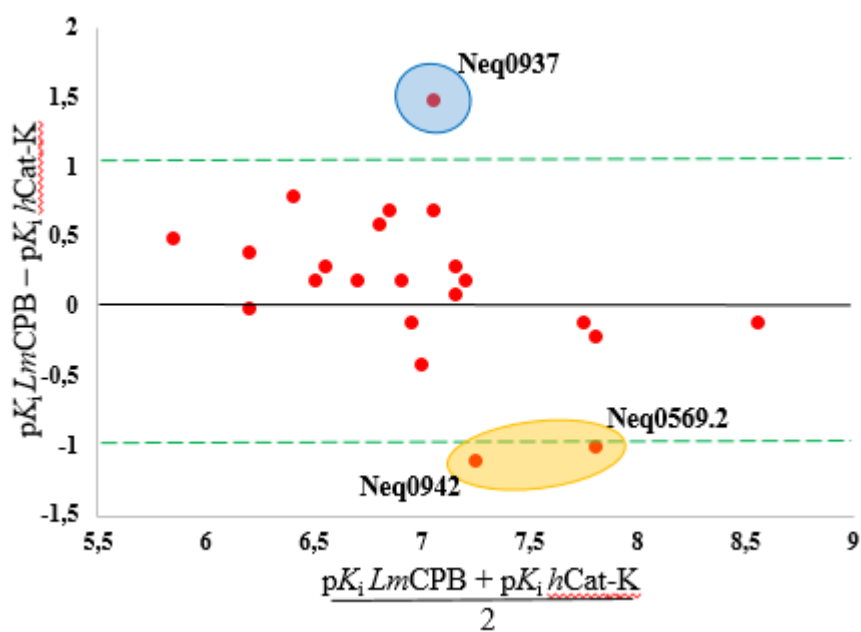
- Scatter plot of *h*CatL against *h*CatS and *h*CatK.

Figure 21: Scatter plot obtained when analyzed *h*CatL with *h*CatS.
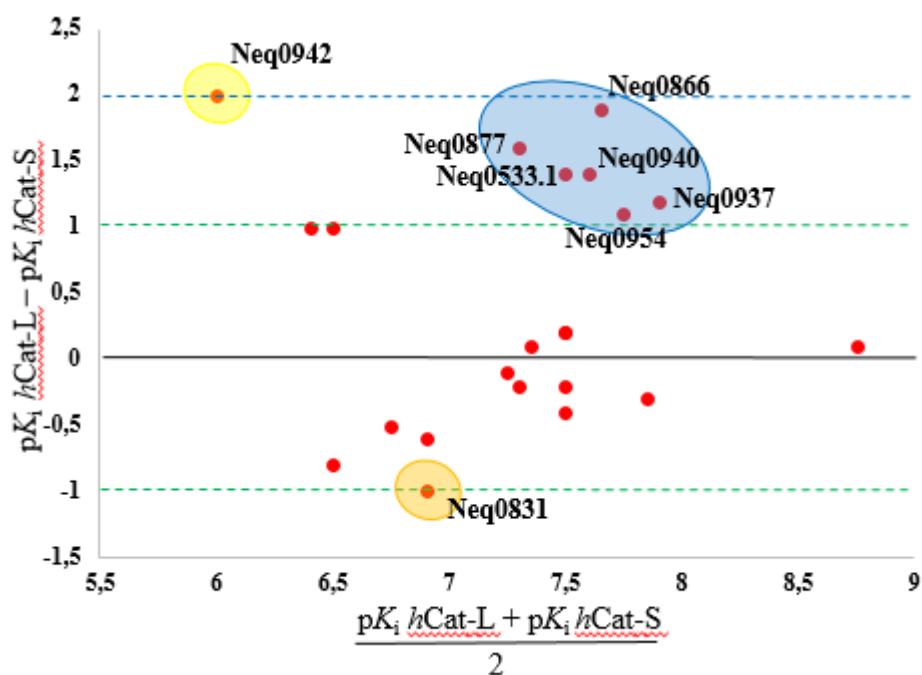
Figure 22: Scatter plot obtained when analyzed *h*CatL with *h*CatK.
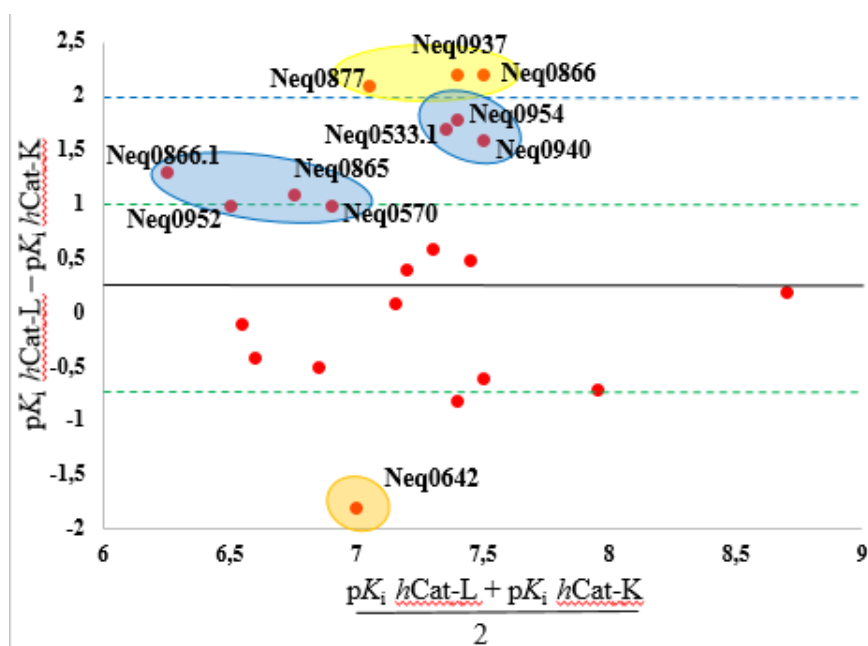
- Scatter plot *h*CatS against *h*CatK.

