

UNIVERSITY OF SÃO PAULO
INSTITUTE OF CHEMISTRY OF SÃO CARLOS
POST-GRADUATION PROGRAM IN CHEMISTRY

Vinícius Bonatto

**MOLECULAR MODELING OF REVERSIBLE COVALENT
INHIBITORS OF CYSTEINE PROTEASES**

**São Carlos, Brazil
2023**

Vinícius Bonatto

**MOLECULAR MODELING OF REVERSIBLE COVALENT
INHIBITORS OF CYSTEINE PROTEASES**

Ph.D. thesis presented to the post-graduation program in chemistry of the Institute of Chemistry of São Carlos for the achievement of the title of Doctor of Science.

Concentration Area: Organic and biological chemistry

Supervisor: Prof. Dr. Carlos Alberto Montanari

Co-supervisor: Prof. Dr. Jerônimo Lameira Silva

Supervisor in Spain: Prof. Dr. Francisco Javier Luque Garriga

Exemplar revisado

O exemplar original encontra-se em
acervo reservado na Biblioteca do IQSC-USP

São Carlos

2023

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

Assinatura:

Data:

Ficha Catalográfica elaborada pela Seção de Referência e Atendimento ao Usuário do SBI/IQSC

Bonatto, Vinícius

Molecular modeling of reversible covalent inhibitors of cysteine proteases /
Vinícius Bonatto. — São Carlos, 2023.

179 f.

Tese (Doutorado em Química Orgânica e Biológica) — Instituto de Química
de São Carlos / Universidade de São Paulo, 2023.

Edição revisada

Orientador: Prof. Dr. Carlos Alberto Montanari

Coorientador: Prof. Dr. Jerônimo Lameira

1. Molecular modeling. 2. Drug design. 3. Covalent inhibitors. 4. Molecular
dynamics. 5. Free energy perturbation. I. Título.



DEDICATION

Essa tese é dedicada com profunda gratidão aos meus pais, sem eles isto não seria possível, e eu não teria chegado a este momento.

This thesis is dedicated with deep gratitude to my parents, without them this would not have been possible, and I would not have reached this moment.

ACKNOWLEDGMENT

First, I have to thank my parents Ana and João, obrigado por tudo.

I would like to sincerely thank Prof. Jerônimo Lameira for all the help he gave to me and for his importance throughout my Ph.D. journey. The support he gave me and continues to give me all this time was crucial for me to develop my research and gradually become more independent.

To Prof. Carlos Montanari, for accepting to be my advisor and patiently teaching me more and more about medicinal chemistry, providing a comprehensive and interdisciplinary view of the process of developing new inhibitors and potential drug candidates. I also thank him for his understanding, teachings, and the conversations we had during all this time.

I would like to express my gratitude to Prof. Javier Luque, one of the brightest people I've had the opportunity to talk to and work with, for kindly having me in his lab. The time I spent in his laboratory was essential in my journey. I learned a lot from him and the other members of his lab.

I also want to thank Dr. Peter Kenny and Prof. Andrei Leitão for all our conversations, discussions, and teachings, especially in medicinal chemistry.

To the NEQUIMED group, especially those I had the opportunity to work closely. To Felipe, Shamim, Wellington, Gabriel, and Lorenzo for conducting the syntheses of the compounds proposed by me and collaborating on various projects. To Fernanda and Fabiana for the biochemical assays conducted. To Pedro for the analyses performed and for the various collaborative works we did. And, of course, to my computational crew, Rafael and Thiago GM, in the many projects, discussions, and adventures we embarked on to model covalent inhibitors.

I would also like to thank all the Computational Biology & Drug Design Group members led by Prof. Javi. They were very welcoming during the time I spent in Barcelona/Spain, and I learned a lot from each one of them through the many conversations we had, whether about computational chemistry, medicinal chemistry, other scientific subjects, or the cultural differences between our countries. Fue un placer estar con ustedes en este período, Mattia, prof. Carolina, Rhys, Toni, Javi Vázquez, Katerina, Mireia y Cristina.

Gracias también a Robin y Julio.

To all my family, especially my grandfather Armelindo and my uncle Nézinho, who unfortunately passed away before I could complete my thesis, but will always be present in my memory.

I would also like to thank my cousins Léo, Lucas, Bruno, and Sofia, who always provided the best moments outside of academia.

To Iza, for all the love, understanding, support, and patience during all the moments we spent together. Thanks to her support, completing this stage in my life was much easier. Also, to her entire family, especially Denílson, Sônia, Cesar, Fafá, and Isabela.

To my friends Bruna, Pudim, Amador, and Rubão.

Special thanks to the funding agencies, CAPES, CNPq, FAPESP, and the European Union program HPC-EUROPA3.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001

The work has also been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, I gratefully acknowledge the support of the Nutrition, Food Science department and the computer resources and technical support provided by BSC.

I also thank the high-performance computing provided by the Superintendência de Tecnologia da Informação from USP (STI-USP) and the Institute of Chemistry of São Carlos.

EPIGRAPH

*“He that riseth late
must trot all day.”*

(Benjamin Franklin)

RESUMO

Cisteíno proteases (CPs) são importantes alvos em diversas doenças infecciosas e câncer, sendo atrativas para o desenvolvimento de novos inibidores. Os inibidores covalentes direcionados, do inglês *Targeted covalent inhibitors* (TCIs) são extensivamente aplicados na inibição de CPs, pois promovem a modificação covalente destas enzimas. Entretanto, o planejamento e modelagem molecular de inibidores covalentes apresenta um desafio computacional, uma vez que o processo de inibição covalente consiste em múltiplos passos.

Este trabalho inicia com uma revisão sistemática da literatura de modo a analisar o uso da nitrila como grupos reativos para o desenvolvimento de inibidores covalentes. As nitrilas apresentam um grande potencial como *warheads* em alvos como as CPs e enzimas que contenham resíduos catalíticos de cisteína ou serina. Além disso, as nitrilas podem aumentar a reatividade de outros grupos, como é o caso da cianoacrilamida, e podem ser usadas em inibidores covalentes reversíveis que tem como alvo resíduos não catalíticos.

Em uma próxima etapa, foram empregados métodos da mecânica clássica (MM) e mecânica quântica (QM) para entender e prever a afinidade de ligantes por CPs específicas. Começando com os cálculos quânticos, descritores baseados em QM foram obtidos, como índices de eletrofilia local, para estimar a reatividade de diversos *warheads*, constatando evidente correlação com a constante de inibição enzimática experimental. Estes descritores poderão ser usados em modelos de aprendizado de máquina ou em modelos de relação estrutura-atividade.

Ainda usando métodos QM para modelar o *warhead*, a metodologia QM-cluster *approach* foi empregada para estimar a relativa diferença de potência de dois inibidores pares com *warheads* distintos (nitrila e aldeído). Com um modelo do sítio ativo contendo pouco mais de 300 átomos, essa técnica mostrou-se eficaz para capturar a diferença de energia entre os ligantes com diferentes *warheads*. Essa metodologia promissora pode ser empregada em futuras campanhas de desenvolvimento de novos inibidores covalentes, onde o alvo também é levado em consideração.

Por fim, foram realizadas simulações de dinâmica molecular para compreender o modo de ligação de diversos inibidores. Além disso, cálculos de perturbação de energia livre (FEP) foram realizados para estimar a energia livre de ligação relativa (RBFEE). Os resultados sugerem que o estado covalente do processo de interação bimolecular é fundamental e mais relevante para a energia livre do que do complexo não covalente.

No geral, as descobertas realizadas nesta tese contribuem para o desenvolvimento de inibidores covalentes reversíveis e abrem caminho para novos avanços no campo.

ABSTRACT

Cysteine proteases (CPs) are essential targets in various infectious diseases and cancer, making them attractive for developing new inhibitors. Targeted covalent inhibitors (TCIs) are extensively applied in the inhibition of CPs, as they promote the covalent modification of these enzymes. However, the molecular design and modeling of covalent inhibitors present a computational challenge since the process of covalent inhibition is complex.

This work begins with a systematic literature review to analyze the use of nitriles as reactive groups for the development of covalent inhibitors. Nitriles have great potential as warheads in targets such as CPs and enzymes containing catalytic cysteine or serine residues. Additionally, nitriles can enhance the reactivity of other groups, such as cyanoacrylamide, and can be used in reversible covalent inhibitors targeting non-catalytic residues.

Next, molecular mechanics (MM) and quantum mechanics (QM) methods were employed to understand and predict ligand affinity for specific CPs. Starting with quantum calculations, QM-based descriptors, such as local electrophilicity indices, were obtained to estimate the reactivity of various warheads, showing a clear correlation with experimental enzymatic inhibition constants. These descriptors can be used in machine learning or quantitative structure-activity relationship models.

Still, using QM methods to model the warhead, the QM-cluster approach was employed to estimate the relative potency difference between two pairs of inhibitors with distinct warheads (nitrile and aldehyde). With a binding site model containing just over 300 atoms, this technique effectively captured the energy difference between ligands with different warheads. This promising methodology can be applied in future campaigns to develop new covalent inhibitors, considering the target protein.

Finally, molecular dynamics simulations were performed to understand the binding modes of various inhibitors. Free energy perturbation (FEP) calculations were carried out to estimate the relative binding free energy (RBFEE). The results suggest that the covalent state of the bimolecular interaction process is fundamental and more relevant to the free energy than the non-covalent complex.

Overall, the findings of this thesis contribute to the development of reversible covalent inhibitors and pave the way for new advancements in the field.

LIST OF PUBLICATIONS

List of primary papers presented in this thesis:

I. BONATTO, V.; LAMEIRO, R. F.; ROCHO, F. R.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A. Nitriles: an attractive approach to the development of covalent inhibitors. **RSC Medicinal Chemistry**, Cambridge, v. 14, p. 201-217, 2023. DOI: 10.1039/D2MD00204C

II. BONATTO, V.; BATISTA, P. H. J.; CIANNI, L.; DE VITA, D. ; SILVA, D. G. ; CEDRON, R.; TEZUKA, D. Y.; DE ALBUQUERQUE, S.; MORAES, C. B.; FRANCO, C. H.; J.; LEITÃO, A.; MONTANARI, C. A. . On the intrinsic reactivity of highly potent trypanocidal cruzain inhibitors. **RSC Medicinal Chemistry**, Cambridge v. 11, p. 1275-1284, 2020. DOI: 10.1039/D0MD00097C

III. BONATTO, V.; SHAMIM, A.; ROCHO, F. R.; LEITÃO, A.; LUQUE, F. J.; LAMEIRA, J.; MONTANARI, C. A. Predicting the Relative Binding Affinity for Reversible Covalent Inhibitors by Free Energy Perturbation Calculations. **Journal of Chemical Information and Modeling**, Washington, v. 61, p. 4733-4744, 2021. DOI: 10.1021/acs.jcim.1c00515

Related articles:

LAMEIRA, J.; BONATTO, V.; CIANNI, L.; ROCHO, F. R.; LEITÃO, A.; MONTANARI, C. A. . Predicting the affinity of halogenated reversible covalent inhibitors through relative binding free energy. **Physical Chemistry Chemical Physics**, Cambridge, v. 21, p. 24723-24730, 2019.

DA COSTA, C. H. S.; BONATTO, V.; DOS SANTOS, A. M.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A. . Evaluating QM/MM Free Energy Surfaces for Ranking Cysteine Protease Covalent Inhibitors. **Journal of Chemical Information and Modeling**, Washington, v. 60, p. 880-889, 2020.

CIANNI, L.; ROCHO, F. R.; ROSINI, F.; BONATTO, V.; RIBEIRO, J.F.R.; LAMEIRA, J.; LEITÃO, A.; SHAMIM, A.; MONTANARI, C. A. . Optimization strategy of single-digit nanomolar cross-class inhibitors of mammalian and protozoa cysteine proteases. **Bioorganic Chemistry**, San Diego, v. 101, p. 104039, 2020.

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

ROCHO, F. R. ; BONATTO, V.; LAMEIRO, R. F. ; LAMEIRA, J. ; LEITÃO, A. ; MONTANARI, C. A. . A patent review on cathepsin K inhibitors to treat osteoporosis (2011 - 2021). **Expert Opinion On Therapeutic Patents**, Abingdon, v. 32, p. 1-13, 2022.

LIST OF ABBREVIATIONS AND ACRONYMS

CPs – Cysteine proteases
Cat – Cathepsin
Cz – Cruzain
SARS-CoV-2 – Acute respiratory syndrome coronavirus 2
COVID-19 – Coronavirus disease 2019
M^{pro} – Main protease of SARS-CoV-2
PL^{pro} – Papain-like protease of SARS-CoV-2
MoB – Mode of binding
MoA – Mode of action
TS – Transition state
PDB – Protein Data Bank
E – Free enzyme/apoenzyme
HIV - Human immunodeficiency virus
FDA – Food and Drug Administration
I – Inhibitor
TCIs – Targeted covalent inhibitors
DPP-4 – Dipeptidyl peptidase 4
T2DM - Type 2 diabetes mellitus
CPKIs - Covalent protein-kinase inhibitors
 τ – Residence time
NCE – New Chemical Entities
VS- Virtual Screening
DEL – DNA-Encoded Library Screenings
SBDD – Structure-Based Drug Design
CADD - Computer-Aided Drug Discovery
LBDD – Ligand-Based Drug Design
FEP – Free Energy Perturbation
MD – Molecular Dynamics
MM – Molecular mechanics
RMSD – Root Mean Square Deviation
RBF E - Relative binding free energy
ABFE - Absolute binding free energy
QM – Quantum mechanics
QSAR - Quantitative structure-activity relationship
QM/MM – Quantum mechanics/molecular mechanics
AI – Artificial intelligence
cpHMD – Constant pH molecular dynamics
EVB - Empirical valence bond

ϵ – Dielectric constant

MMP - Matched molecular pairs

SP – Single-point

ZPE – Zero-points energy

TABLE OF CONTENTS

DEDICATION.....	
ACKNOWLEDGMENT.....	
EPIGRAPH.....	
RESUMO.....	
ABSTRACT.....	
LIST OF PUBLICATIONS.....	
LIST OF ABBREVIATIONS AND ACRONYMS	
1. INTRODUCTION.....	1
1.1 Cysteine Proteases as Pharmacological Targets.....	1
1.2 Catalytic Mechanism of Cysteine Proteases	3
1.3 Reversible Covalent Inhibitors	8
1.3.1 History of the Covalent Drugs.....	8
1.3.2 Benefits and Limitations.....	12
1.3.3 Irreversible and Reversible Warheads.....	13
1.3.4 Reaction Scheme of the Covalent Inhibition and Residence Time ...	15
1.4 Computational Chemistry and Drug Design.....	18
1.4.1 Rational Drug Design.....	18
1.4.2 Computer-Aided Drug Discovery.....	20
1.4.3 Designing Reversible Covalent Inhibitors.....	23
REFERENCES.....	28
2. Published Articles and Their Context in the Doctoral Project	43
3. Nitriles: An Attractive Approach to the Development of Covalent Inhibitors.....	45
4. On the Intrinsic Reactivity of Highly Potent Trypanocidal Cruzain Inhibitors.....	63

5. QM-cluster Approach to Model the Warhead of Reversible Covalent Inhibitors	74
REFERENCES	83
6. Predicting the Relative Binding Affinity for Reversible Covalent Inhibitors by Free Energy Perturbation Calculations	86
7. CONCLUSIONS.....	99
APPENDIX A – Supporting Information for On the Intrinsic Reactivity of Highly Potent Trypanocidal Cruzain Inhibitors.....	101
APPENDIX B – Cartesian Coordinates for the QM-cluster Systems	118
APPENDIX C – Supporting Information for Predicting the Relative Binding Affinity for Reversible Covalent Inhibitors by Free Energy Perturbation Calculations	128

1. INTRODUCTION

1.1 Cysteine Proteases as Pharmacological Targets

Proteases represent a class of enzymes responsible for catalyzing the hydrolytic cleavage of peptide bonds.^{1,2} They are involved in many physiological processes and can be related to various pathological conditions and disease states when dysregulated.³ With almost 600 proteases in humans, this class of enzymes can be divided into four main subclasses based on their specificities and catalytic mechanisms: aspartic, cysteine, serine and metalloproteases.⁴ This work focuses on cysteine proteases (CPs), which possess a cysteine residue in their catalytic dyad or triad.^{1,5}

Encompassing a total of 11 members, the cathepsins (Cat) are the representative proteins in humans for the CPs family, which consists of Cat B, C, F, H, K, L, O, S, V, W and X.^{1,6} CatB and CatL are particularly noteworthy as they are widely distributed throughout our body. In contrast, the remaining cysteine cathepsins are found in more specific tissues or cell types.⁷ Many pathological conditions observed in humans are associated with the malfunctioning of these enzymes, making them attractive targets for developing of new drug candidates.^{6,8}

Lately, studies have observed increased expression of cathepsins in cancer cells, such as Cat (B, L, S, and K),^{9,10} indicating their potential involvement in neoplastic progression and cancer advancement.^{9,11} CatB has been implicated in neurodegenerative disorders such as Parkinson's and Alzheimer's. Microglia, cells responsible for defending the central nervous system, typically produce CatB and can secrete excessive amounts of this enzyme, leading to neuronal apoptosis.^{12,13} CatL can be involved in the mechanism by which coronaviruses introduce their genetic material into host cells,¹⁴ this process will be further elucidated. CatK is specifically linked to bone diseases such as osteoporosis and osteoarthritis, as it degrades collagen, a major component of the organic bone matrix.^{15,16} CatS, on the other hand, plays a crucial role in the body's immune response to tumors, promoting the activation of CD4+T cells over CD8+T cells, thereby facilitating tumor growth and increasing tumor volume.^{17,18} Cathepsins are also related to cardiovascular diseases, where patients have exhibited higher enzyme activity in the heart and arterial walls.⁷ Moreover, they contribute directly and indirectly to cardiovascular inflammation by regulating adaptive and innate

immune responses. The secretion of cathepsins is also associated with the inflammatory response in several diseases, including rheumatoid arthritis and pancreatitis.³

Expanding beyond human cysteine cathepsins, there are other biologically relevant CPs that are attractive to serve as drug targets. Parasitic and viral proteases are examples of proteins that are being targeted for the treatment of infectious diseases.¹⁹

Illustrative examples of parasitic proteases include the enzyme cruzain (Cz), a recombinant form of the enzyme cruzipain, from *Trypanosoma cruzi*, the etiological agent of Chagas disease,^{20,21} which is essential for the development and survival of the parasite inside and outside the host cells. Rhodesain is another non-human CP of interest in developing new drugs. It is the most expressed CP and a validated target for *T. brucei rhodesiense*, the causative agent of African sleeping sickness, a parasitic disease endemic to sub-Saharan Africa.²² Additionally, in the context of malaria, the parasite *Plasmodium falciparum* contains an essential CP for survival within the host cells, the falcipain-2.²³

Regarding viral proteases, the most recent and notorious example proteins are associated with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), responsible for the coronavirus disease 2019 (COVID-19) pandemic.^{24,25} The papain-like protease (PL^{pro}) and main protease (M^{pro}) are two essential enzymes associated with the release of nonstructural proteins responsible for viral replication and transcription.^{26–28} M^{pro} is one of the most explored targets in order to the development of novel antiviral drugs and so far, it has one drug that targeted it approved worldwide, which is Paxlovid (nirmatrelvir + ritonavir) and one more drug approved in Japan, ensitrelvir.^{25,27,29,30} The case of nirmatrelvir, present in paxlovid as an inhibitor of the M^{pro}, will be discussed in more detail later on.