Figure 23: Scatter plot obtained when analyzed *h*CatS with *h*CatK.
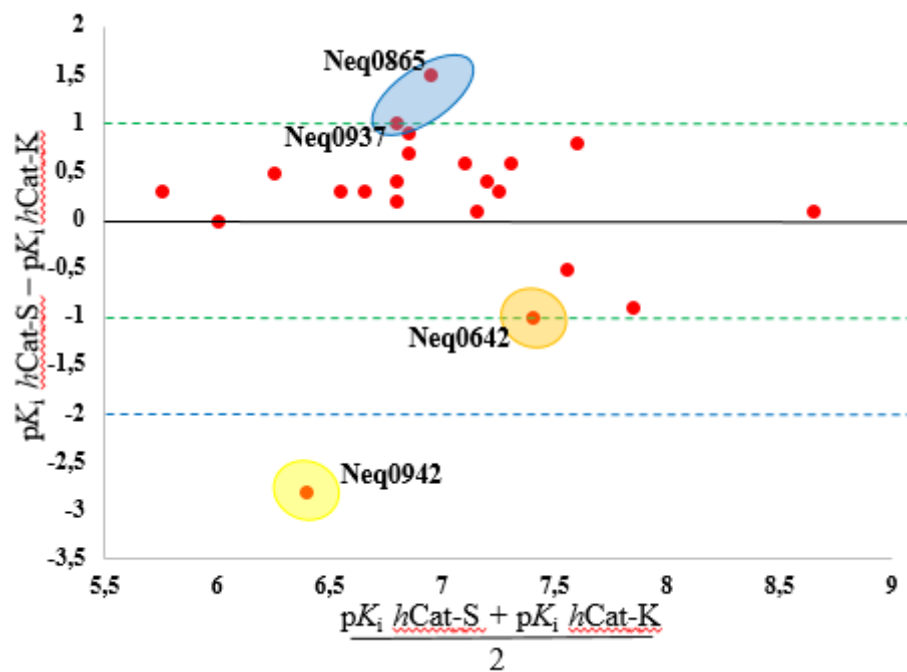
APPENDIX C- THERMAL STABILITY STUDY OF CRUZAIN AGAINST FOURTEEN LIGANDS

Figure 24: Cruzain unfolding peak when incubated with Neq0533.1 obtained after the correction of the blank (buffer - buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 μM.
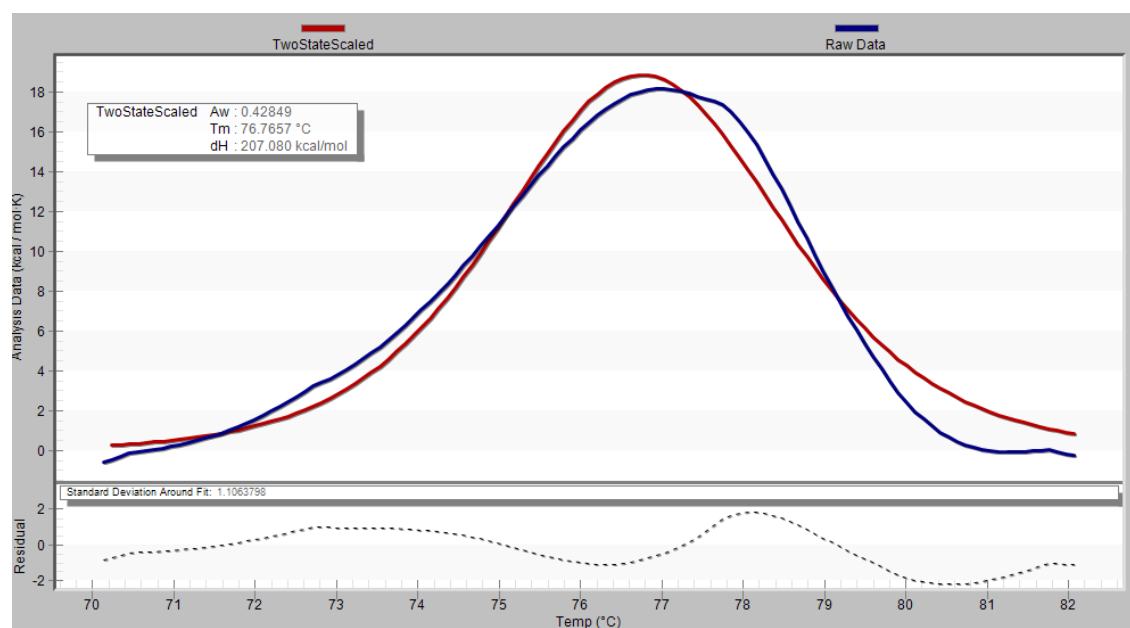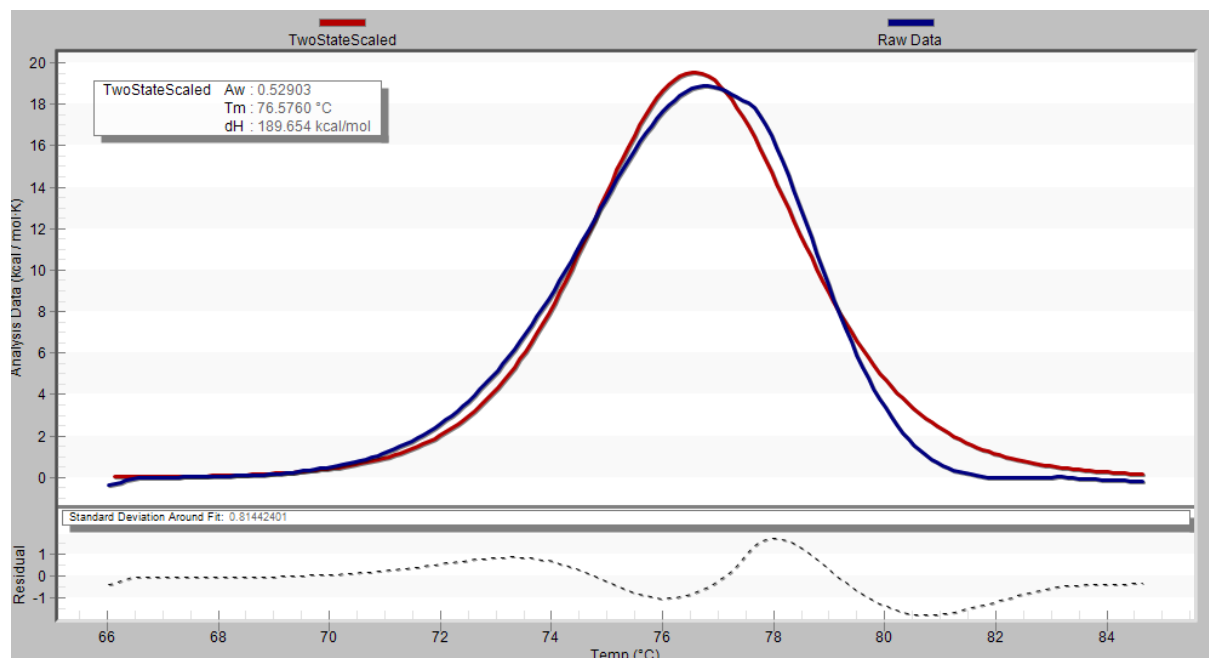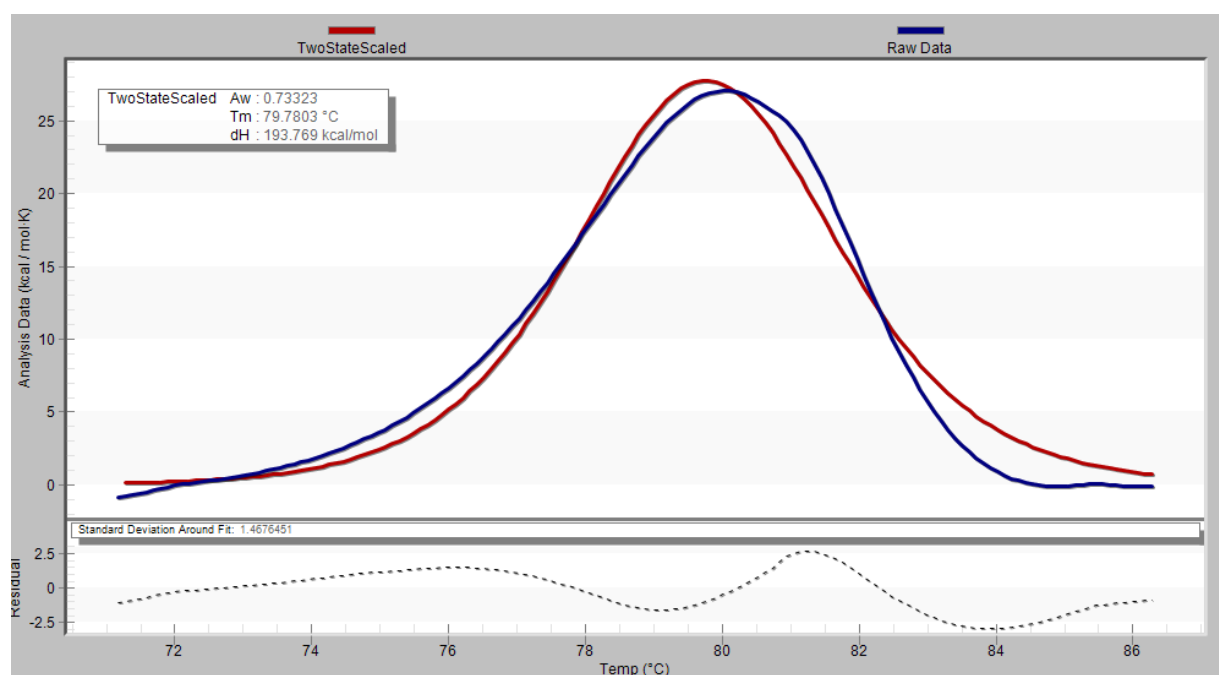

Figure 25: Cruzain unfolding peak when incubated with Neq0570 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 μM.