Additionally, another strategy to impair viral replication is the inhibition of host enzymes.³¹ Regarding SARS-CoV-2, interestingly, after viral endocytosis, a human enzyme with potential inhibitory targets is CatL.^{14,32} Once the virus is present within the endosome, this cathepsin cleaves the S protein of SARS-CoV-2, releasing the viral genetic material into the host cell's cytoplasm, and initiating the mechanism for the production of new viral particles.^{14,25,33}

Thus, the inhibition of cysteine proteases can be directly related to the cure or treatment of various diseases.^{3,8,11} Therefore, due to their association with certain illnesses, CPs are highly attractive targets for discovering and developing new drug candidates. Other factors that make these enzymes appealing for the search for new chemical entities include their ubiquity in different organisms, the known 3D structures with modes of interaction (MoB) and mechanisms of action (MoA) of a myriad of inhibitors.^{4,6,34,35} Moreover, elucidating the catalytic mechanism of these enzymes also holds invaluable value for developing new inhibitors.

1.2 Catalytic Mechanism of Cysteine Proteases

The chemical processes that occur within cells are typically controlled by a type of protein called enzymes. Enzymes are considered biological catalysts, allowing specific reactions to occur at biologically feasible rates and are essential for life. They are highly substrate-specific, and therefore, the active site of enzymes is tailored to interact with a particular substrate.³⁶

This specificity for particular substrates is directly related to the stabilization of their respective transition state (TS), rather than the substrate itself. Therefore, enzymes are responsible for stabilizing the transition state of a reaction, reducing the free energy of activation. They achieve this by providing an alternative and more energetically favorable reaction pathway.³⁷

Therefore, it is through the stabilization of the substrate's TS that enzymes are able to catalyze specific reactions within biologically feasible time scales and are responsible for numerous biochemical processes in living organisms.³⁶ Thus, understanding the mechanism of enzyme action is crucial to comprehend why enzymes can exhibit greater specificity for one substrate over another.

This knowledge brings one of the most beneficial applications for human beings, the development of medications. Since enzymes catalyze a wide range of cellular processes, inhibiting them can interrupt specific enzymatic reactions, leading to a desired therapeutic effect.

Before exploring the catalytic mechanism of cysteine proteases, it is pertinent to highlight two fundamental observations that warrant attention:

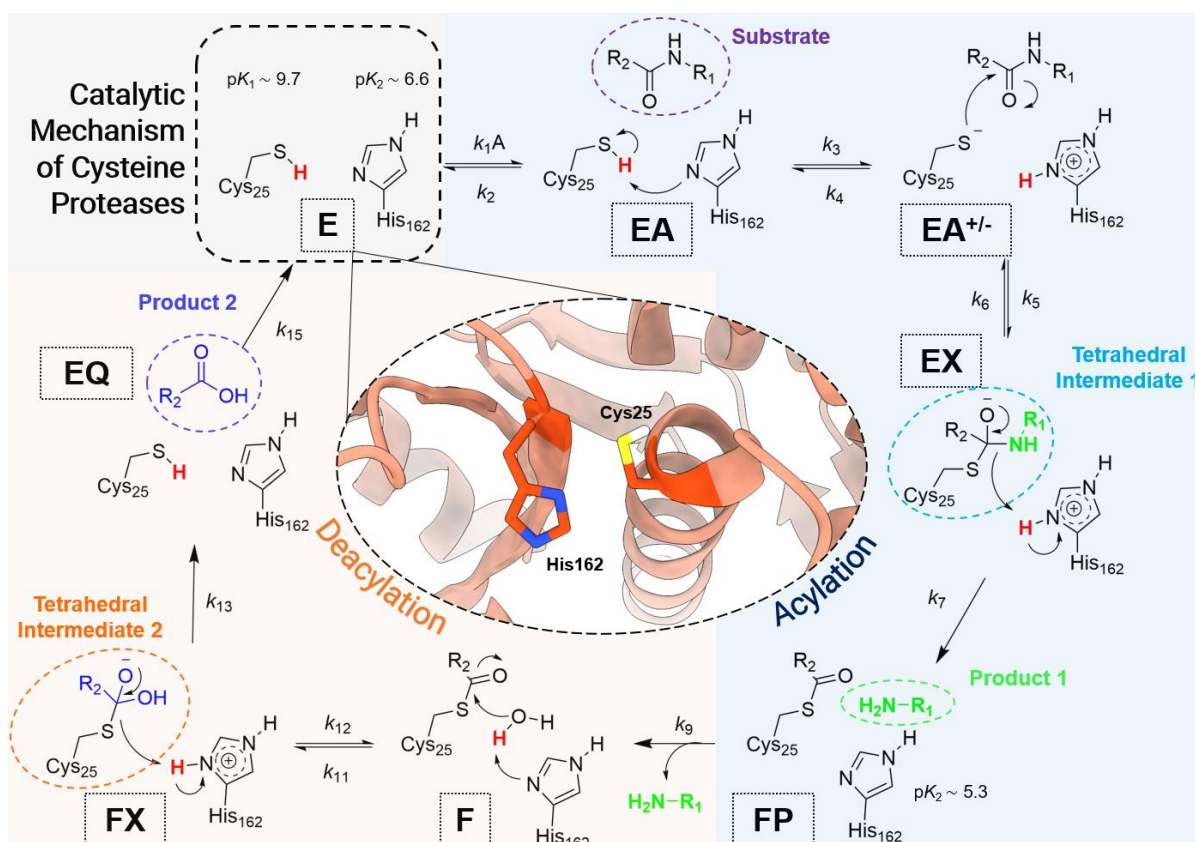
1 – Typically, CPs possess a catalytic triad consisting of Cys/His/Asn, which is essential for their proteolytic activity.^{2,38} This triad is found in human

cathepsins as well as in parasitic proteases like cruzain, falcipain, and rhodesain. However, the Asn residue is sometimes absent from the active site, and its function is compensated by a water molecule.³⁹ This water molecule effectively polarizes the His residue, enabling it to capture the proton from the Cys residue.^{39,40} A notable example is the M^{pro} of SARS-CoV-2, which exhibits a catalytic dyad of Cys/His instead of the catalytic triad.^{40,41} In the mechanism presented below, the depiction will focus solely on the direct interaction/reaction between the Cys/His residues with the substrate/inhibitor, without illustrating the involvement of the Asn residue or this water molecule;

2 – The other point concerns the timing of Cys-His ionic pair is formed in relation to ligand binding: whether it occurs before (free enzyme or apoenzyme) or after the binding of the ligand. This topic remains an open question in the literature and appears to vary across different enzymes.^{38,42–44} Zhai and Meek conducted studies on cruzain and determined the pK_a of the catalytic Cys/His residues.⁴⁵ They found that both residues in the free enzyme form exist in a neutral state. Thus, it is believed that the ligand binding induces the formation of the ion pair.^{38,45} Similar observations have been made for other CPs, as observed for Villamil and colleagues.⁴⁶ Interestingly, computational studies have also shown that the catalytic dyad Cys/His in free M^{pro} will be at a neutral state.^{47–50} However, a recent study solved the free enzyme of M^{pro} through neutron crystallography and showed that the ionic pair Cys⁻/His⁺ is present in the apoenzyme form.⁴⁴ The presence of ion pair prior to ligand binding has also been observed for papain.⁵¹ In summary, the formation of the ionic pair in cysteine proteases can occur either before or after ligand binding and varies among different enzymes.

Next, the catalytic mechanism of cysteine proteases will be explained in greater detail. Cysteine proteases possess a catalytic cysteine residue that performs a nucleophilic attack on the carbonyl carbon of a scissile peptide bond. Their catalytic mechanism is a complex process that can be divided into two stages: 1) acylation and 2) deacylation.^{38,45} The general mechanism of action for CPs is illustrated in Figure 1, using cruzain as an example, in line with the experiments performed by Zhai and Meek.⁴⁵

Figure 1 – Catalytic mechanism for CPs, the amino acid residue numbering (Cys25 and His162) is according to cruzain. The mechanism can be divided into two stages, acylation (right part with blue background) and deacylation (left part with orange background). The 3D representation in the center is the apocruzain from the PDB code: 6N3S.⁵²



Source: Image created by the author.

Herein, it is being considered that the enzyme in the free form (E) possesses the catalytic dyad, Cys25 and His162, in the neutral state, according to their pK_a values, 9.7 and 6.6, respectively.⁴⁵ The nomenclature of kinetic constants is the same used for Zhai and Meek⁴⁵ in case of interest in further reading. Accordingly, with the approach of the substrate in the active site (EA complex), the first part of the mechanism begins with the acylation of the enzyme (the right part of Figure 1, blue background). With the substrate close to the catalytic dyad, the His residue, through its imidazole group, captures the proton from Cys via a general base catalysis mechanism. Once the ionic pair $\text{Cys}^-/\text{His}^+$ is formed ($\text{EA}^{+/-}$), the cysteine acts as a strong nucleophile, utilizing the thiolate present in its side chain to perform a nucleophilic attack on the carbonyl carbon of the peptide bond.

Following this attack, the tetrahedral intermediate 1 (EX complex) is formed, wherein the oxygen previously present in the carbonyl of the scissile peptide bond becomes negatively charged and is stabilized through hydrogen bonding in a region of the enzyme known as the oxyanion hole.³⁸ Subsequently, the His⁺ residue will be deprotonated, returning to its neutral form, while the proton is transferred to the substrate, releasing the first product, H₂N-R₁.

Thus, the acylation stage of the mechanism is completed, with the formation of product 1 and the acyl-enzyme complex (FP). The next step in the mechanism involves the deacylation of the enzyme (left part of Figure 1, orange background), where the protein will be regenerated through hydrolysis.

In the first step of the deacylation half-reaction (F), once again, general base catalysis takes place, with the participation of a water molecule, which is crucial for the mechanism and is the reason that proteases are also called hydrolases. His162 will capture a proton from the water molecule, becoming positively charged (His⁺), while the resulting hydroxide (OH⁻) will act as a nucleophile, attacking the carbonyl carbon of the acyl-enzyme, forming the tetrahedral intermediate 2 (FX complex).

Similar to the first formed intermediate, this one will also be stabilized through intermolecular interactions in the oxyanion hole, where the oxygen of the former carbonyl group becomes negatively charged after the nucleophilic attack. Then, the proton from His⁺ will be transferred to the sulfur of Cys25, and in a concerted manner, it will lead to the formation and release of the second product (EQ).

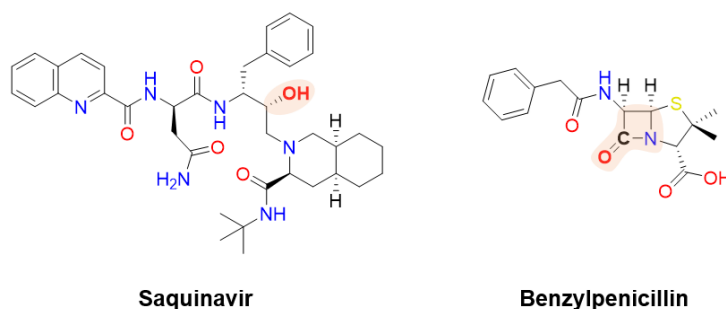
Therefore, with the release of the second product, the deacylation half-reaction of the enzyme is completed, and the catalytic dyad can be found in the neutral form again (Cys/His) in the free enzyme. Thus, with the enzyme regenerated to its initial state (E), the peptide bond hydrolysis mechanism can begin once again, involving general acid-base catalysis, covalent catalysis, and nucleophilic catalysis.

It is important to highlight that the rate-limiting step of the reaction is the deacylation, which occurs after the rapid acylation.^{38,45} Furthermore, for Cz, the Cys/His⁺ ion pair is only formed in the presence of the substrate/inhibitor in the active site.

From understanding the reaction mechanism and its associated steps, several inhibitors have been proposed for CPs over time. Many of these inhibitors have been designed to mimic the transition state and contain functional groups such as diazomethyl ketones, vinyl sulfones, haloalkyl ketones, α -ketoesters, nitriles, aldehydes, among many others, which undergo nucleophilic attack by the catalytic cysteine.¹ Hence, the tetrahedral intermediate is formed and subsequently forms a covalent adduct, either irreversibly or reversibly,^{53,54} as will be discussed in the next topic.

Remarkably, the development of new inhibitors based on the enzyme's mechanism of action is not limited to CPs.⁵⁵ For instance, by understanding the mechanism of aspartic proteases, another family of proteases, it has been possible to develop new drugs for treating human immunodeficiency virus (HIV) infections, such as Saquinavir (Figure 2),⁵⁶ the first protease inhibitor approved by the Food and Drug Administration (FDA). Researchers have designed compounds containing a strategic hydroxyl group to mimic the substrate's transition state, creating a competitive inhibitor for the enzyme's active site, thus interfering in the enzymatic proteolytic process.⁵⁷

Figure 2 – Chemical structures of Saquinavir, used as an antiretroviral drug, and benzylpenicillin, used as an antibacterial agent. The highlighted groups are essential to mimic the substrate according to the mechanism of action of the respective enzymes that these drugs inhibit.



Source: Image created by the author.

Another successful example of a class of drugs that inhibit target proteins based on the catalytic mechanism is antibiotics, one of the greatest discoveries ever made by humans. Although the first antibiotics, such as Fleming's penicillin, were serendipitously discovered, one of the most effective approaches to treating bacterial infections is through the use of β -lactam drugs (Figure 2).⁵⁸ The β -

lactam antibiotic function by inhibiting the enzyme transpeptidase, which features a catalytic Ser residue in its active site. The four-membered β -lactam ring in these drugs mimics the carbonyl group of the peptide bond in peptidoglycan, the substrate upon which transpeptidase acts.⁵⁸ By inhibiting this enzyme, bacterial cell wall formation is disrupted or compromised, resulting in cell lysis.

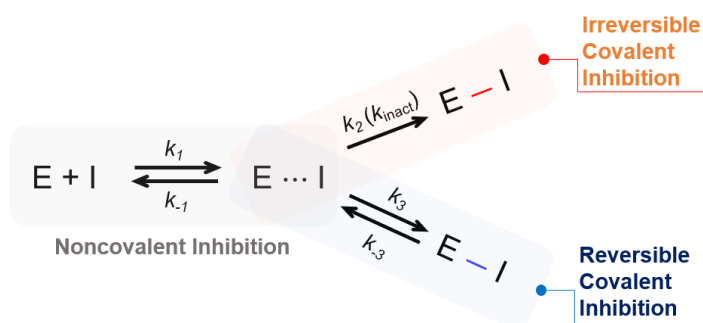
1.3 Reversible Covalent Inhibitors

1.3.1 History of the Covalent Drugs

By understanding the reaction mechanism of a particular enzyme, it is possible to develop new competitive inhibitors based on this process.⁵⁹ Generally, there are two groups of inhibitors: non-covalent and covalent.

In Scheme 1, we can distinguish these two types of inhibitors. Non-covalent inhibitors (gray in Scheme 1) do not form any chemical bond with their target enzyme; their inhibition is due to the formation of a stable enzyme-inhibitor complex through intermolecular/non-bonded interactions between the species involved ($E \cdots I$). On the other hand, in addition to intermolecular interactions, covalent inhibitors also form a chemical bond with the macromolecular target ($E-I$). The covalent inhibitors can be classified as irreversible (orange in Scheme 1) or reversible (blue in Scheme 1), depending on the reactive group used in the ligand to form the covalent bond with the protein,^{54,60} *vide infra*.

Scheme 1 – General mechanism of enzymatic inhibition between an inhibitor (I) and its target (E). Here, k_1/k_{-1} are the constants of noncovalent inhibition (gray), k_2 (k_{inact}) will guide irreversible covalent inhibition (orange), and k_3/k_{-3} will guide reversible covalent inhibition (blue).