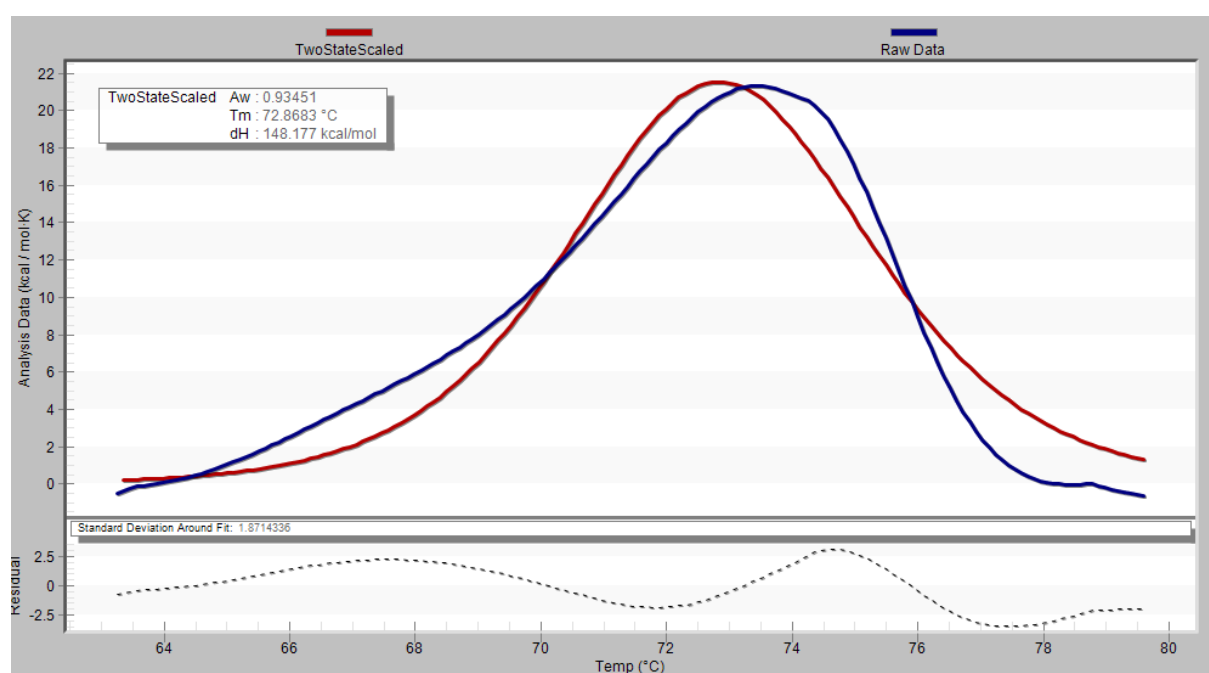
Figure 26: Cruzain unfolding peak when incubated with Neq0635.1 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 μM.
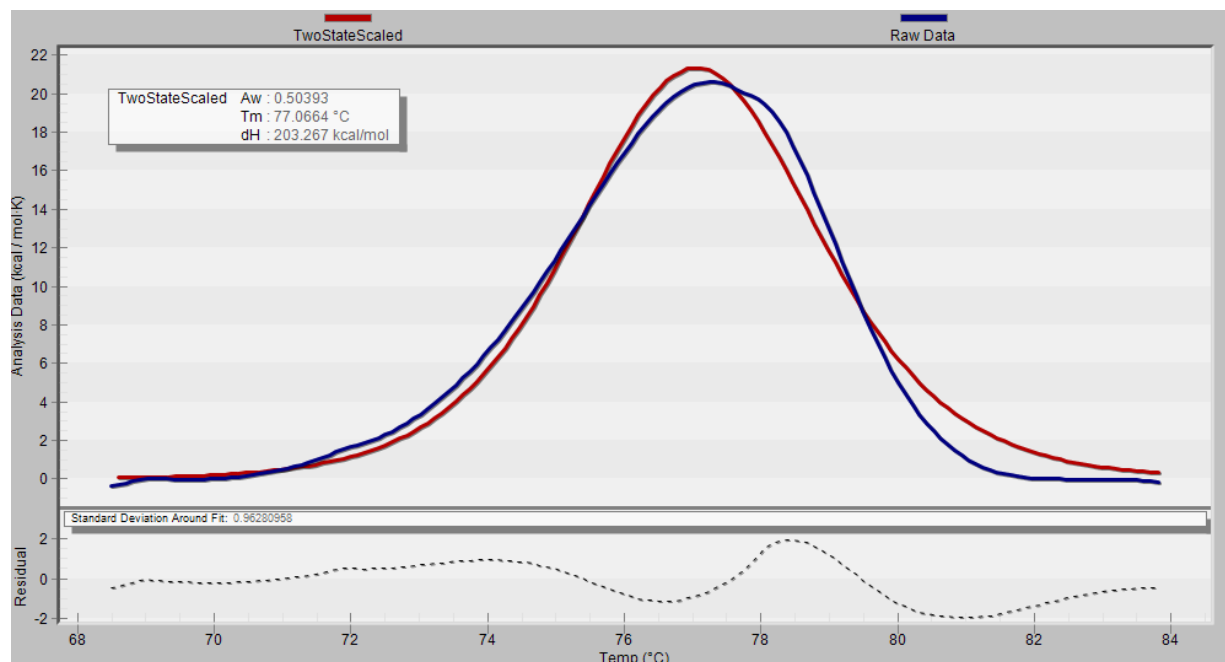

Figure 27: Cruzain unfolding peak when incubated with Neq0643.1 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 μM.