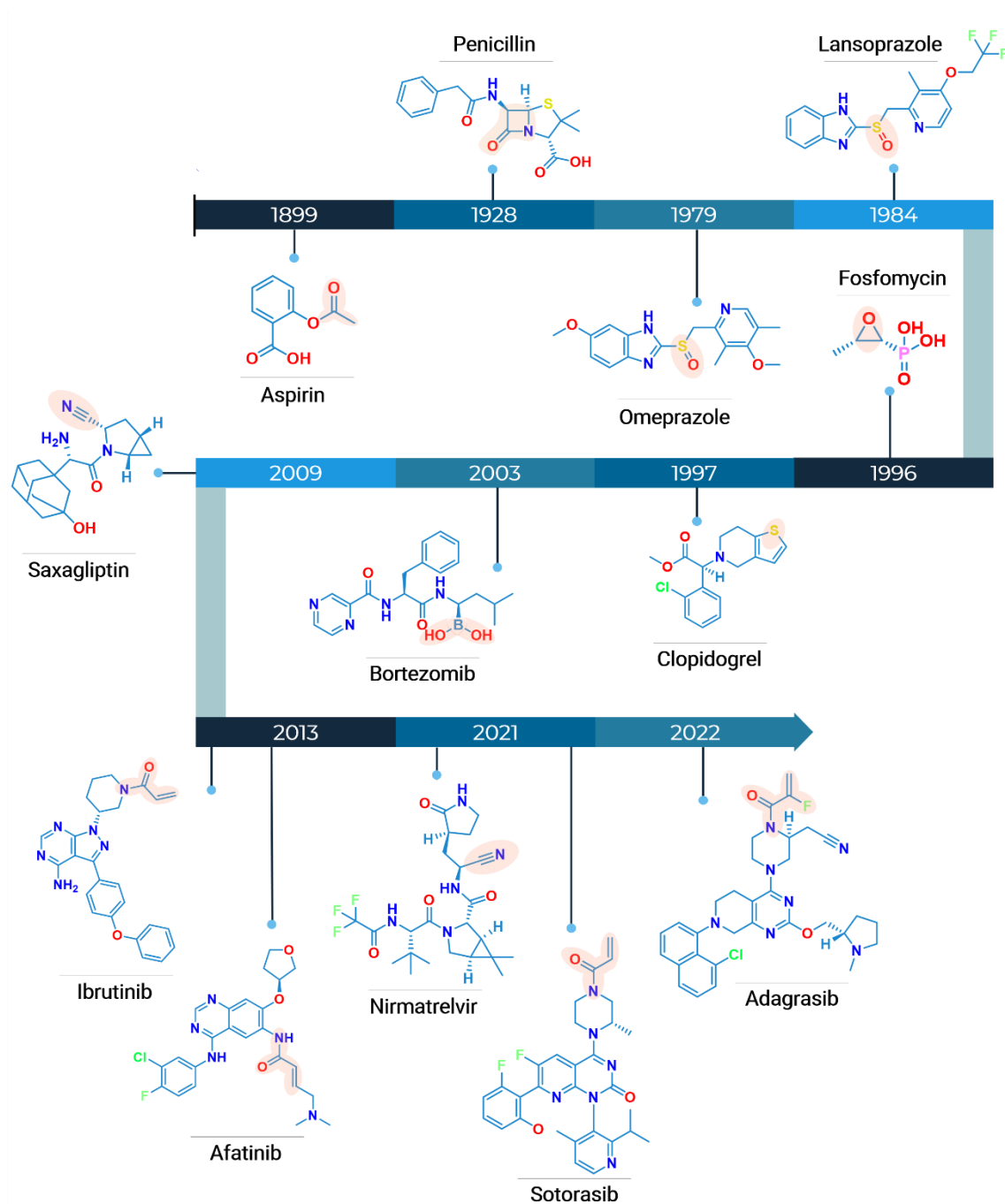


Source: Image created by the author.

Covalent inhibitors or targeted covalent inhibitors (TCIs) have been on the market for over a century and, currently, there are more than 50 FDA-approved drugs utilizing this mechanism of action.^{54,60–66} These drugs have made

significant contributions to the history of medicine and are considered milestones in the field. Figure 3 shows a timeline of the major covalent drugs.^{54,60–66}

Figure 3 – Timeline of important drugs with covalent inhibition as the mechanism of action. The reactive group or atom of each drug is highlighted.



Source: Image created by the author.

Starting with aspirin, a trademark registered by Bayer in 1899, which marked the beginning of its production and distribution.^{65,67} Although aspirin is a widely used medication with various therapeutic effects its primary mechanism of

action involves inhibiting the production of prostaglandins, which are signaling molecules involved in pain, inflammation and blood clotting processes.^{64,67,68}

A few years later, in 1928, one of the most famous and pivotal medications in our history was discovered at Hospital St. Mary in London, the aforementioned penicillin. During World War II, there was a crucial need for large-scale production of this β -lactam drug. The United States took the initiative to meet the demand and ensure an adequate supply of this life-saving medication for treating not only their soldiers but also the soldiers of allied countries.^{69,70}

In the late 1970s, a groundbreaking discovery was made: omeprazole. This proton pump inhibitor works by reducing the production of stomach acid, relieving individuals suffering from acid-related conditions and revolutionizing the treatment of gastrointestinal reflux.⁷¹ Building on this success, lansoprazole was patented shortly after, offering similar therapeutic benefits.⁶²

In 1996, a new class of antibiotics entered the market with the introduction of fosfomycin, an epoxide-containing drug.⁷² This antibiotic functions through a different mechanism than penicillin but shares the same goal of inhibiting bacterial cell wall synthesis.⁷³ Its introduction provided healthcare professionals with an additional tool for combating bacterial infections.

The year following fosfomycin, a remarkable advancement in the treatment of vascular disorders occurred with the introduction of the anticoagulant clopidogrel. It offers improved management of conditions such as heart attacks and strokes.^{62,74} Interestingly, these drugs with the sulfur atom in their scaffolds (omeprazole, lansoprazole and clopidogrel) are prodrugs. Upon metabolism, their metabolites are thiol-based compounds that permanently create disulfide bonds to inactivate the macromolecular target.^{65,71,74}

Moving to the first decade of the 21st century, the discovery of bortezomib provided another option for cancer treatment. Bortezomib is a proteasome inhibitor that induces programmed cell death in cancer cells.⁷⁵ And in 2009, the antidiabetic drug saxagliptin began to be marketed for the medication of type 2 diabetes mellitus (T2DM).^{76,77} With the successful design of saxagliptin, other nitrile-based covalent inhibitors of dipeptidyl peptidase 4 (DPP-4) were developed and approved worldwide (e.g. vildagliptin, anagliptin and trelagliptin).⁷⁸

In the following decade, the most fruitful period of covalent inhibitors began with significant advancements in oncology treatments through kinase

inhibition.^{79–81} In 2013, ibrutinib and afatinib emerged as groundbreaking covalent protein-kinase inhibitors (CPKIs), revolutionizing targeted therapies for cancer.^{54,79} These two pioneering drugs paved the way for the development of six additional CPKIs that have since gained FDA approval.^{79,80} These eight CPKIs represent 11.3% of all FDA-approved drugs that act by inhibiting kinases (71 drugs).⁸²

Furthermore, countless lives were lost during the arduous battle against COVID-19 in recent years, underscoring the urgent need for effective treatments. While vaccines were developed at an unprecedented pace, Pfizer's scientists took an interesting approach by revisiting an old peptidomimetic drug candidate initially intended for treating SARS-CoV-1.^{25,83} Recognizing its potential, they repurposed this candidate to develop a new antiviral drug named nirmatrelvir, which MoA consists of impairing viral replication in host cells through the inhibition of the M^{pro}.^{84,85} Nirmatrelvir is marketed under the trademark name paxlovid, a combination of nirmatrelvir and ritonavir. The latter is necessary to inhibit the metabolization of nirmatrelvir by CYP3A4.⁸⁶ The fantastic success of paxlovid is evident in its revenue. In 2022, paxlovid ranked among the top four best-selling pharmaceutical products, generating a staggering \$18.933 billion in sales.^{87,88} Furthermore, the projected sales for paxlovid in 2023 are expected to surpass the \$15 billion mark.⁸⁹

Finally, to conclude the timeline of noteworthy covalent drugs, notable advancements include sotorasib and adagrasib for the treatment of solid tumors.^{90,91} These two breakthrough drugs are known to target “undruggable” proteins, since the inhibition of mutant KRAS (G12C), a GTPase and member of the RAS family, faced 40 years of inefficiency until the covalent approach.^{65,92} They target the Cys12 of KRAS (G12C) to form the covalent bond and then inactivate the enzyme, reshaping cancer treatment strategies.⁹²

An intriguing aspect of medications that contain TCIs as their active substances is that a significant portion of them was discovered serendipitously rather than through a rational project involving the inhibition of a specific target with a covalent bond formation.^{62,65,93} In numerous covalent drugs discovered before 2000, their precise mechanism of action remained undisclosed for many years following the medication's introduction to the market. For instance, it was not until the 1970s that the main therapeutic effect of aspirin, involving

cyclooxygenase 1 inactivation through the acetylation of Ser529 of the enzyme, was fully understood.^{81,94} A similar situation occurred with penicillin and omeprazole, where the elucidation of their respective mechanisms of action came several years after their initial discovery, despite recognizing the drugs' beneficial properties.

Interestingly, until the recent success of saxagliptin (as well as the other “gliptins”) and the CPKIs (mainly ibrutinib), the preference has always been for the use of non-covalent inhibitors.^{95,96} This preference stemmed from the potential complications associated with the reactive group of covalent compounds, as it could potentially bind to unintended human proteins apart from the specific target. Therefore, despite the historical use of TCIs, the deliberate development of drugs that engage in a chemical bond with a specific target is a relatively recent advancement. Of course, like any approach, it has both positive and negative aspects.

1.3.2 Benefits and Limitations

Considering these two categories of inhibitors (non-covalent and covalent), there are several advantages and disadvantages associated with each of them. One significant benefit of covalent inhibitors, in contrast to non-covalent inhibitors, is their prolonged therapeutic response.^{54,64} The chemical bond formed by covalent inhibitors with the enzyme allows for an extended duration of the ligand-target interaction, ensuring a longer-lasting desired effect.⁹⁷ This effect is related to the longer residence time (τ) of the reversible covalent inhibitors and the re-synthesis rate of the enzyme for the irreversible ones,⁵⁴ as will be discussed further in the next topic.

Furthermore, covalent drugs offer the advantage of higher potency and efficiency, allowing for lower dosages to achieve the desired therapeutic effect.^{54,62} The enhanced potency and efficiency of covalent ligands are interconnected, resulting in the additional benefit of reduced dosing frequency. This feature can improve patient tolerability and compliance to covalent drugs.^{62,65}

Another notable advantage of covalent inhibitors is the reduced potential for developing resistance mechanisms.^{54,62} TCIs require substantial modifications to the enzyme's catalytic site or the emergence of entirely new enzymes to

inactive the covalent inhibitor. This occurrence, as demonstrated by the emergence of β -lactamases in bacteria,⁵⁸ is infrequent and may require a significant amount of time to occur.

Finally, with the advent of TCIs, it became possible to target enzymes that were once deemed "undruggable", as in the case of the sotorasib and adagrasib to inhibit the mutant KRAS (G12C).^{63,92} These inhibitors have expanded the scope of therapeutic possibilities by effectively engaging with challenging enzyme targets that were previously considered inaccessible or difficult to target.

However, despite the numerous advantages, inhibitors that form a chemical bond with the target enzyme have some potential issues, particularly regarding the reactivity of the group that undergoes the reaction with the macromolecule.^{54,62} This reactivity may lead to off-target effects, which have historically hindered the exploration of ligands in this class for a significant period of time.⁹⁵

This disadvantage, for instance, can result in the permanent inactivation of unintended enzymes, leading to idiosyncratic toxicity. Another potential issue is the risk of protein haptization, which can activate the immune system and potentially trigger an adverse immune response in humans. Additionally, there is also the possibility of inhibitor depletion by glutathione (GSH).^{54,62,63}

However, it is possible to mitigate the aforementioned potential issues with the proper selection of the reactive group in the covalent inhibitor. This requires a meticulous choice of the warhead that will be integrated into the covalent inhibitor's structure and react with the target enzyme.

1.3.3 Irreversible and Reversible Warheads

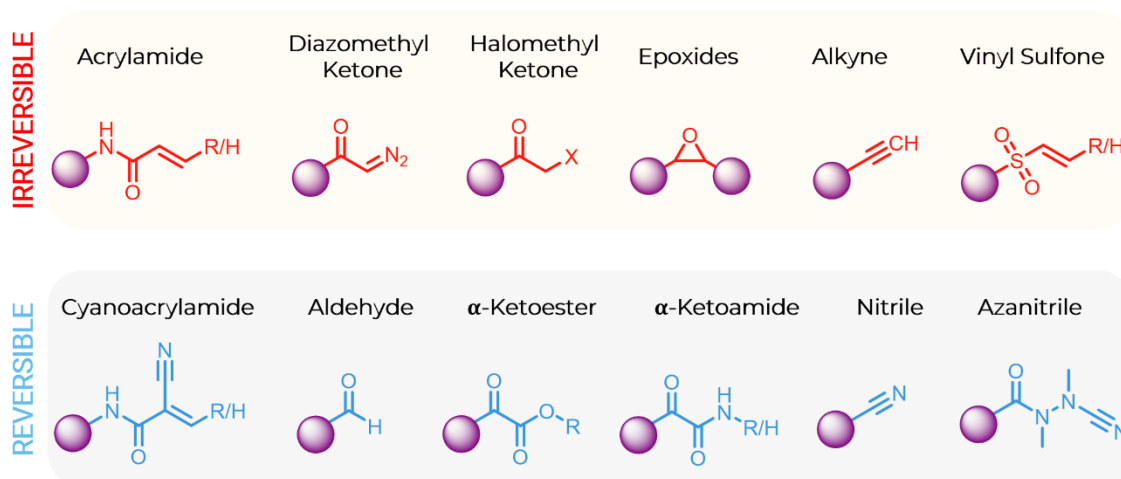
The group in the inhibitor that will react with the target enzyme is the warhead. This critical group within an inhibitor has high electrophilicity and is therefore suitable to participate in a nucleophilic attack by the enzyme.⁹⁸

Warheads can bind to the enzyme either reversibly or irreversibly, depending on their electrophilicity.^{55,98} The reactivity of the warhead plays a crucial role in determining how easily it forms a bond with the enzyme.^{99,100} For instance, in the case of irreversible inhibitors, the inhibitor forms an irreversible bond with the enzyme's nucleophilic group, resulting in the permanent inactivation of the target. The formation of this reversibility of the reaction can be

determined by the energy barrier of the reaction and the stability of the resulting adduct.^{101,102}

Groups such as α,β -unsaturated amide (acrylamide), diazomethyl ketone, halomethyl ketones ($X = \text{Cl}, \text{Br}$), epoxide, alkyne (terminal) and vinyl sulfone are known to form an irreversible bond (more exothermic reaction) with the target macromolecule. They are represented in red in Figure 4.^{98–100,103} Also, some groups bind to the enzyme reversibly. Examples of such groups include α,β -unsaturated cyanoamide (cyanoacrylamide), aldehyde, α -ketoester, α -ketoamide, nitrile and azanitrile, which form reversible bonds with the macromolecule target and are highlighted in blue in Figure 4.^{98–100}

Figure 4 – Examples of warheads responsible for irreversible (red) and reversible (blue) binding with the target enzyme.



Source: Image created by the author.

One question that may arise is: when to choose irreversible warheads and when to choose reversible warheads. The decision between these two types of warheads in covalent inhibition depends on various considerations. Irreversible warheads are typically preferred when sustained and permanent inactivation of the target enzyme is desired.^{104,105} This approach is suitable for situations where long-lasting inhibition is necessary for therapeutic efficacy. On the other hand, reversible warheads offer the advantage of flexibility and adjustability in modulating enzyme activity and are also related to lower toxicity issues.^{98,106}

In addition, another important factor to consider when choosing the type of warhead is the ease of modeling. Irreversible warheads pose a challenge in terms

of modeling due to the direct relationship between the transition state and the kinetic rate constant (k_{inact}).^{61,102,107} Modeling this class of inhibitors requires a more complex approach. In contrast, reversible covalent inhibitors can be more readily modeled using X-ray co-crystallized structures of the inhibitor with the target protein. Given this, and due to the toxicity effects, the reversible warheads should be of preference when possible.

However, there are situations where the use of a reversible warhead may not be suitable for the target protein. In such cases, an irreversible warhead becomes necessary. For instance, more reactive and electrophilic warheads are typically required when the target amino acid is not a catalytic residue, such as an allosteric Cys or a non-catalytic Cys in the active site.^{106,108} This is exemplified by the use of irreversible warheads in inhibitors like ibrutinib, which targets a non-catalytic Cys, and in the case of KRAS G12C inhibitors, which target a Cys residue in an allosteric site. These examples highlight the need for irreversible warheads in specific scenarios where reversible warheads may not be practical or feasible. Nonetheless, recent works by Tauton's and Ojida's groups^{106,108–111} overcome this issue with the design of very reactive reversible warheads (e.g. cyanoacrylamide and dihaloacetamide) and these moieties are valuable options to target non-catalytic Cys residues without using irreversible warheads.

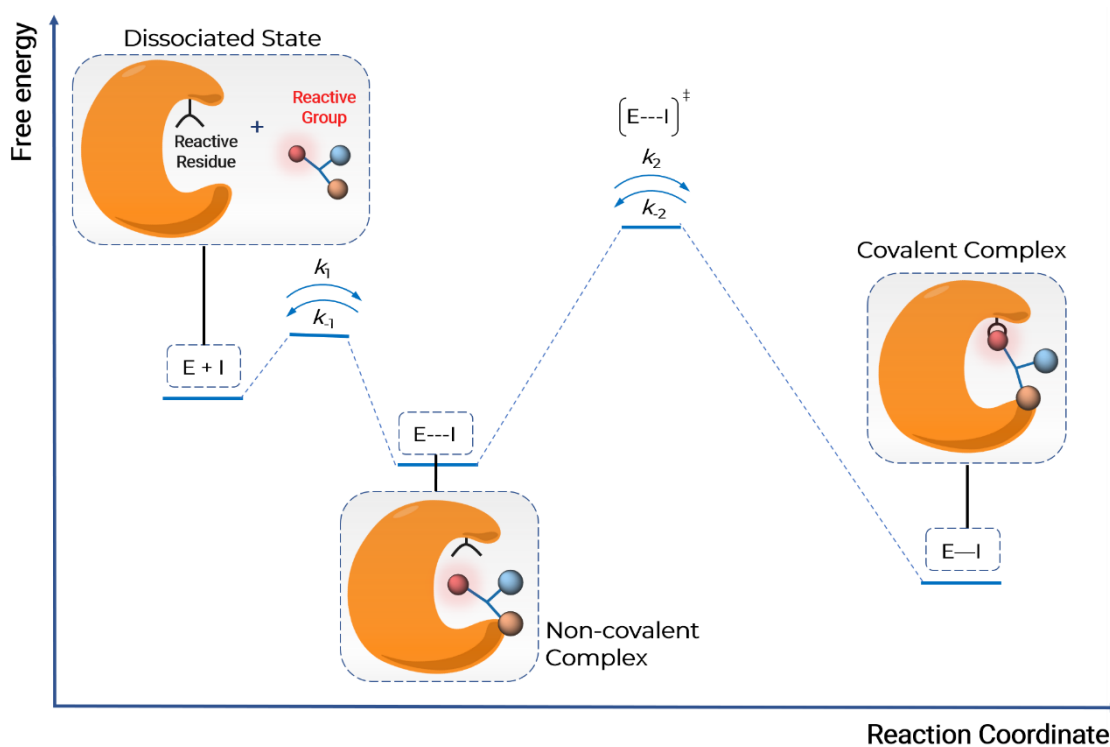
Overall, the choice of the appropriate warhead relies on several factors, including enhancing specificity for the target enzyme, minimizing off-target effects, achieving the desired duration of inhibition, and aligning with the therapeutic goals of the treatment.^{99,100} By carefully considering these factors, researchers can optimize the design of covalent inhibitors and maximize their potential for therapeutic success.

1.3.4 Reaction Scheme of the Covalent Inhibition and Residence Time

The process of covalent inhibition involves two steps, as depicted in Figure 5 for a nucleophilic attack of a reactive residue (e.g. Cys or Ser) on a TCI. Initially, after the dissociated state, the non-covalent complex (also called the Michaelis complex) allows the inhibitor's reactive electrophilic group to be positioned in close proximity to the catalytic nucleophilic group of the enzyme in the active site. Subsequently, the covalent complex is formed through the nucleophilic attack of

the reactive residue of the protein, resulting in the formation of the inhibitor-enzyme adduct (covalent complex).^{112,113} The rate constants k_2 will govern the reversibility of the reaction and factors such as the acidity of the α -proton in the adduct, conformation of the covalent complex and the noncovalent interactions may also influence the rate of the reverse reaction. Hydrolysis is another possible factor that can contribute to the cleavage of the adduct formed, contributing to the reversible mechanism.^{54,61,108}

Figure 5 - Representation of the reaction scheme involving a ligand possessing a reactive group (I) with a target enzyme (E) containing a reactive residue to perform the nucleophilic attack. The rate constant k_2 will determine the reversibility of the covalent adduct [E-I] to be irreversible or reversible.



Source: Image created by the author.

In general, when the adduct is highly stable, the reaction cannot proceed in the reverse direction, and the energy barrier of the reverse reaction cannot be overcome. As a result, irreversible inhibition occurs, leading to the inactivation of the enzyme. In contrast, when the covalent complex is reversibly inhibited, the energy barrier can be overcome, allowing the reversible inhibitor to dissociate from the enzyme and restore its activity.¹⁰¹ Therefore, the residence time (τ) or dissociation rate is strictly related to the energy of the reverse reaction barrier (k).

2). Residence time refers to the duration of a compound remaining bound within a binding site.^{97,114} Non-covalent inhibitors typically have a short residence time, ranging from seconds to a few hours, while reversible covalent inhibitors have a longer residence time, usually up to 10 hours.¹¹⁴ However, there are exceptional cases of reversible TCIs with residence time exceeding 6 days.¹⁰⁸

Since the residence time is associated with the covalent bond reversibility (eq. 1), when determining the value of k_{-2} it is possible to calculate the reverse reaction barrier (ΔG^\ddagger), according to the Eyring-Polanyi equation (eq. 2).

$$\tau = (k_{-2})^{-1} \quad (1)$$

$$k = \frac{\kappa k_B T}{h} e^{-\frac{\Delta G^\ddagger}{RT}} \quad (2)$$

Herein, the k in eq. 2 will be the rate constant k_{-2} , κ is the transmission coefficient, k_B is the Boltzmann constant, T is the temperature, h is the Planck's constant and R is the universal gas constant.

Considering that typically the reversible covalent inhibitors present $\tau = 10\text{h}$,^{81,114} they are not likely to exceed the value of around 23.5 kcal.mol⁻¹ for the reverse reaction barrier. Under computational studies, investigations employing multiscale simulations have successfully determined reverse reaction barrier values below 23 kcal.mol⁻¹ for reversible inhibition, while values higher than 23 kcal.mol⁻¹ are observed for irreversible inhibition.^{101,115} Thus, if the energy barrier for the reverse reaction exceeds around 23 kcal.mol⁻¹, it is expected that enzyme inactivation will be achieved through an irreversible inhibition mechanism.

Consequently, once the enzyme is inactivated, the residence time becomes less relevant for the therapeutic effect since the compound will remain permanently bound to the enzyme. Instead, the activity of an irreversible inhibitor *in vitro* or *in vivo* studies will be closely linked to the re-synthesis rate.⁵⁴ The re-synthesis rate refers to the rate at which a cell or organism synthesizes new functional proteins to replace those that have been inactivated.

All these factors and a scheme reaction of a nitrile-based inhibitor are presented and further discussed in the results sections.⁸³

1.4 Computational Chemistry and Drug Design

1.4.1 Rational Drug Design

The discovery and development of new drugs is a laborious and time-consuming process, involving various fields of study. Recent studies estimate that it takes around eight years and costs millions to billions of dollars to find a new drug, depending on the therapeutic area.^{116,117}

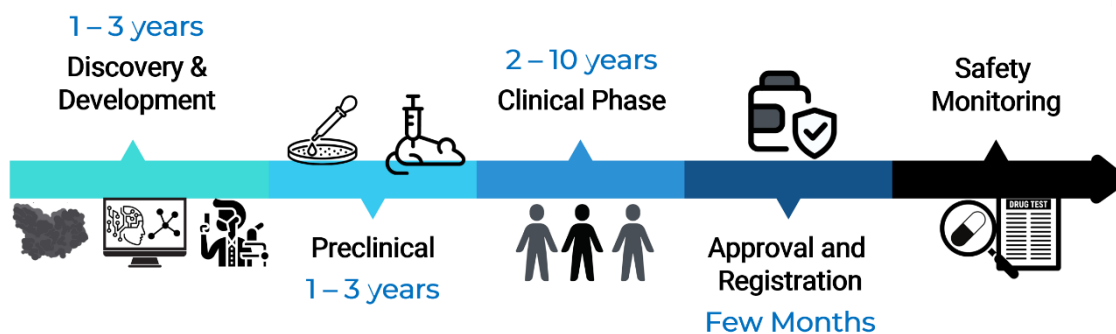
The classical stages of drug development can be summarized in Figure 6.^{118,119} It begins with the identification and validation of a macromolecular target for a specific disease, followed by the discovery and optimization of new inhibitors. The most promising candidates are then synthesized and tested in enzymatic assays. This initial stage typically takes 1 to 3 years.

Those compounds showing promising results move on to the preclinical phase, which lasts 1 to 3 years. During this phase, *in vitro* assays are conducted, followed by *in vivo* testing using animal models. Eventually, the molecules that have the potential to become drugs, exhibiting desirable pharmacokinetic and pharmacodynamic properties, are selected to proceed to clinical trials.

The stage involving human testing is the most time-consuming and costly, accounting for about 50 to 60% of the project's budget.^{116,120} Clinical trials are divided into three phases: Phase 1 to assess safety, Phase 2 to evaluate the efficacy, and Phase 3 to test these factors on a larger scale. Following successful clinical trials, the drug candidate undergoes registration and approval by the local regulatory agency before it can be marketed.

Once approved and with the drug already in circulation, the final stage of drug development begins, known as Phase 4 clinical trials. During this phase, safety monitoring of the new product is conducted to detect any possible side effects or adverse events that may not have been previously observed.

Figure 6 – Stages of drug development. A process that takes years and involves multimillion or even billion-dollar investments to reach a new product.



Source: Image created by the author.

In the initial stage of the process depicted in Figure 6, the main objective is to obtain active new chemical entities (NCEs) to develop lead compounds. Many strategies can be employed to identify new lead candidates, including starting from known compounds, random screening, directed screening, fragment screening, virtual screening (VS) and DNA-encoded library screening (DEL).¹²¹

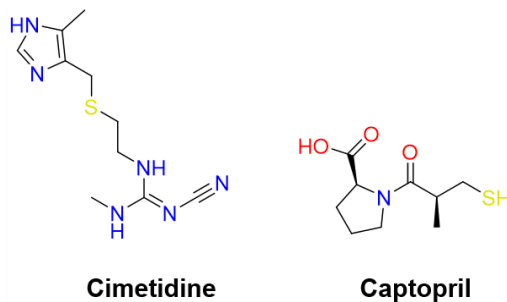
However, this rational process to obtain NCEs was not always like this. Until the early 1970s, this process was done differently, often through serendipity from plant-derived products or microbial extracts.^{122,123} The turning point in drug design came with the discovery of cimetidine (Figure 7), a project that started in mid-1964 and ended in 1976 with the commercialization of the drug.¹²⁴

This drug was planned from the discovery of a new target (whose 3D structure was unknown) to the development of various ligands in order to arrive at the best option. Cimetidine solved a problem for which there was no solution, as it was recommended for the treatment of gastric ulcers and became the best-selling and most prescribed drug in the world within a few years, until omeprazole, shown in Figure 3, took its place.^{123,124} For the study and development of cimetidine, Sir James Black was awarded the Nobel Prize in Physiology or Medicine in 1988.¹²⁵

Another rational approach to the discovery of new inhibitors emerged with the advancements in X-ray crystallography techniques, where new projects began to be carried out based on the structure of the macromolecular target, known as structure-based drug design (SBDD).¹²⁶

In this context, in the 1980s, computer-aided structure-activity relationship studies led to the discovery of captopril, Figure 7.¹²⁷ This is considered by many as the first drug discovered with the assistance of a computer, marking the beginning of a new era in the discovery and development of new drugs, known as computer-aided drug discovery (CADD).^{127,128}

Figure 7 – Chemical structures of cimetidine and captopril.



Source: Image created by the author.

1.4.2 Computer-Aided Drug Discovery

Several computational methods can be utilized to discover NCEs and perform analysis between the structures of these new compounds and their biological activity. Nowadays, it is possible to state that computational methods are indispensable in finding or optimizing new compounds.

This process can be divided into two main areas: SBDD and ligand-based drug design (LBDD). In SBDD, the structure of the macromolecular target is utilized to guide the design of new compounds. On the other hand, LBDD focuses solely on the chemical structures of the ligands to identify potential lead candidates.^{129,130}

By employing these tactics, particularly when combined (SBDD + LBDD), the success in developing new inhibitors can be significantly enhanced. The integrated approach of SBDD and LBDD allows for a more comprehensive understanding of the molecular target and the chemical properties of the ligands, facilitating the identification of promising compounds with high affinity and selectivity for the desired target.^{129–132}

A wide range of computational methods can be applied to identify new active compounds or optimize known ones. The choice of which technique to use

depends on the specific problem to be addressed and the feasible computational cost for that task.

Large-scale virtual screening and the creation of target-focused virtual libraries are commonly employed in campaigns to discover new inhibitors. Additionally, predicting the binding modes of inhibitors in proteins is a common practice, often accomplished through molecular docking and molecular dynamics (MD) simulations.^{130,133,134}

These physics-based methods, such as molecular mechanics (MM), utilize force fields to describe the atoms in the system, enabling the capture of intra- and intermolecular interactions.¹³⁵ Molecular docking, in particular, offers the advantage of generating high-quality binding poses for ligands within macromolecular targets, with root mean square deviation (RMSD) values of up to 2 Å when compared to co-crystallized ligands.^{136,137}

Furthermore, molecular docking offers the advantage of being highly efficient and extremely fast in generating binding poses, making it particularly useful for docking ultra-large libraries that contain billions of compounds, allowing for comprehensive coverage of specific chemical spaces.^{138–140} Each pose generated by docking is associated with a score; however, these scores often deviate significantly from the experimentally observed values. This discrepancy arises due to the limitations of docking methods in considering important factors such as protein flexibility, proper entropy description, and solvent effects.^{136,141,142}

Nevertheless, recent research efforts have aimed to overcome these limitations and enhance the accuracy of docking scores.^{143,144} Despite these ongoing challenges, docking remains a valuable tool for rapidly and efficiently generating poses.

In contrast, molecular dynamics simulations involve solvating the entire complex and allowing it to move dynamically over time based on the forces described by the chosen force field.^{130,135} Therefore, by taking into account these missing thermodynamic terms in the scoring functions, such as protein motion, entropic considerations, explicit water molecules, and other factors like desolvation effects and ligand strain, MD-based methods for obtaining binding free energy are considered the most rigorous and suitable among all-atom methods with explicit solvent.^{131,142,145}

These MD-based methods for obtaining the free energy of binding are also commonly referred to as alchemical, perturbative, or free energy perturbation (FEP) methods.^{131,142,145} Since free energy is a state function, FEP-based methods utilize a thermodynamic cycle to connect two distinct physical states through intermediate non-physical states. This approach enables the determination of the free energy for a given molecule or a set of molecules based on a predefined thermodynamic cycle. By carefully designing the cycle and simulating the required transformations, FEP methods allow for the calculation of the desired free energy values.^{131,142,145}

The calculation of free energy can be approached in two ways: relative (RBFE – relative binding free energy) and absolute (ABFE – absolute binding free energy).¹³¹ In relative calculations, the difference in free energy between a target molecule and a reference molecule is determined. On the other hand, absolute calculations involve the computation of the free energy for a single molecule independently.

However, all these improvements and benefits come at a much higher computational cost compared to methods such as molecular docking. Moreover, physics-based methods are not suitable for describing processes involving electron transfer and chemical reactions. Thus, in certain situations, quantum mechanics-based (QM) methods are required to accurately capture these effects with high accuracy.^{61,107}

Nonetheless, due to their significant computational demands, QM-based methods are generally avoided in drug discovery projects and are typically restricted to a limited number of atoms. Still, there are situations where QM-based methods can be of interest, such as when determining quantum parameters for developing quantitative structure-activity relationship (QSAR) models.^{61,107} In this approach, specific descriptors are derived using QM techniques, focusing solely on the studied ligand.

Another approach to incorporate quantum methods is through multiscale or hybrid calculations (QM/MM), which involve describing part of the system with molecular mechanics and a specific region of the system with quantum mechanics.¹⁴⁶ In this method, the active/catalytic site of the enzyme is treated with quantum mechanics to capture electron transfer and bond formation/breakage accurately. At the same time, the rest of the system is

described using molecular mechanics to account for the protein environment and explicit solvent.¹⁴⁶ This allows for a more detailed and accurate representation of the system, considering both quantum effects and the larger molecular environment.

Nevertheless, QM/MM calculations are still computationally expensive for drug discovery projects. An alternative approach that has been developed is the QM clusters, wherein only a small portion of the protein-ligand system is explicitly described (typically around ~300 atoms), while the remaining system is represented using a dielectric constant to approximate the protein environment.^{147–149} This strategy allows for a more computationally feasible calculation by focusing on the critical interactions within the active site.

This type of methodology tends to be faster than QM/MM simulations and has been successfully applied to describe enzymatic catalysis reactions.^{147,149–152} However, its use for the development of new inhibitors is still in its early stages and has rarely been employed for this purpose.

Furthermore, the emerging use of artificial intelligence (AI) in drug discovery holds tremendous promise.^{153,154} AI can play a significant role in identifying potential targets for specific diseases, predicting three-dimensional protein structures such as AlphaFold,¹⁵⁵ and generating and selecting new inhibitors for specific targets.¹⁵⁶ Various AI-based methodologies can be employed, either focusing solely on the ligand or in combination with the target protein, to create robust and effective models for drug discovery.¹⁵⁷ The AI has the advantage that it can be very fast after the model is trained, but its limitation is the necessary amount of data to create new models.¹⁵⁴

The application of generative AI models for generating novel molecules, coupled with affinity prediction models (regression or classification), has the potential to revolutionize the process of discovering new inhibitors.¹⁵⁸ This has already been demonstrated in the accelerated development of new antibiotics and compounds targeting idiopathic pulmonary fibrosis, where AI has played a crucial role.^{159–161}

1.4.3 Designing Reversible Covalent Inhibitors

In this way, many computational methods can be helpful in the development of new covalent inhibitors. Schirmeister and colleagues published

a protocol that focuses on the rational design of covalent inhibitors, emphasizing the significance of quantum methods in treating the warhead moiety.¹⁶² In a different study, they employed QM-based methods in the gas phase to comprehend the reactivity differences among warheads based on the reaction mechanism for cysteine and serine/threonine residues.¹⁶³ Additionally, covalent docking was utilized to predict the MoBs for the compounds under investigation.

Other studies have also suggested several quantum methods and QM descriptors to assess and predict the reactivity of warheads. Furthermore, Lameira et al. have consistently demonstrated the significance of QM/MM methods in modeling and/or explaining the differences in inhibitor potency with different warheads.^{101,164} Notably, through QM/MM methodology, it was possible to distinguish reversible covalent inhibition from irreversible inhibition based on differences in reaction energy.⁵³ Hence, the significance of quantum methods becomes apparent when modeling the warhead in covalent ligands.