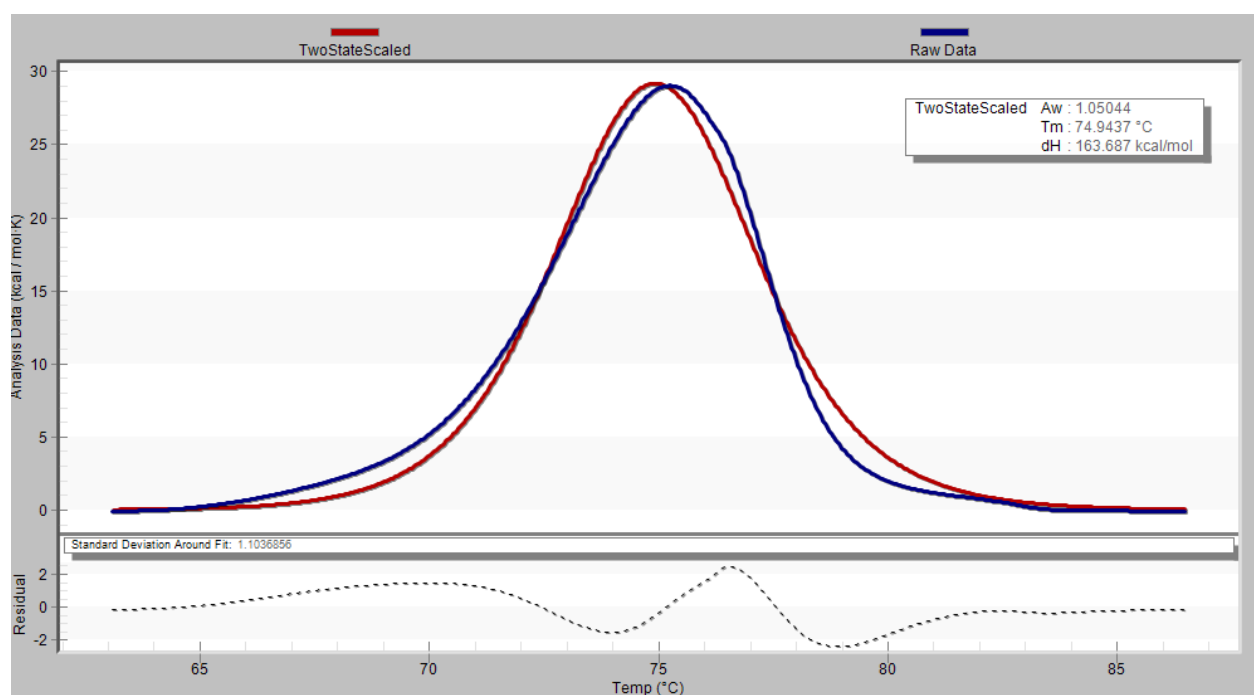
Figure 28: Cruzain unfolding peak when incubated with Neq0818 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 µM.
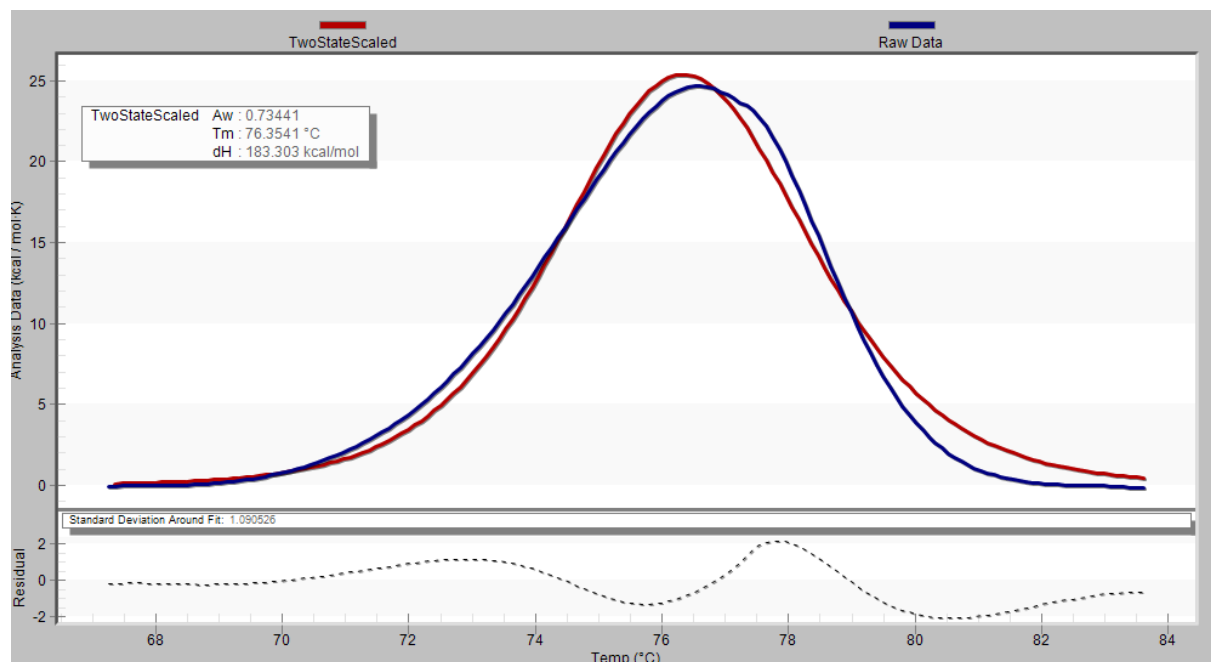

Figure 29: Cruzain unfolding peak when incubated with Neq0865 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 µM.

Figure 30: Cruzain unfolding peak when incubated with Neq0866 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 µM.
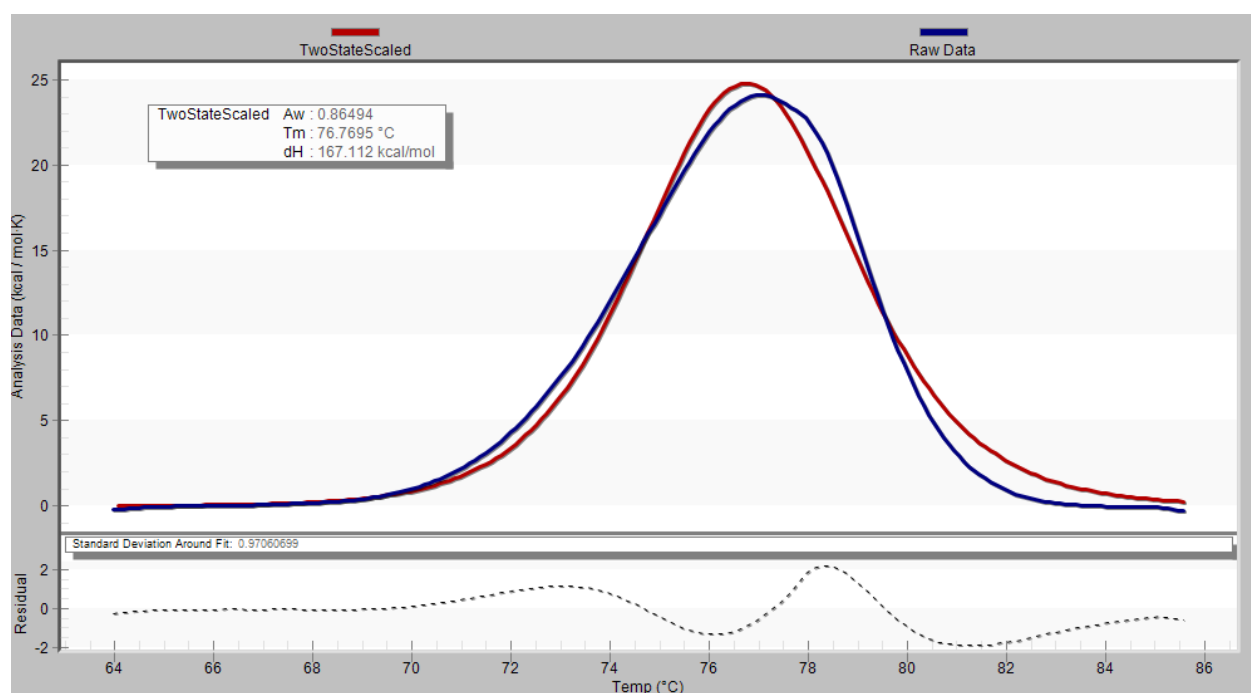

Figure 31: Cruzain unfolding peak when incubated with Neq0922 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 µM.