Similarly, Zhou et al. presented a workflow for the development of covalent inhibitors targeting non-catalytic Cys residues in kinases.¹⁶⁵ A critical aspect in this context is the determination of pK_a values for the target Cys residues to assess their potential reactivity through deprotonation. Constant pH molecular dynamics (cpHMD) was highlighted as a possible approach to determine the pK_a values of amino acids with reactive tendencies within an enzyme. In a recent study, Awoonor-Williams et al. benchmarked in silico tools for pK_a determination of Cys residues and also highlighted cpHMD as a viable option, although it was not the best methodology employed by them.¹⁶⁶ Once confirmed that the residue could exhibit reactivity, Zhou et al. utilized covalent docking in their workflow to guide the selection of compounds for synthesis and subsequent experimental testing.¹⁶⁵

However, once the warhead is defined and it is known that the target residue is capable of performing a nucleophilic attack (e.g., catalytic Cys), various methods can be employed to enhance the affinity of a reversible covalent ligand for the target by optimizing non-covalent interactions. Unlike irreversible inhibitors, reversible covalent ligands are not heavily reliant on the transition state for modulating their activity, making their optimization a relatively simpler task.^{102,105} In this regard, physics-based methods provide valuable tools for this optimization process.

Covalent docking has proven to be valuable in predicting and optimizing the intermolecular interactions between reversible covalent inhibitors and macromolecular targets¹⁶⁷. However, although covalent docking programs generate excellent poses, the reliability of the scoring function is still a concern.¹⁴¹ Therefore, a combination of covalent docking and molecular dynamics simulations is recommended to confirm the stability of the generated pose.

Continuing, perturbative methods for calculating free energy have emerged as a powerful tool for determining the relative binding free energy of covalent inhibitors.¹⁴⁵ Kuhn pioneered the application of this methodology in reversible TCIs in 2017, focusing on CatL inhibitors.¹⁶⁸ In this study, only the covalent complex was considered for performing mutations in the molecules and obtaining the relative binding free energy among the 10 molecules under investigation.

Also, in 2017, the research group led by Prof. Luo at the Western University of Health Sciences demonstrated that it is possible to obtain the relative binding free energy for a series of covalent inhibitors by separately considering the non-covalent and covalent states, as well as predict selectivity towards specific targets.¹⁶⁹ Nevertheless, to obtain a more accurate estimation of the binding free energy for covalent inhibitors, it would be necessary to consider and combine the values obtained for both states (non-covalent, ΔG_{dm} and covalent, ΔG_{dc}), according to the following equations:

$$\Delta G_{dc} = \Delta \Delta G_{dc} + \Delta G_{dc}^{core} \quad (3)$$

$$\Delta G_{dm} = \Delta \Delta G_{dm} + \Delta G_{dm}^{core} \quad (4)$$

$$\frac{1}{K_d} = e^{-\frac{(\Delta \Delta G_{dc} + \Delta G_{dc}^{core})}{RT}} + e^{-\frac{(\Delta \Delta G_{dm} + \Delta G_{dm}^{core})}{RT}} \quad (5)$$

However, obtaining the binding free energy of the warhead (ΔG^{core}) presents some challenges. For the non-covalent state, FEP methods can be employed, but for the covalent state, quantum-based methodologies are necessary. In the same study, it was mentioned that if the contribution of the covalent state to the free energy is at least -5.5 kcal/mol greater than the non-covalent state, it is possible to use only the covalent state for the calculations.¹⁶⁹

However, it would still be necessary to perform calculations for both states, which can be computationally demanding.

In another study conducted by Prof. Luo's group, they employed a specific thermodynamic cycle to overcome the computational challenges associated with QM-based methods, allowing them to determine the absolute binding free energy for a series of reversible covalent inhibitors.¹⁶⁷ By combining the FEP values obtained for the non-covalent and covalent states, they were able to achieve accurate results. However, it should be noted that this thermodynamic cycle may not be valid if the warhead is different between the compounds.

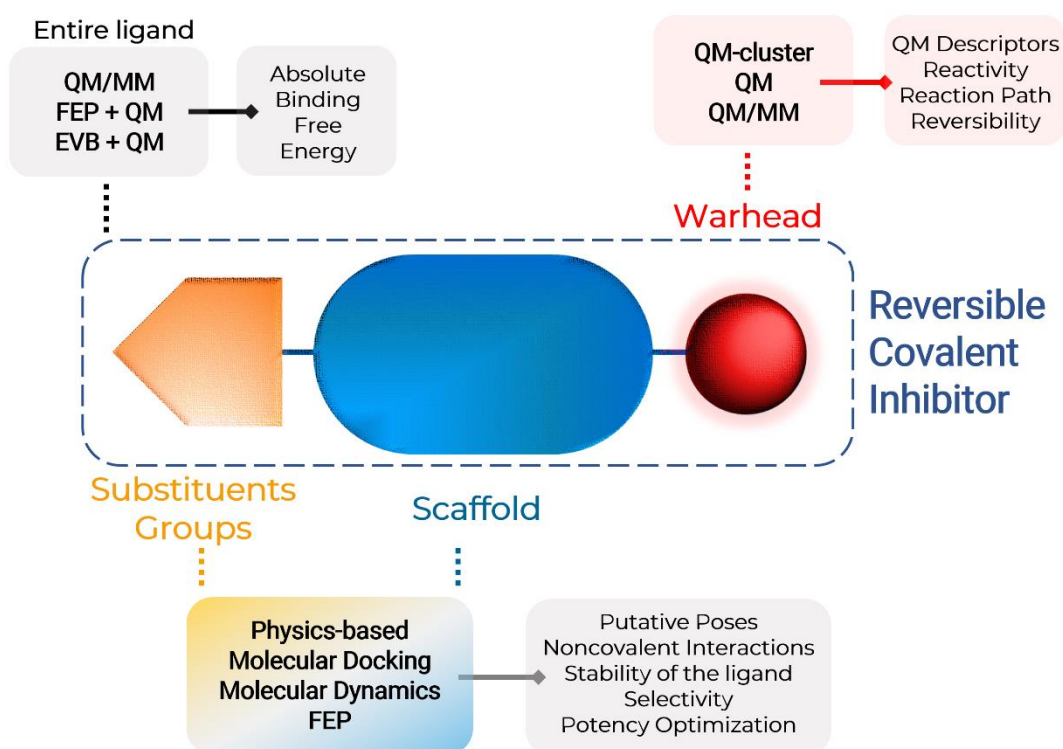
Our research group also applied the FEP methodology to investigate the potency difference among a series of halogenated-based reversible covalent inhibitors for CatL, and we obtained satisfactory results by considering the contributions from both the non-covalent and covalent states.¹⁷⁰

More recently, Prof. Warshel also embarked on the quest to determine the absolute binding free energy for reversible covalent inhibitors. For this purpose, he utilized the empirical valence bond (EVB) method to obtain absolute values for the non-covalent interactions between the ligands and the studied proteins.¹⁷¹ Additionally, simple QM calculations in the gas phase were employed to determine the energy values associated with the warhead ($\Delta G_{\text{cov}} = E_{\text{product}} - E_{\text{reactants}}$). The isolated ligand and SHCH₃ were considered reactants in the gas phase, and the formed covalent adduct was considered the product. By combining the results, it was possible to obtain satisfactory absolute energy values for the ligands, with errors of up to 2.3 kcal/mol when compared to experimental values.¹⁷¹

Overall, physics-based methodologies are great choices for modeling non-covalent interactions and accurately assessing energy differences between similar molecules.^{61,107} But, studying covalent inhibitors through FEP can be computationally expensive since it requires calculations in both the non-covalent and covalent states.

With this in mind, when modeling reversible covalent inhibitors, the choice of methodologies will depend on the specific objectives, and different approaches can yield excellent results. Figure 8 summarizes the computational calculations that can be employed to model reversible covalent inhibitors.

Figure 8 – Computational approaches for modeling reversible covalent inhibitors. These methodologies play a crucial role in understanding the behavior and interactions of reversible covalent inhibitors, aiding in the design and optimization of these compounds.



Source: Image created by the author.

These computational methodologies can be beneficial in reversible covalent inhibitor development projects. Determining the overall project objective is essential to select the most suitable method for a given task. By employing these methodologies, researchers can gain valuable insights into the behavior of reversible covalent inhibitors and accelerate drug discovery projects.

REFERENCES

- 1 CIANNI, L.; FELDMANN, C. W.; GILBERG, E.; GÜTSCHOW, M.; JULIANO, L.; LEITÃO, A.; BAJORATH, J.; MONTANARI, C. A. Can cysteine protease cross-class inhibitors achieve selectivity? **Journal of Medicinal Chemistry**, Washington, v. 62, n. 23, p. 10497–10525, 2019.
- 2 EREZ, E.; FASS, D.; BIBI, E. How intramembrane proteases bury hydrolytic reactions in the membrane. **Nature**, Berlin, v. 459, p. 371–378, 2009.
- 3 KRAMER, L.; TURK, D.; TURK, B. The future of cysteine cathepsins in disease management. **Trends in Pharmacological Sciences**, London, v. 38, n. 10, p. 873–898, 2017.
- 4 LECAILLE, F.; KALETA, J.; BROMME, D. Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. **Chemical Reviews**, Washington, v. 102, p. 4459–4488, 2002.
- 5 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, p. 1–3, 2018.
- 6 PATEL, S.; HOMAIEI, A.; EL-SEEDI, H. R.; AKHTAR, N. Cathepsins: proteases that are vital for survival but can also be fatal. **Biomedicine & Pharmacotherapy**, Issy-les-Moulineaux, v. 105, p. 526–532, 2018.
- 7 LIU, C. L.; GUO, J.; ZHANG, X.; SUKHOVA, G. K.; LIBBY, P.; SHI, G.-P. Cysteine protease cathepsins in cardiovascular disease: from basic research to clinical trials. **Nature Reviews Cardiology**, Berlin, v. 15, n. 6, p. 351–370, 2018.
- 8 TURK, B. Targeting proteases: successes, failures and future prospects. **Nature Reviews Drug Discovery**, Berlin, v. 5, n. 9, p. 785–799, 2006.
- 9 MOHAMED, M. M.; SLOANE, B. F. Cysteine cathepsins: multifunctional enzymes in cancer. **Nature Reviews Cancer**, Berlin, v. 6, p. 764–775, 2006.
- 10 SIKLOS, M.; BENAÏSSA, M.; THATCHER, G. R. J. Cysteine proteases as therapeutic targets: does selectivity matter? a systematic review of calpain and cathepsin inhibitors. **Acta Pharmaceutica Sinica B**, Beijing, v. 5, n. 6, p. 506–519, 2015.
- 11 RAGHAV, N.; JANGRA, S.; KUMAR, A.; BHATTACHARYYA, S.; WADHWA, D.; SINDHU, J. Cathepsin B, H and L inhibitors as cell proliferating agents: design, synthesis, computational and pharmacological studies of some novel 2-(2-naphthoyl)-6,6-dimethyl-3-aryl-2,3,6,7-tetrahydrobenzofuran-4(5H)-ones. **RSC Advances**, Cambridge, v. 21, p. 34588–34599, 2016.
- 12 BERNSTEIN, H. G.; KEILHOFF, G. Putative roles of cathepsin B in Alzheimer's disease pathology: the good, the bad, and the ugly in one? **Neural Regeneration Research**, Mumbai, v. 13, n. 12, p. 2100–2101, 2018.
- 13 NAKANISHI, H. Microglial cathepsin B as a key driver of inflammatory brain diseases and brain aging. **Neural Regeneration Research**, Mumbai, v. 15, n. 1, p. 25–29, 2020.

- 14 LIU, T.; LUO, S.; LIBBY, P.; SHI, G. P. Cathepsin L-selective inhibitors: a potentially promising treatment for COVID-19 patients. **Pharmacology & Therapeutics**, Oxford, v. 213, p. 107587, 2020.
- 15 ROCHO, F. R.; BONATTO, V.; LAMEIRO, R. F.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A. A patent review on cathepsin K inhibitors to treat osteoporosis (2011 – 2021). **Expert Opinion on Therapeutic Patents**, Oxon, v. 32, n. 5, p. 561–573, 2022.
- 16 LU, J.; WANG, M.; WANG, Z.; FU, Z.; LU, A.; ZHANG, G. Advances in the discovery of cathepsin K inhibitors on bone resorption. **Journal of Enzyme Inhibition and Medicinal Chemistry**, Oxon, v. 33, n. 1, p. 890–904, 2018.
- 17 BARARIA, D.; HILDEBRAND, J. A.; STOLZ, S.; HAEBE, S.; ALIG, S.; TREVISANI, C. P.; OSORIO-BARRIOS, F.; BARTOSCHEK, M. D.; MENTZ, M.; PASTORE, A.; GAITZSCH, E.; HEIDE, M.; JURINOVIC, V.; RAUTTER, K.; GUNAWARDANA, J.; SABDIA, M. B.; SZCZEPANOWSKI, M.; RICHTER, J.; KLAPPER, W.; LOUISSAINT, A.; LUDWIG, C.; BULTMANN, S.; LEONHARDT, H.; EUSTERMANN, S.; HOPFNER, K.-P.; HIDDEMANN, W.; VON BERGWELT-BAILDON, M.; STEIDL, C.; KRIDEL, R.; TOBIN, J. W. D.; GANDHI, M. K.; WEINSTOCK, D. M.; SCHMIDT-SUPPRIAN, M.; SÁROSI, M. B.; RUDELIUS, M.; PASSERINI, V.; MAUTNER, J.; WEIGERT, O. Cathepsin S alterations induce a tumor-promoting immune microenvironment in follicular lymphoma. **Cell Reports**, Cambridge, v. 31, n. 5, p. 107522, 2020.
- 18 DHEILLY, E.; BATTISTELLO, E.; KATANAYEVA, N.; SUNGALEE, S.; MICHAUX, J.; DUNS, G.; WEHRLE, S.; SORDET-DESSIMOZ, J.; MINA, M.; RACLE, J.; FARINHA, P.; COUKOS, G.; GFELLER, D.; MOTTOK, A.; KRIDEL, R.; CORREIA, B. E.; STEIDL, C.; BASSANI-STERNBERG, M.; CIRIELLO, G.; ZOETE, V.; ORICCHIO, E. Cathepsin S regulates antigen processing and T cell activity in non-Hodgkin lymphoma. **Cancer Cell**, Cambridge, v. 37, n. 5, p. 674-689, 2020.
- 19 AGBOWURO, A. A.; HUSTON, W. M.; GAMBLE, A. B.; TYNDALL, J. D. A. Proteases and protease inhibitors in infectious diseases. **Medicinal Research Reviews**, Hoboken, v. 38, n. 4, p. 1295–1331, 2018.
- 20 MCKERROW, J. H. Update on drug development targeting parasite cysteine proteases. **PLoS Neglected Tropical Diseases**, San Francisco, v. 12, p. 1–4, 2018.
- 21 CIANNI, L.; SARTORI, G.; ROSINI, F.; VITA, D. De; PIRES, G.; REBELO, B.; LEITÃO, A.; BURTOLOSO, A. C. B.; MONTANARI, C. A. Leveraging the cruzain S3 subsite to increase affinity for reversible covalent inhibitors. **Bioorganic & Medicinal Chemistry**, Oxford, v. 79, p. 285–292, 2018.
- 22 ETTARI, R.; TAMBORINI, L.; ANGELO, I. C.; MICALE, N.; PINTO, A.; DE MICHELI, C.; CONTI, P. Inhibition of rhodesain as a novel therapeutic modality for human African trypanosomiasis. **Journal of Medicinal Chemistry**, Washington, v. 56, n. 14, p. 5637–5658, 2013.
- 23 PANDEY, K. C.; DIXIT, R. Structure-function of falcipains: malarial cysteine proteases. **Journal of Tropical Medicine**, London, v. 2012, p. 345195, 2012.

- 24 HOLMES, E. C.; GOLDSTEIN, S. A.; RASMUSSEN, A. L.; ROBERTSON, D. L.; CRITS-CHRISTOPH, A.; WERTHEIM, J. O.; ANTHONY, S. J.; BARCLAY, W. S.; BONI, M. F.; DOHERTY, P. C.; FARRAR, J.; GEOGHEGAN, J. L.; JIANG, X.; LEIBOWITZ, J. L.; NEIL, S. J. D.; SKERN, T.; WEISS, S. R.; WOROBEY, M.; ANDERSEN, K. G.; GARRY, R. F.; RAMBAUT, A. The origins of SARS-CoV-2: A critical review. **Cell**, Cambridge, v. 184, n. 19, p. 4848–4856, 2021.
- 25 LI, G.; HILGENFELD, R.; WHITLEY, R.; DE CLERCQ, E. Therapeutic strategies for COVID-19: Progress and lessons learned. **Nature Reviews Drug Discovery**, Berlin, v. 22, n. 6, p. 449–475, 2023.
- 26 LIU, C.; ZHOU, Q.; LI, Y.; GARNER, L. V.; WATKINS, S. P.; CARTER, L. J.; SMOOT, J.; GREGG, A. C.; DANIELS, A. D.; JERVEY, S.; ALBAIU, D. Research and development on therapeutic agents and vaccines for Covid-19 and related human coronavirus diseases. **ACS Central Science**, Washington, v. 6, n. 3, p. 315–331, 2020.
- 27 PANG, X.; XU, W.; LIU, Y.; LI, H.; CHEN, L. The research progress of SARS-CoV-2 main protease inhibitors from 2020 to 2022. **European Journal of Medicinal Chemistry**, Issy-les-Moulineaux, v. 257, p. 115491, 2023.
- 28 ZHANG, L.; LIN, D.; SUN, X.; CURTH, U.; DROSTEN, C.; SAUERHERING, L.; BECKER, S.; ROX, K.; HILGENFELD, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved a-ketoamide inhibitors. **Science**, Washington, v. 368, n. 6489, p. 409–412, 2020.
- 29 XIONG, M.; SU, H.; ZHAO, W.; XIE, H.; SHAO, Q.; XU, Y. What coronavirus 3C-like protease tells us: From structure, substrate selectivity, to inhibitor design. **Medicinal Research Reviews**, Hoboken, v. 41, n. 4, p. 1965–1998, 2021.
- 30 TAN, H.; HU, Y.; JADHAV, P.; TAN, B.; WANG, J. Progress and challenges in targeting the SARS-CoV-2 papain-like protease. **Journal of Medicinal Chemistry**, Washington, v. 65, n. 11, p. 7561–7580, 2022.
- 31 MAZZON, M.; MARSH, M. Targeting viral entry as a strategy for broad-spectrum antivirals. **F1000Research**, Oxon, v. 8, n. 1628, 2019.
- 32 OU, X.; LIU, Y.; LEI, X.; LI, P.; MI, D.; REN, L.; GUO, L.; GUO, R.; CHEN, T.; HU, J.; XIANG, Z.; MU, Z.; CHEN, X.; CHEN, J.; HU, K.; JIN, Q.; WANG, J.; QIAN, Z. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. **Nature Communications**, Berlin, v. 11, n. 1620, p. 1–12, 2020.
- 33 SACCO, M. D.; MA, C.; LAGARIAS, P.; GAO, A.; TOWNSEND, J. A.; MENG, X.; DUBE, P.; ZHANG, X.; HU, Y.; KITAMURA, N.; HURST, B.; TARBET, B.; MARTY, M. T.; KOLOCOURIS, A.; XIANG, Y.; CHEN, Y.; WANG, J. Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against M^{pro} and cathepsin L. **Science Advances**, Washington, v. 6, n. 50, eabe0751, 2020.
- 34 PALERMO, C.; JOYCE, J. A. Cysteine cathepsin proteases as pharmacological targets in cancer. **Trends in Pharmacological Sciences**, London, v. 29, p. 22–28, 2007.
- 35 OTTO, H. H.; SCHIRMEISTER, T. Cysteine proteases and their inhibitors. **Chemical Reviews**, Washington, v. 97, n. 1, p. 133–171, 1997.

- 36 NELSON, D. L.; COX, M. M. **Lehninger principles of biochemistry**. 6th ed. New York: W.H. Freeman and Company, 2013.
- 37 HAYNIE, D. T. **Biological thermodynamics**. 2nd ed. Cambridge: Cambridge University Press, 2008.
- 38 OANCA, G.; ASADI, M.; SAHA, A.; RAMACHANDRAN, B.; WARSHEL, A. Exploring the catalytic reaction of cysteine proteases. **The Journal of Physical Chemistry B**, Washington, v. 124, n. 50, p. 11349–11356, 2020.
- 39 GORBALENYA, A. E.; SNIJDER, E. J. Viral cysteine proteinases. **Perspectives in Drug Discovery and Design**, Dordrecht, v. 6, n. 1, p. 64–86, 1996.
- 40 KNELLER, D. W.; PHILLIPS, G.; O'NEILL, H. M.; JEDRZEJCZAK, R.; STOLS, L.; LANGAN, P.; JOACHIMIAK, A.; COATES, L.; KOVALEVSKY, A. Structural plasticity of SARS-CoV-2 3CL M^{pro} active site cavity revealed by room temperature X-ray crystallography. **Nature Communications**, Berlin, v. 11, n. 1, p. 3202, 2020.
- 41 GRIGORENKO, B. L.; POLYAKOV, I. V.; KHRENOVA, M. G.; GIUDETTI, G.; FARAJI, S.; KRYLOV, A. I.; NEMUKHIN, A. V. Multiscale simulations of the covalent inhibition of the SARS-CoV-2 main protease: Four compounds and three reaction mechanisms. **Journal of the American Chemical Society**, Washington, v. 145, n. 24, p. 13204–13214, 2023.
- 42 SHOKHEN, M.; KHAZANOV, N.; ALBECK, A. Challenging a paradigm: theoretical calculations of the protonation state of the Cys25-His159 catalytic diad in free papain. **Proteins: Structure, Function, and Bioinformatics**, Hoboken, v. 77, n. 4, p. 916–926, 2009.
- 43 HENDERSON, J. A.; VERMA, N.; HARRIS, R. C.; LIU, R.; SHEN, J. Assessment of proton-coupled conformational dynamics of SARS and MERS coronavirus papain-like proteases: Implication for designing broad-spectrum antiviral inhibitors. **The Journal of Chemical Physics**, Melville, v. 153, n. 11, p. 115101, 2020.
- 44 KNELLER, D. W.; PHILLIPS, G.; WEISS, K. L.; ZHANG, Q.; COATES, L.; KOVALEVSKY, A. Direct observation of protonation state modulation in SARS-CoV-2 main protease upon inhibitor binding with neutron crystallography. **Journal of Medicinal Chemistry**, Washington, v. 64, n. 8, p. 4991–5000, 2021.
- 45 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, 2018.
- 46 VILLAMIL, M. A.; CHEN, J.; LIANG, Q.; ZHUANG, Z. A noncanonical cysteine protease USP1 is activated through active site modulation by USP1-associated factor 1. **Biochemistry**, Washington, v. 51, n. 13, p. 2829–2839, 2012.
- 47 RAMOS-GUZMÁN, C. A.; RUIZ-PERNÍA, J. J.; TUÑÓN, I. Inhibition mechanism of SARS-CoV-2 main protease with ketone-based inhibitors unveiled by multiscale simulations: Insights for improved designs. **Angewandte Chemie International Edition**, Weinheim, v. 60, n. 49, p. 25933–25941, 2021.

- 48 MARTÍ, S.; ARAFET, K.; LODOLA, A.; MULHOLLAND, A. J.; ŚWIDEREK, K.; MOLINER, V. Impact of warhead modulations on the covalent inhibition of SARS-CoV-2 M^{pro} explored by QM/MM simulations. **ACS Catalysis**, Washington, v. 12, n. 1, p. 698–708, 2022.
- 49 ARAFET, K.; SERRANO-APARICIO, N.; LODOLA, A.; MULHOLLAND, A. J.; GONZÁLEZ, F. V.; ŚWIDEREK, K.; MOLINER, V. Mechanism of inhibition of SARS-CoV-2 M^{pro} by N3 peptidyl Michael acceptor explained by QM/MM simulations and design of new derivatives with tunable chemical reactivity. **Chemical Science**, Cambridge, v. 12, n. 4, p. 1433–1444, 2021.
- 50 CHAN, H. T. H.; MOESSER, M. A.; WALTERS, R. K.; MALLA, T. R.; TWIDALE, R. M.; JOHN, T.; DEEKS, H. M.; JOHNSTON-WOOD, T.; MIKHAILOV, V.; SESSIONS, R. B.; DAWSON, W.; SALAH, E.; LUKACIK, P.; STRAIN-DAMERELL, C.; OWEN, C. D.; NAKAJIMA, T.; ŚWIDEREK, K.; LODOLA, A.; MOLINER, V.; GLOWACKI, D. R.; SPENCER, J.; WALSH, M. A.; SCHOFIELD, C. J.; GENOVESE, L.; SHOEMARK, D. K.; MULHOLLAND, A. J.; DUARTE, F.; MORRIS, G. M. Discovery of SARS-CoV-2 M^{pro} peptide inhibitors from modelling substrate and ligand binding. **Chemical Science**, Cambridge, v. 12, n. 41, p. 13686–13703, 2021.
- 51 LEWIS, S. D.; JOHNSON, F. A.; SHAFER, J. A. Potentiometric determination of ionizations at the active site of papain. **Biochemistry**, Washington, v. 15, n. 23, p. 5009–5017, 1976.
- 52 BARBOSA DA SILVA, E.; DALL, E.; BRIZA, P.; BRANDSTETTER, H.; FERREIRA, R. S. Cruzain structures: apocruzain and cruzain bound to S-methyl thiomethanesulfonate and implications for drug design. **Acta Crystallographica Section F Structural Biology Communications**, Chester, v. 75, n. 6, p. 419–427, 2019.
- 53 DOS SANTOS, A. M.; OLIVEIRA, A. R. S.; DA COSTA, C. H. S.; KENNY, P. W.; MONTANARI, C. A.; VARELA, J. de J. G.; LAMEIRA, J. Assessment of reversibility for covalent cysteine protease inhibitors using quantum mechanics/molecular mechanics free energy surfaces. **Journal of Chemical Information and Modeling**, Washington, v. 62, n. 17, p. 4083–4094, 2022.
- 54 SINGH, J.; PETTER, R. C.; BAILLIE, T. A.; WHITTY, A. The resurgence of covalent drugs. **Nature Reviews Drug Discovery**, Berlin, v. 10, n. 4, p. 307–317, 2011.
- 55 BAILLIE, T. A. Targeted covalent inhibitors for drug design. **Angewandte Chemie International Edition**, Weinheim, v. 55, n. 43, p. 13408–13421, 2016.
- 56 VELLA, S. Update on a proteinase inhibitor. **AIDS**, Philadelphia, v. 8, 1994.
- 57 GHOSH, A. K.; OSSWALD, H. L.; PRATO, G. Recent progress in the development of HIV-1 protease inhibitors for the treatment of HIV/AIDS. **Journal of Medicinal Chemistry**, Washington, v. 59, n. 11, p. 5172–5208, 2016.
- 58 MORA-OCHOMOGO, M.; LOHANS, C. T. β -Lactam antibiotic targets and resistance mechanisms: from covalent inhibitors to substrates. **RSC Medicinal Chemistry**, Cambridge, v. 12, n. 10, p. 1623–1639, 2021.
- 59 HOLDGATE, G. A.; MEEK, T. D.; GRIMLEY, R. L. Mechanistic enzymology in drug discovery: a fresh perspective. **Nature Reviews Drug Discovery**, Berlin, v. 17, n. 2, p. 115–132, 2018.

- 60 TULEY, A.; FAST, W. The taxonomy of covalent inhibitors. **Biochemistry**, Washington, v. 57, n. 24, p. 3326–3337, 2018.
- 61 AWOONOR-WILLIAMS, E.; WALSH, A. G.; ROWLEY, C. N. Modeling covalent-modifier drugs. **Biochimica et Biophysica Acta: proteins and proteomics**, Amsterdam, v. 1865, n. 11, Part B, p. 1664–1675, 2017.
- 62 SUTANTO, F.; KONSTANTINIDOU, M.; DÖMLING, A. Covalent inhibitors: a rational approach to drug discovery. **RSC Medicinal Chemistry**, Cambridge, v. 11, n. 8, p. 876–884, 2020.
- 63 GHOSH, A. K.; SAMANTA, I.; MONDAL, A.; LIU, W. R. Covalent inhibition in drug discovery. **ChemMedChem**, Weinheim, v. 14, n. 9, p. 889–906, 2019.
- 64 KIM, H.; HWANG, Y. S.; KIM, M.; PARK, S. B. Recent advances in the development of covalent inhibitors. **RSC Medicinal Chemistry**, Cambridge, v. 12, n. 7, p. 1037–1045, 2021.
- 65 BOIKE, L.; HENNING, N. J.; NOMURA, D. K. Advances in covalent drug discovery. **Nature Reviews Drug Discovery**, Berlin, v. 21, n. 12, p. 881–898, 2022.
- 66 SINGH, J. The ascension of targeted covalent inhibitors. **Journal of Medicinal Chemistry**, Washington, v. 65, n. 8, p. 5886–5901, 2022.
- 67 MONTINARI, M. R.; MINELLI, S.; DE CATERINA, R. The first 3500 years of aspirin history from its roots – a concise summary. **Vascular Pharmacology**, New York, v. 113, p. 1–8, 2019.
- 68 CHANDRASEKHARAN, N.; SIMMONS, D. L. The cyclooxygenases. **Genome Biology**, London, v. 5, n. 9, p. 241, 2004.
- 69 AMERICAN CHEMICAL SOCIETY INTERNATIONAL HISTORIC CHEMICAL LANDMARKS. Discovery and Development of Penicillin. **American Chemical Society**, 1999. Disponível em: <http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/flemingpenicillin.html>. Acesso em: 04 jul. 2023.
- 70 MOHR, K. I. History of antibiotics research. In: STADLER, M.; DERSCH, P. (ed.) **How to overcome the antibiotic crisis: facts, challenges, technologies and future perspectives**. New York: Springer Cham, 2016. p. 237–272.
- 71 OLBE, L.; CARLSSON, E.; LINDBERG, P. A proton-pump inhibitor expedition: the case histories of omeprazole and esomeprazole. **Nature Reviews Drug Discovery**, Berlin, v. 2, n. 2, p. 132–139, 2003.
- 72 DIJKMANS, A. C.; ZACARÍAS, N. V. O.; BURGGRAAF, J.; MOUTON, J. W.; WILMS, E.; VAN NIEUWKOOP, C.; TOUW, D. J.; STEVENS, J.; KAMERLING, I. M. C. Fosfomicin: pharmacological, clinical and future perspectives. **Antibiotics**, Basel, v. 6, n. 4, p. 24, 2017.
- 73 KAHAN, F. M.; KAHAN, J. S.; CASSIDY, P. J.; KROPP, H. The mechanism of action of fosfomicin (phosphonomycin). **Annals of the New York Academy of Sciences**, Hoboken, v. 235, n. 1, p. 364–386, 1974.

- 74 SAVI, P.; PEREILLO, J. M.; UZABIAGA, M. F.; COMBALBERT, J.; PICARD, C.; MAFFRAND, J. P.; PASCAL, M.; HERBERT, J. M. Identification and biological activity of the active metabolite of clopidogrel. **Thrombosis and Haemostasis**, Stuttgart, v. 84, n. 11, p. 891–896, 2000.
- 75 CURRAN, M. P.; MCKEAGE, K. Bortezomib: a review of its use in patients with multiple myeloma. **Drugs**, Auckland, v. 69, n. 7, p. 859–888, 2009.
- 76 VAN DER VEKEN, P.; HAEMERS, A.; AUGUSTYNS, K. Prolyl peptidases related to dipeptidyl peptidase IV: potential of specific inhibitors in drug discovery. **Current Topics in Medicinal Chemistry**, Sharjah, v. 7, p. 621–635, 2007.
- 77 METZLER, W. J.; YANCHUNAS, J.; WEIGELT, C.; KISH, K.; KLEI, H. E.; XIE, D.; ZHANG, Y.; CORBETT, M.; TAMURA, J. K.; HE, B.; HAMANN, L. G.; KIRBY, M. S.; MARCINKEVICIENE, J. Involvement of DPP-IV catalytic residues in enzyme-saxagliptin complex formation. **Protein science**, Hoboken, v. 17, n. 2, p. 240–250, 2008.
- 78 FOLEY, J. E.; AHRÉN, B. The vildagliptin experience - 25 years since the initiation of the novartis glucagon-like peptide-1 based therapy programme and 10 years since the first vildagliptin registration. **European Journal of Endocrinology**, Bristol, v. 13, n. 2, p. 56–61, 2017.
- 79 COHEN, P.; CROSS, D.; JÄNNE, P. A. Kinase drug discovery 20 years after imatinib: progress and future directions. **Nature Reviews Drug Discovery**, Berlin, v. 20, n. 7, p. 551–569, 2021.
- 80 FERGUSON, F. M.; GRAY, N. S. Kinase inhibitors: the road ahead. **Nature Reviews Drug Discovery**, Berlin, v. 17, n. 5, p. 353–377, 2018.
- 81 ABDELDAYEM, A.; RAOUF, Y. S.; CONSTANTINESCU, S. N.; MORIGGL, R.; GUNNING, P. T. Advances in covalent kinase inhibitors. **Chemical Society Reviews**, Cambridge, v. 49, n. 9, p. 2617–2687, 2020.
- 82 XERXA, E.; MILJKOVIĆ, F.; BAJORATH, J. Data-driven global assessment of protein kinase inhibitors with emphasis on covalent compounds. **Journal of Medicinal Chemistry**, Washington, v. 66, n. 11, p. 7657–7665, 2023.
- 83 BONATTO, V.; LAMEIRO, R. F.; ROCHO, F. R.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A. Nitriles: an attractive approach to the development of covalent inhibitors. **RSC Medicinal Chemistry**, Cambridge, v. 14, n. 2, p. 201–217, 2023.
- 84 OWEN, D. R.; ALLERTON, C. M. N.; ANDERSON, A. S.; ASCHENBRENNER, L.; AVERY, M.; BERRITT, S.; BORAS, B.; CARDIN, R. D.; CARLO, A.; COFFMAN, K. J.; DANTONIO, A.; DI, L.; ENG, H.; FERRE, R.; GAJIWALA, K. S.; GIBSON, S. A.; GREASLEY, S. E.; HURST, B. L.; KADAR, E. P.; KALGUTKAR, A. S.; LEE, J. C.; LEE, J.; LIU, W.; MASON, S. W.; NOELL, S.; NOVAK, J. J.; OBACH, R. S.; OGILVIE, K.; PATEL, N. C.; PETTERSSON, M.; RAI, D. K.; REESE, M. R.; SAMMONS, M. F.; SATHISH, J. G.; SINGH, R. S. P.; STEPPAN, C. M.; STEWART, A. E.; TUTTLE, J. B.; UPDYKE, L.; VERHOEST, P. R.; WEI, L.; YANG, Q.; ZHU, Y. An oral SARS-CoV-2 M^{pro} inhibitor clinical candidate for the treatment of COVID-19. **Science**, v. 374, n. 6575, p. 1586–1593, 2021.