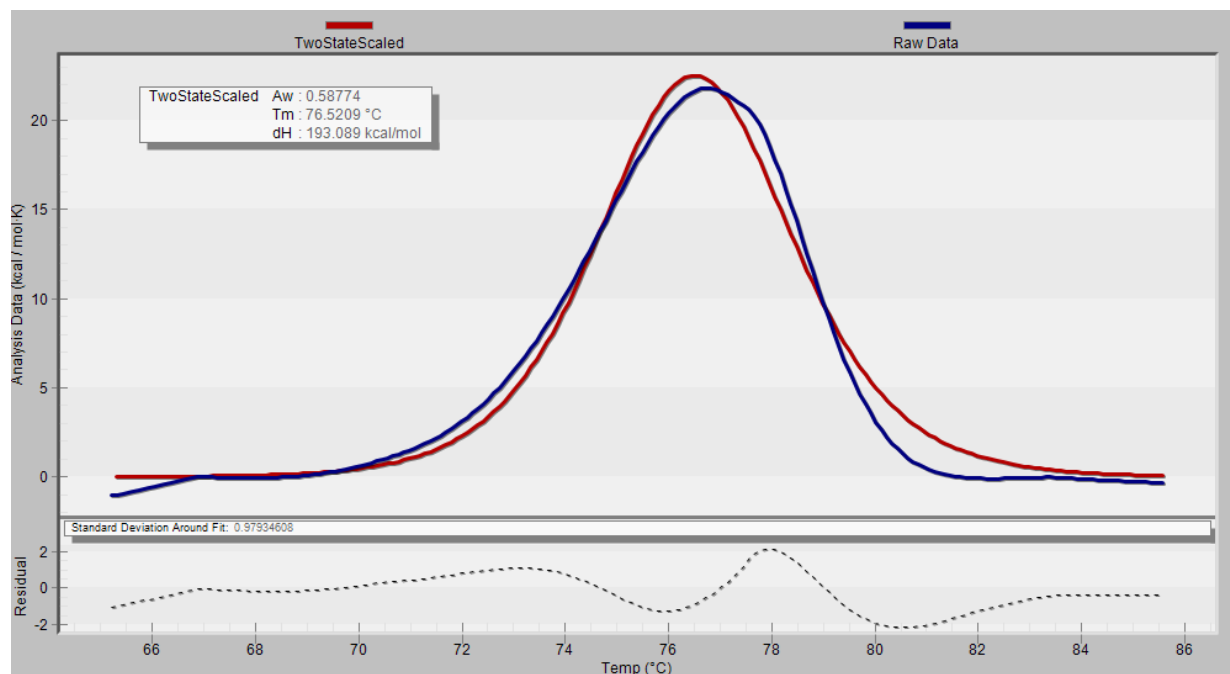
Figure 32: Cruzain unfolding peak when incubated with Neq0925 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 μM.

Figure 33: Cruzain unfolding peak when incubated with Neq0937 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 μM.
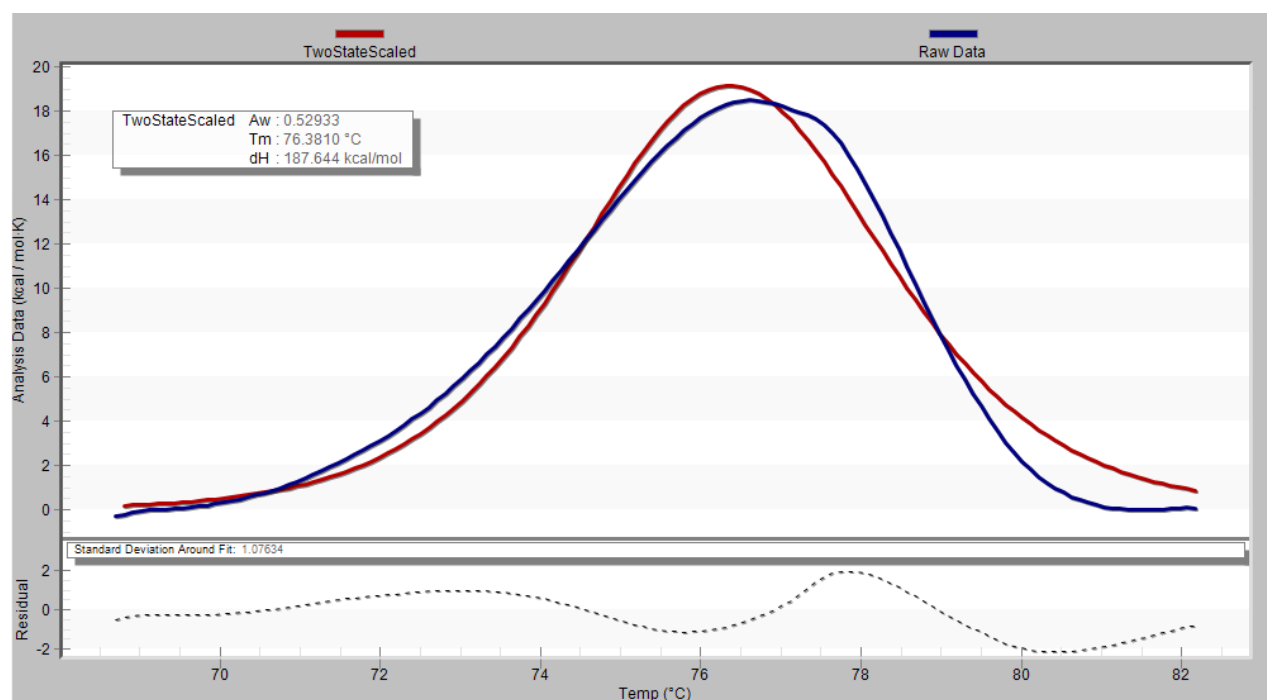
Figure 34: Cruzain unfolding peak when incubated with Neq0940 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 µM.