- 85 HALFORD, B. The path to paxlovid. **ACS Central Science**, Washington, v. 8, n. 4, p. 405–407, 2022.
- 86 LAMB, Y. N. Nirmatrelvir plus ritonavir: first approval. **Drugs**, Auckland, v. 82, n. 5, p. 585–591, 2022.
- 87 DRUG DISCOVERY & DEVELOPMENT. The 50 best-selling pharmaceuticals of 2022: COVID-19 vaccines poised to take a step back. **Drug Discovery & Development**, 2023. Disponível em: : <https://www.drugdiscoverytrends.com/50-of-2022s-best-selling-pharmaceuticals/>. Acesso em: 04 jul. 2023.
- 88 URQUHART, L. Top companies and drugs by sales in 2022. **Nature Reviews Drug Discovery**, Berlin, v. 22, n. 4, p. 260–260, 2023.
- 89 BROWN, A. Top product forecasts for 2023. **Nature Reviews Drug Discovery**, Berlin, v. 22, n. 1, p. 8–8, 2023.
- 90 BLAIR, H. A. Sotorasib: first approval. **Drugs**, Auckland, v. 81, n. 13, p. 1573–1579, 2021.
- 91 DHILLON, S. Adagrasib: first approval. **Drugs**, Auckland, v. 83, n. 3, p. 275–285, 2023.
- 92 HUANG, L.; GUO, Z.; WANG, F.; FU, L. KRAS mutation: From undruggable to druggable in cancer. **Signal Transduction and Targeted Therapy**, London, v. 6, n. 1, p. 386, 2021.
- 93 BAUER, R. A. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. **Drug Discovery Today**, Oxon, v. 20, n. 9, p. 1061–1073, 2015.
- 94 VANE, J. R. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. **Nature New Biology**, Berlin v. 231, n. 25, p. 232–235, 1971.
- 95 HALFORD, B. Covalent drugs go from fringe field to fashionable endeavor. **Chemical & Engineering News**, 2020. Disponível em: <https://cen.acs.org/pharmaceuticals/drug-discovery/Covalent-drugs-fringe-field-fashionable/98/i43>. Acesso em: 04 jul. 2023.
- 96 DE VITA, E. 10 years into the resurgence of covalent drugs. **Future Medicinal Chemistry**, London, v. 13, n. 2, p. 193–210, 2020.
- 97 COPELAND, R. A.; POMPLIANO, D. L.; MEEK, T. D. Drug–target residence time and its implications for lead optimization. **Nature Reviews Drug Discovery**, Berlin, v. 5, n. 9, p. 730–739, 2006.
- 98 GEHRINGER, M.; LAUFER, S. A. Emerging and re-emerging warheads for targeted covalent inhibitors: Applications in medicinal chemistry and chemical biology. **Journal of Medicinal Chemistry**, Washington, v. 62, n. 12, p. 5673–5724, 2019.
- 99 ÁBRÁNYI-BALOGH, P.; PETRI, L.; IMRE, T.; SZIJJ, P.; SCARPINO, A.; HRAST, M.; MITROVIĆ, A.; FONOVIČ, U. P.; NÉMETH, K.; BARRETEAU, H.; ROPER, D. I.; HORVÁTI, K.; FERENCZY, G. G.; KOS, J.; ILAŠ, J.; GOBEC, S.; KESERŰ, G. M. A road map for prioritizing warheads for cysteine targeting covalent inhibitors. **European Journal of Medicinal Chemistry**, Issy-les-Moulineaux, v. 160, p. 94–107, 2018.

- 100 ÁBRÁNYI-BALOGH, P.; KESERŐ, G. M. Warheads for designing covalent inhibitors and chemical probes. *In*: YAO, X. (ed.) **Advances in Chemical Proteomics**. Amsterdam: Elsevier, 2021. p. 47–73.
- 101 DOS SANTOS, A. M.; OLIVEIRA, A. R. S.; DA COSTA, C. H. S.; KENNY, P. W.; MONTANARI, C. A.; VARELA, J. de J. G.; LAMEIRA, J. Assessment of reversibility for covalent cysteine protease inhibitors using quantum mechanics/molecular mechanics free energy surfaces. **Journal of Chemical Information and Modeling**, Washington, v. 62, n. 17, p. 4083–4094, 2022.
- 102 YU, H. S.; GAO, C.; LUPYAN, D.; WU, Y.; KIMURA, T.; WU, C.; JACOBSON, L.; HARDER, E.; ABEL, R.; WANG, L. Toward atomistic modeling of irreversible covalent inhibitor binding kinetics. **Journal of Chemical Information and Modeling**, Washington, v. 59, n. 9, p. 3955–3967, 2019.
- 103 POWERS, J. C.; ASGIAN, J. L.; EKICI, Ö. D.; JAMES, K. E. Irreversible inhibitors of serine, cysteine, and threonine proteases. **Chemical Reviews**, Washington, v. 102, p. 4639–4750, 2002.
- 104 FLANAGAN, M. E.; ABRAMITE, J. A.; ANDERSON, D. P.; AULABAUGH, A.; DAHAL, U. P.; GILBERT, A. M.; LI, C.; MONTGOMERY, J.; OPPENHEIMER, S. R.; RYDER, T.; SCHU, B. P.; UCCELLO, D. P.; WALKER, G. S.; WU, Y.; BROWN, M. F.; CHEN, J. M.; HAYWARD, M. M.; NOE, M. C.; OBACH, R. S.; PHILIPPE, L.; SHANMUGASUNDARAM, V.; SHAPIRO, M. J.; STARR, J.; STROH, J.; CHE, Y. Chemical and computational methods for the characterization of covalent reactive groups for the prospective design of irreversible inhibitors. **Journal of Medicinal Chemistry**, Washington, v. 57, p. 10072–10079, 2014.
- 105 STRELOW, J. M. A perspective on the kinetics of covalent and irreversible inhibition. **SLAS Discovery**, New York, v. 22, n. 1, p. 3–20, 2017.
- 106 SERAFIMOVA, I. M.; PUFALL, M. A.; KRISHNAN, S.; DUDA, K.; COHEN, M. S.; MAGLATHLIN, R. L.; MCFARLAND, J. M.; MILLER, R. M.; FRÖDIN, M.; TAUNTON, J. Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. **Nature Chemical Biology**, Berlin, v. 8, n. 5, p. 471–476, 2012.
- 107 WARD, R. A. Modeling covalent protein-ligand interactions. *In*: WOLKENHAUER, O. (ed.) **Systems medicine: integrative, qualitative and computational approaches**. Amsterdam: Academic Press, 2021. p. 174-189.
- 108 BRADSHAW, J. M.; MCFARLAND, J. M.; PAAVILAINEN, V. O.; BISCONTE, A.; TAM, D.; PHAN, V. T.; ROMANOV, S.; FINKLE, D.; SHU, J.; PATEL, V.; TON, T.; LI, X.; LOUGHHEAD, D. G.; NUNN, P. A.; KARR, D. E.; GERRITSEN, M. E.; FUNK, J. O.; OWENS, T. D.; VERNER, E.; BRAMELD, K. A.; HILL, R. J.; GOLDSTEIN, D. M.; TAUNTON, J. Prolonged and tunable residence time using reversible covalent kinase inhibitors. **Nature Chemical Biology**, Berlin, v. 11, n. 7, p. 525–531, 2015.
- 109 KRISHNAN, S.; MILLER, R. M.; TIAN, B.; MULLINS, R. D.; JACOBSON, M. P.; TAUNTON, J. Design of reversible, cysteine-targeted Michael acceptors guided by kinetic and computational analysis. **Journal of the American Chemical Society**, Washington, v. 136, n. 36, p. 12624–12630, 2014.

- 110 MILLER, R. M.; PAAVILAINEN, V. O.; KRISHNAN, S.; SERAFIMOVA, I. M.; TAUNTON, J. Electrophilic fragment-based design of reversible covalent kinase inhibitors. **Journal of the American Chemical Society**, Washington, v. 135, n. 14, p. 5298–5301, 2013.
- 111 YAMANE, D.; TETSUKAWA, R.; ZENMYO, N.; TABATA, K.; YOSHIDA, Y.; MATSUNAGA, N.; SHINDO, N.; OJIDA, A. Expanding the chemistry of dihaloacetamides as tunable electrophiles for reversible covalent targeting of cysteines. **Journal of Medicinal Chemistry**, Washington, 2023.
- 112 BONATTO, V.; SHAMIM, A.; ROCHO, F. dos R.; LEITÃO, A.; LUQUE, F. J.; LAMEIRA, J.; MONTANARI, C. A. Predicting the relative binding affinity for reversible covalent inhibitors by free energy perturbation calculations. **Journal of Chemical Information and Modeling**, Washington, v. 61, n. 9, p. 4733–4744, 2021.
- 113 OANCA, G.; ASADI, M.; SAHA, A.; RAMACHANDRAN, B.; WARSHEL, A. Exploring the catalytic reaction of cysteine proteases. **The Journal of Physical Chemistry B**, Washington, v. 124, n. 50, p. 11349–11356, 2020.
- 114 TUMMINO, P. J.; COPELAND, R. A. Residence time of receptor–ligand complexes and its effect on biological function. **Biochemistry**, Washington, v. 47, n. 20, p. 5481–5492, 2008.
- 115 MONDAL, D.; WARSHEL, A. Exploring the mechanism of covalent inhibition: simulating the binding free energy of α -ketoamide inhibitors of the main protease of SARS-CoV-2. **Biochemistry**, Washington, v. 59, n. 48, p. 4601–4608, 2020.
- 116 WOUTERS, O. J.; MCKEE, M.; LUYTEN, J. Estimated research and development investment needed to bring a new medicine to market, 2009-2018. **Journal of the American Medical Association**, Chicago, v. 323, n. 9, p. 844–853, 2020.
- 117 SCANNELL, J. W.; BLANCKLEY, A.; BOLDON, H.; WARRINGTON, B. Diagnosing the decline in pharmaceutical R&D efficiency. **Nature Reviews Drug Discovery**, Berlin, v. 11, n. 3, p. 191–200, 2012.
- 118 PAUL, S. M.; MYTELKA, D. S.; DUNWIDDIE, C. T.; PERSINGER, C. C.; MUNOS, B. H.; LINDBORG, S. R.; SCHACHT, A. L. How to improve R&D productivity: The pharmaceutical industry's grand challenge. **Nature Reviews Drug Discovery**, Berlin, v. 9, n. 3, p. 203–214, 2010.
- 119 DICKSON, M.; GAGNON, J. P. Key factors in the rising cost of new drug discovery and development. **Nature Reviews Drug Discovery**, Berlin, v. 3, n. 5, p. 417–429, 2004.
- 120 DIMASI, J. A.; GRABOWSKI, H. G.; HANSEN, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. **Journal of Health Economics**, Amsterdam, v. 47, p. 20–33, 2016.
- 121 BROWN, D. G. An analysis of successful hit-to-clinical candidate pairs. **Journal of Medicinal Chemistry**, Washington, v. 66, n. 11, p. 7101–7139, 2023.
- 122 KAUL, P. N. Drug discovery: Past, present and future. *In*: JUCKER, E. **Progress in Drug Research**. Basel: Birkhäuser Basel, 1998. p. 9–105.
- 123 MONTANARI, C. A. **Química medicinal: métodos e fundamentos em planejamento de fármacos**. São Paulo: EDUSP, 2011.

- 124 MOLINDER, H. K. M. The development of cimetidine: 1964-1976: a human story. **Journal of Clinical Gastroenterology**, Philadelphia, v. 19, n. 3, 1994.
- 125 THE NOBEL PRIZE IN PHYSIOLOGY OR MEDICINE 1988. Sir James W. Black Biographical. **The Nobel Prize**, 2010. Disponível em: <https://www.nobelprize.org/prizes/medicine/1988/black/biographical/>. Acesso em: 7 jul. 2023.
- 126 ZHENG, H.; HOU, J.; ZIMMERMAN, M. D.; WLODAWER, A.; MINOR, W. The future of crystallography in drug discovery. **Expert Opinion on Drug Discovery**, Oxon, v. 9, n. 2, p. 125–137, 2014.
- 127 TALELE, T. T.; RIGBY, S. A. K. and A. C. Successful applications of computer aided drug discovery: Moving drugs from concept to the clinic. **Current Topics in Medicinal Chemistry**, Sharjah, v 10 n. 1, p. 127–141, 2010.
- 128 PRIETO-MARTÍNEZ, F. D.; LÓPEZ-LÓPEZ, E.; EURÍDICE JUÁREZ-MERCADO, K.; MEDINA-FRANCO, J. L. Chapter 2 - Computational drug design methods—current and future perspectives. *In*: ROY, K. (ed.) **In silico drug design: repurposing techniques and methodologies**. Amsterdam: Academic Press, 2019. p. 19–44.
- 129 ROCHA, S. F. L. S.; OLANDA, C. G.; FOKOUE, H. H.; SANT'ANNA, C. M. R. Virtual screening techniques in drug discovery: review and recent applications. **Current Topics in Medicinal Chemistry**, Sharjah, v. 19 n. 19, p. 1751–1767, 2019.
- 130 DE VIVO, M.; MASETTI, M.; BOTTEGONI, G.; CAVALLI, A. Role of molecular dynamics and related methods in drug discovery. **Journal of Medicinal Chemistry**, Washington, v. 59, n. 9, p. 4035–4061, 2016.
- 131 COURNIA, Z.; ALLEN, B. K.; BEUMING, T.; PEARLMAN, D. A.; RADAK, B. K.; SHERMAN, W. Rigorous free energy simulations in virtual screening. **Journal of Chemical Information and Modeling**, Washington, v. 60, n. 9, p. 4153–4169, 2020.
- 132 COURNIA, Z.; CHIPOT, C.; ROUX, B.; YORK, D. M.; SHERMAN, W. Free energy methods in drug discovery. *In*: ARMACOST, K. A.; THOMPSON, D. C. (ed.) **Free energy methods in drug discovery: current state and future directions**. Washington: American Chemical Society, 2021. p. 1–38.
- 133 WALTERS, W. P. Virtual chemical libraries. **Journal of Medicinal Chemistry**, Washington, v. 62, n. 3, p. 1116–1124, 2019.
- 134 LIONTA, E.; SPYROU, G.; VASSILATIS, D.; COURNIA, z. Structure-based virtual screening for drug discovery: Principles, applications and recent advances. **Current Topics in Medicinal Chemistry**, Sharjah, v. 14, n. 16, p. 1923–1938, 2014.
- 135 DURRANT, J. D.; MCCAMMON, J. A. Molecular dynamics simulations and drug discovery. **BMC Biology**, London, v. 9, n. 1, p. 71, 2011.
- 136 FERREIRA, L.; DOS SANTOS, R.; OLIVA, G.; ANDRICOPULO, A. Molecular docking and structure-based drug design strategies. **Molecules**, Basel, v. 20, n. 7, p. 13384–13421, 2015.
- 137 HOUSTON, D. R.; WALKINSHAW, M. D. Consensus docking: improving the reliability of docking in a virtual screening context. **Journal of Chemical Information and Modeling**, Washington, v. 53, n. 2, p. 384–390, 2013.

138 LYU, J.; WANG, S.; BALIUS, T. E.; SINGH, I.; LEVIT, A.; MOROZ, Y. S.; O'MEARA, M. J.; CHE, T.; ALGAA, E.; TOLMACHOVA, K.; TOLMACHEV, A. A.; SHOICHET, B. K.; ROTH, B. L.; IRWIN, J. J. Ultra-large library docking for discovering new chemotypes. **Nature**, Berlin, v. 566, n. 7743, p. 224–229, 2019.

139 LUTTENS, A.; GULLBERG, H.; ABDURAKHMANOV, E.; VO, D. D.; AKABERI, D.; TALIBOV, V. O.; NEKHOTIAEVA, N.; VANGEEL, L.; DE JONGHE, S.; JOCHMANS, D.; KRAMBRICH, J.; TAS, A.; LUNDGREN, B.; GRAVENFORS, Y.; CRAIG, A. J.; ATILAW, Y.; SANDSTRÖM, A.; MOODIE, L. W. K.; LUNDKVIST, Å.; VAN HEMERT, M. J.; NEYTS, J.; LENNERSTRAND, J.; KIHLEBERG, J.; SANDBERG, K.; DANIELSON, U. H.; CARLSSON, J. Ultralarge virtual screening identifies SARS-CoV-2 main protease inhibitors with broad-spectrum activity against coronaviruses. **Journal of the American Chemical Society**, Washington, v. 144, n. 7, p. 2905–2920, 2022.

140 SINGH, I.; LI, F.; FINK, E. A.; CHAU, I.; LI, A.; RODRIGUEZ-HERNÁNDEZ, A.; GLENN, I.; ZAPATERO-BELINCHÓN, F. J.; RODRIGUEZ, M. L.; DEVKOTA, K.; DENG, Z.; WHITE, K.; WAN, X.; TOLMACHOVA, N. A.; MOROZ, Y. S.; KANISKAN, H. Ü.; OTT, M.; GARCÍA-SASTRE, A.; JIN, J.; FUJIMORI, D. G.; IRWIN, J. J.; VEDADI, M.; SHOICHET, B. K. Structure-based discovery of inhibitors of the SARS-CoV-2 nsp14 n7-methyltransferase. **Journal of Medicinal Chemistry**, Washington, v. 66, n. 12, p. 7785–7803, 2023.

141 FISCHER, A.; SMIEŠKO, M.; SELNER, M.; LILL, M. A. Decision making in structure-based drug discovery: visual inspection of docking results. **Journal of Medicinal Chemistry**, Washington, v. 64, n. 5, p. 2489–2500, 2021.

142 BREZNIK, M.; GE, Y.; BLUCK, J. P.; BRIEM, H.; HAHN, D. F.; CHRIST, C. D.; MORTIER, J.; MOBLEY, D. L.; MEIER, K. Prioritizing small sets of molecules for synthesis through *in-silico* tools: a comparison of common ranking methods. **ChemMedChem**, Weinheim, 2022. No prelo.

143 CORSO, G.; STÄRK, H.; JING, B.; BARZILAY, R.; JAAKKOLA, T. DiffDock: diffusion steps, twists, and turns for molecular docking. arXiv preprint arXiv.2210.01776, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2210.01776>. Acesso em: 06 jul. 2023.