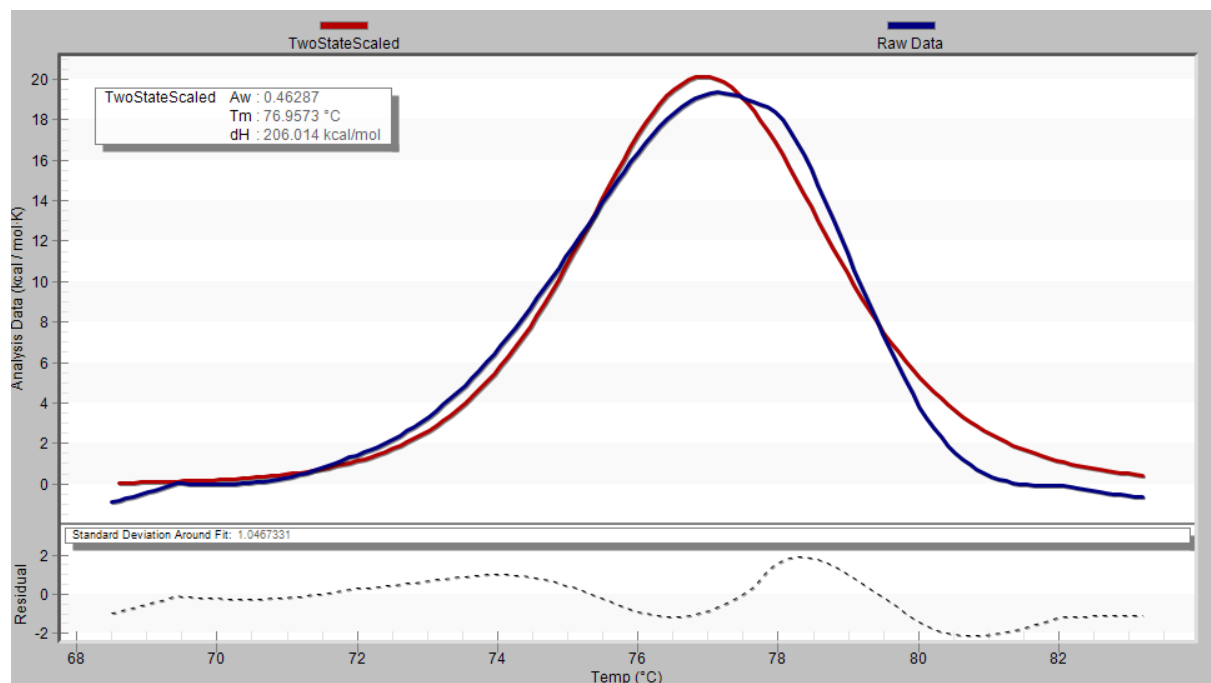

Figure 35: Cruzain unfolding peak when incubated with Neq0942 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 µM.
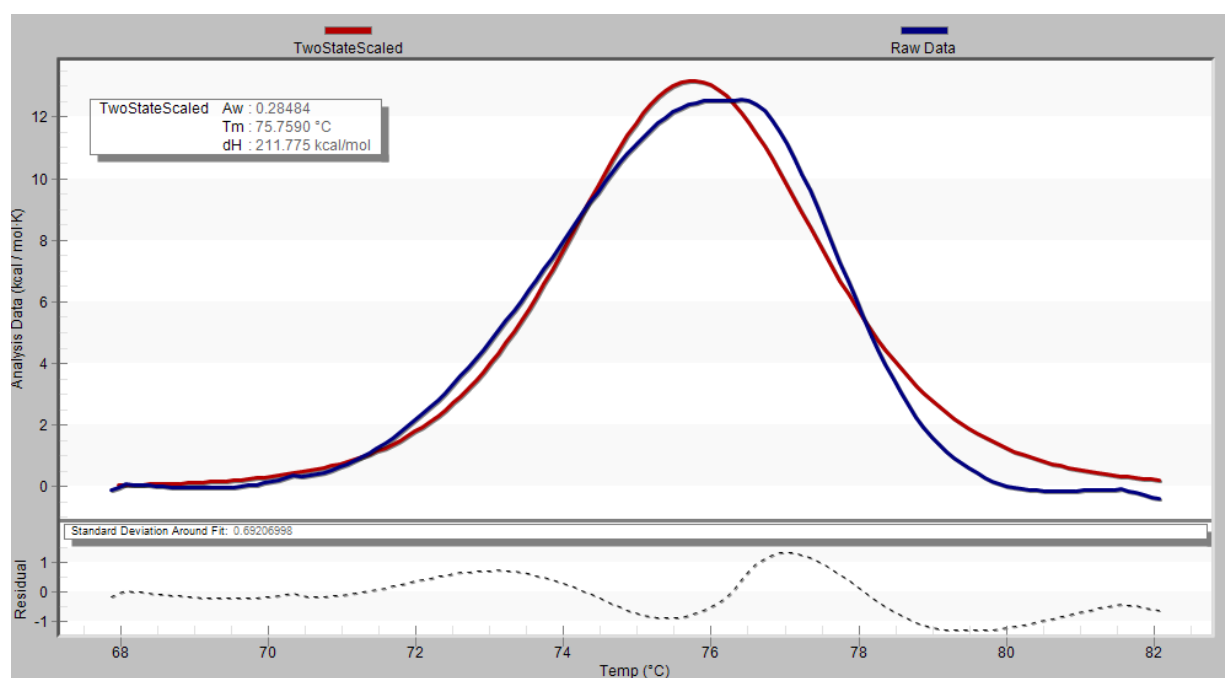
Figure 36: Cruzain unfolding peak when incubated with Neq0954 obtained after the correction of the blank (buffer-buffer + ligand). The solution was heated from 0 to 90 °C in acetate buffer pH = 5.5, 3 atm and a rate of 2 °C/min. The ligand concentration was 120 μM.