144 GENTILE, F.; YAACOUB, J. C.; GLEAVE, J.; FERNANDEZ, M.; TON, A. T.; BAN, F.; STERN, A.; CHERKASOV, A. Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking. **Nature Protocols**, Berlin, v. 17, n. 3, p. 672–697, 2022.

145 MEY, A. S. J. S.; ALLEN, B.; MACDONALD, H. E. B.; CHODERA, J. D.; KUHN, M.; MICHEL, J.; MOBLEY, D. L.; NADEN, L. N.; PRASAD, S.; RIZZI, A.; SCHEEN, J.; SHIRTS, M. R.; TRESADERN, G.; XU, H. Best practices for alchemical free energy calculations. arXiv preprint arXiv.2008.03067, 2020. Disponível em: <https://doi.org/10.33011/livecoms.2.1.18378>. Acesso em: 06 jul. 2023.

146 SENN, H. M.; THIEL, W. QM/MM methods for biomolecular systems. **Angewandte Chemie International Edition**, Weinheim, v. 48, n. 7, p. 1198–1229, 2009.

147 SHENG, X.; KAZEMI, M.; PLANAS, F.; HIMO, F. Modeling enzymatic enantioselectivity using quantum chemical methodology. **ACS Catalysis**, Washington, v. 10, n. 11, p. 6430–6449, 2020.

- 148 SIEGBAHN, P. E. M.; HIMO, F. The quantum chemical cluster approach for modeling enzyme reactions. **Wiley Interdisciplinary Reviews: computational molecular science**, Hoboken, v. 1, n. 3, p. 323–336, 2011.
- 149 SHENG, X.; HIMO, F. The quantum chemical cluster approach in biocatalysis. **Accounts of Chemical Research**, Washington, v. 56, n. 8, p. 938–947, 2023.
- 150 SHENG, X.; HIMO, F. Mechanism of 3-methylglutaconyl CoA decarboxylase AibA/AibB: Pericyclic reaction versus direct decarboxylation. **Angewandte Chemie International Edition**, Weinheim, v. 59, n. 51, p. 22973–22977, 2020.
- 151 LOSEV, T. V.; GERASIMOV, I. S.; PANOVA, M. V.; LISOV, A. A.; ABDYUSHEVA, Y. R.; RUSINA, P. V.; ZALETSKAYA, E.; STROGANOV, O. V.; MEDVEDEV, M. G.; NOVIKOV, F. N. Quantum mechanical-cluster approach to solve the bioisosteric replacement problem in drug design. **Journal of Chemical Information and Modeling**, Washington, v. 63, n. 4, p. 1239–1248, 2023.
- 152 PREJANÒ, M.; ŠKERLOVÁ, J.; STENMARK, P.; HIMO, F. Reaction mechanism of human PAICS elucidated by quantum chemical calculations. **Journal of the American Chemical Society**, Washington, v. 144, n. 31, p. 14258–14268, 2022.
- 153 SCHNEIDER, G. Mind and machine in drug design. **Nature Machine Intelligence**, Berlin, v. 1, n. 3, p. 128–130, 2019.
- 154 YANG, X.; WANG, Y.; BYRNE, R.; SCHNEIDER, G.; YANG, S. Concepts of artificial intelligence for computer-assisted drug discovery. **Chemical Reviews**, Washington, v. 119, n. 18, p. 10520–10594, 2019.
- 155 REN, F.; DING, X.; ZHENG, M.; KORZINKIN, M.; CAI, X.; ZHU, W.; MANTSYZOV, A.; ALIPER, A.; ALADINSKIY, V.; CAO, Z.; KONG, S.; LONG, X.; MAN LIU, B. H.; LIU, Y.; NAUMOV, V.; SHNEYDERMAN, A.; OZEROV, I. V.; WANG, J.; PUN, F. W.; POLYKOVSKIY, D. A.; SUN, C.; LEVITT, M.; ASPURU-GUZIK, A.; ZHAVORONKOV, A. AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor. **Chemical Science**, Cambridge, v. 14, n. 6, p. 1443–1452, 2023.
- 156 IVANENKOV, Y. A.; POLYKOVSKIY, D.; BEZRUKOV, D.; ZAGRIBELNYY, B.; ALADINSKIY, V.; KAMYA, P.; ALIPER, A.; REN, F.; ZHAVORONKOV, A. Chemistry42: an ai-driven platform for molecular design and optimization. **Journal of Chemical Information and Modeling**, Washington, v. 63, n. 3, p. 695–701, 2023.
- 157 STROOBANTS, A.; MERVIN, L. H.; ENGVIST, O.; ROBB, G. R. An industrial evaluation of proteochemometric modelling: predicting drug-target affinities for kinases. **Artificial Intelligence in the Life Sciences**, Oxford, v. 4, p. 100079, 2023.
- 158 IVANENKOV, Y.; ZAGRIBELNYY, B.; MALYSHEV, A.; EVTEEV, S.; TERENTIEV, V.; KAMYA, P.; BEZRUKOV, D.; ALIPER, A.; REN, F.; ZHAVORONKOV, A. The hitchhiker's guide to deep learning driven generative chemistry. **ACS Medicinal Chemistry Letters**, Washington, v. 14, n. 7, p. 901–915, 2023.

- 159 STOKES, J. M.; YANG, K.; SWANSON, K.; JIN, W.; CUBILLOS-RUIZ, A.; DONGHIA, N. M.; MACNAIR, C. R.; FRENCH, S.; CARFRAE, L. A.; BLOOM-ACKERMANN, Z.; TRAN, V. M.; CHIAPPINO-PEPE, A.; BADRAN, A. H.; ANDREWS, I. W.; CHORY, E. J.; CHURCH, G. M.; BROWN, E. D.; JAAKKOLA, T. S.; BARZILAY, R.; COLLINS, J. J. A deep learning approach to antibiotic discovery. **Cell**, Cambridge, v. 180, n. 4, p. 688-702, 2020.
- 160 LIU, G.; CATACTAN, D. B.; RATHOD, K.; SWANSON, K.; JIN, W.; MOHAMMED, J. C.; CHIAPPINO-PEPE, A.; SYED, S. A.; FRAGIS, M.; RACHWALSKI, K.; MAGOLAN, J.; SURETTE, M. G.; COOMBES, B. K.; JAAKKOLA, T.; BARZILAY, R.; COLLINS, J. J.; STOKES, J. M. Deep learning-guided discovery of an antibiotic targeting acinetobacter baumannii. **Nature Chemical Biology**, Berlin, 2023. In press.
- 161 ZHAVORONKOV, A.; IVANENKOV, Y. A.; ALIPER, A.; VESELOV, M. S.; ALADINSKIY, V. A.; ALADINSKAYA, A. V.; TERENTIEV, V. A.; POLYKOVSKIY, D. A.; KUZNETSOV, M. D.; ASADULAEV, A.; VOLKOV, Y.; ZHOLUS, A.; SHAYAKHMETOV, R. R.; ZHEBRAK, A.; MINAEVA, L. I.; ZAGRIBELNYY, B. A.; LEE, L. H.; SOLL, R.; MADGE, D.; XING, L.; GUO, T.; ASPURU-GUZYK, A. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. **Nature Biotechnology**, Berlin, v. 37, n. 9, p. 1038–1040, 2019.
- 162 SCHIRMEISTER, T.; KESSELRING, J.; JUNG, S.; SCHNEIDER, T. H.; WEICKERT, A.; BECKER, J.; LEE, W.; BAMBERGER, D.; WICH, P. R.; DISTLER, U.; TENZER, S.; JOHÉ, P.; HELLMICH, U. A.; ENGELS, B. Quantum chemical-based protocol for the rational design of covalent inhibitors. **Journal of the American Chemical Society**, Washington, v. 138, n. 27, p. 8332–8335, 2016.
- 163 MÜLLER, P.; META, M.; MEIDNER, J. L.; SCHWICKERT, M.; MEYR, J.; SCHWICKERT, K.; KERSTEN, C.; ZIMMER, C.; HAMMERSCHMIDT, S. J.; FREY, A.; LAHU, A.; DE LA HOZ-RODRÍGUEZ, S.; AGOST-BELTRÁN, L.; RODRÍGUEZ, S.; DIEMER, K.; NEUMANN, W.; GONZÁLEZ, F. V.; ENGELS, B.; SCHIRMEISTER, T. Investigation of the compatibility between warheads and peptidomimetic sequences of protease inhibitors—a comprehensive reactivity and selectivity study. **International Journal of Molecular Sciences**, Basel, v. 24, n. 8, p. 7226, 2023.
- 164 DA COSTA, C.; BONATTO, V.; DOS SANTOS, A.; LAMEIRA, J.; LEITAO, A.; MONTANARI, C. Evaluating QM/MM free energy surfaces for ranking cysteine protease covalent inhibitors. **Journal of Chemical Information and Modeling**, Washington, v. 60, n. 2, p. 880–889, 2020.
- 165 ZHOU, Y.; YU, H.; VIND, A. C.; KONG, L.; LIU, Y.; SONG, X.; TU, Z.; YUN, C.; SMAILL, J. B.; ZHANG, Q.-W.; DING, K.; BEKKER-JENSEN, S.; LU, X. Rational design of covalent kinase inhibitors by an integrated computational workflow (Kin-Cov). **Journal of Medicinal Chemistry**, Washington, v. 66, n. 11, p. 7405–7420, 2023.
- 166 AWOONOR-WILLIAMS, E.; GOLOSOV, A. A.; HORNAK, V. Benchmarking *in silico* tools for cysteine pK_a prediction. **Journal of Chemical Information and Modeling**, Washington, v. 63, n. 7, p. 2170–2180, 2023.
- 167 ZHANG, H.; JIANG, W.; CHATTERJEE, P.; LUO, Y. Ranking reversible covalent drugs: from free energy perturbation to fragment docking. **Journal of Chemical Information and Modeling**, Washington, v. 59, n. 5, p. 2093–2102, 2019.

168 KUHN, B.; TICHÝ, M.; WANG, L.; ROBINSON, S.; MARTIN, R. E.; KUGLSTATTER, A.; BENZ, J.; GIROUD, M.; SCHIRMEISTER, T.; ABEL, R.; DIEDERICH, F.; HERT, J. Prospective evaluation of free energy calculations for the prioritization of cathepsin L inhibitors. **Journal of Medicinal Chemistry**, Washington, v. 60, n. 6, p. 2485–2497, 2017.

169 CHATTERJEE, P.; BOTELLO-SMITH, W. M.; ZHANG, H.; QIAN, L.; ALSAMARAH, A.; KENT, D.; LACROIX, J. J.; BAUDRY, M.; LUO, Y. Can relative binding free energy predict selectivity of reversible covalent inhibitors? **Journal of the American Chemical Society**, Washington, v. 139, n. 49, p. 17945–17952, 2017.

170 LAMEIRA, J.; BONATTO, V.; CIANNI, L.; DOS REIS ROCHO, F.; LEITÃO, A.; MONTANARI, C. A. Predicting the affinity of halogenated reversible covalent inhibitors through relative binding free energy. **Physical Chemistry Chemical Physics**, Cambridge, v. 21, n. 44, p. 24723–24730, 2019.

171 ZHOU, J.; SAHA, A.; HUANG, Z.; WARSHEL, A. Fast and effective prediction of the absolute binding free energies of covalent inhibitors of SARS-CoV-2 main protease and 20S proteasome. **Journal of the American Chemical Society**, Washington, v. 144, n. 17, p. 7568–7572, 2022.

2. Published Articles and Their Context in the Doctoral Project

This chapter summarizes the doctoral project objectives and their alignment with the published articles. This work aimed to apply computational chemistry and cheminformatics methods for the modeling of reversible covalent cysteine protease inhibitors. The specific objectives included determining parameters for QSAR models through quantum calculations, analyzing putative binding modes using molecular dynamics simulations, calculating the affinity of different molecules against CatL using alchemical free energy methods, and investigating the importance of non-covalent and covalent states in the binding free energy of reversible covalent inhibitors. These objectives guided the research conducted and resulted in the publication of three articles within the thesis as the first author.

I - Review Article - "Nitriles: an attractive approach to the development of covalent inhibitors". Royal Society of Chemistry – RSC Medicinal Chemistry, 14, 201-217, 2023.

This article provided a comprehensive review of the use of nitriles as a versatile and promising functional group in the development of covalent inhibitors. It discussed the advantages, scheme reaction, reactivity and selectivity of nitrile-based warheads, drugs with the nitrile in the reactive moiety and recent advances in the field. Overall, this article provides valuable insights and information that aligns with the goals of the thesis in understanding and exploring the potential of covalent inhibitors, including nitriles as a critical component in their design and development.

II - Results Article - "On the intrinsic reactivity of highly potent trypanocidal cruzain inhibitors". Royal Society of Chemistry – RSC Medicinal Chemistry, 11, 1275-1284, 2020.

This article presented the investigation findings into the intrinsic reactivity of potent trypanocidal cruzain inhibitors. The study utilized computational approaches, including the Fukui and electrophilicity indices, to assess the intrinsic reactivity of the compounds and correlate it with enzymatic affinity. The results highlight the significance of the predicted QM descriptors in establishing correlations with experimental potency measurements. These QM descriptors can be further utilized in QSAR models to facilitate the selection of new warheads for cysteine proteases.

III - Results Article - "Predicting the relative binding affinity for reversible covalent inhibitors by free energy perturbation calculations".
American Chemical Society – Journal of Chemical Information and Modeling, 61, 4733-4744, 2021.

This article focused on the development and application of free energy perturbation calculations to predict the relative binding affinity of reversible covalent inhibitors. Molecular modeling of covalent ligands presents challenges due to the need to consider both covalent and noncovalent ligand-bound states in the binding process as discussed before. The results demonstrate that the covalent complex can effectively rank the inhibitors and offer insights for developing new inhibitors in drug discovery projects considering only the covalent state.

In addition to these three articles, one more section is provided relating and discussing the use of the QM-cluster approach to access the difference in the energy when only the warhead is changed between pair compounds. This can be considered a complement of both results article since the protein environment is considered (this was not true in the case of the QM descriptors) and also opens the door to the possibility of obtaining the ABFE for covalent inhibitors in the combination of FEP + QM-cluster approach.

Overall, the published articles on nitriles group to the development of covalent inhibitors, intrinsic reactivity of trypanocidal cruzain inhibitors, and predicting the binding affinity of reversible covalent inhibitors all contribute directly to achieving the objectives of this doctoral project. Also, the relevance of these findings was possible to observe as all these three articles were selected to illustrate the cover of the respective issue of the journal. These findings enhance our understanding of covalent inhibitor design and provide valuable insights for future drug discovery projects.

3. Nitriles: An Attractive Approach to the Development of Covalent Inhibitors

Reproduced from BONATTO, V.; LAMEIRO, R. F.; ROCHO, F. R.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A. Nitriles: an attractive approach to the development of covalent inhibitors. **RSC Medicinal Chemistry**. Cambridge, v. 14, p. 201-217, 2023 with permission from the Royal Society of Chemistry.

Cite this: <https://pubs.rsc.org/en/content/articlelanding/2023/MD/D2MD00204C>



Cite this: *RSC Med. Chem.*, 2023, 14, 201

Nitriles: an attractive approach to the development of covalent inhibitors

Vinícius Bonatto, ^a Rafael F. Lameiro, ^a Fernanda R. Rocho, ^a Jerônimo Lameira, ^{ab} Andrei Leitão ^a and Carlos A. Montanari ^{id} ^{*a}

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

Received 1st July 2022,
Accepted 31st October 2022

DOI: 10.1039/d2md00204c

rsc.li/medchem

Introduction

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

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

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

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

In addition to the nonbonded interactions, nitriles are a remarkable group that can form covalent adducts with proteins, mainly linked to a reactive cysteine or serine side chain.^{6,7} The stability of the covalent bond between the ligand and the target can be modulated by different types of substituents in the vicinity of the nitrile, which also impacts its reactivity. The reactivity modulation is critical to the design of various types of inhibitors, from reversible to irreversible. In addition, the nitrile's electron-withdrawing property can be used in combination with other warheads, as is the case of cyano-acrylamides, in which the covalent bond is formed on the acrylamide β -carbon, while the cyano group increases the reactivity and promotes reversibility.^{8,9} All these

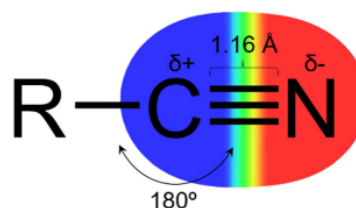


Fig. 1 Representation of the nitrile (cyano) group.

^a Medicinal and Biological Chemistry Group, São Carlos Institute of Chemistry, University of São Paulo, Avenue Trabalhador Sancarlenense, 400, 13566-590, São Carlos/SP, Brazil. E-mail: carlos.montanari@usp.br

^b Institute of Biological Science, Federal University of Pará, Rua Augusto Correa S/ N, Belém, PA, Brazil

Review

characteristics make the nitrile an interesting and versatile moiety to be considered in drug design.

Apart from affecting the PD profile of drug molecules, nitriles can be incorporated to improve pharmacokinetic properties. There are examples in which the addition of a nitrile does not influence the potency of the inhibitor, but makes the compound 10-fold more soluble than its molecular pair by decreasing $\log P$.¹⁰ In addition to this notable contribution to PK, the cyano functional group is considered a metabolically stable group, and may also reduce the susceptibility of hepatic oxidative metabolism.³ Nonetheless, cases relating to the toxicity of nitriles may happen due to a metabolically-induced release of cyanide into the body.¹¹ Despite being unusual, this reaction can occur in a few cases, such as after ingesting fruit pits containing cyanogenic glycosides, although in a minimal amount. Overall, nitriles are considered a non-toxic group and are usually eliminated unchanged from the human body.

This review focuses on compounds in which the nitrile group is key to the ligand's reactivity, and is inspired by the successful development and approval examples of saxagliptin and nirmatrelvir (US approval in 2009 and 2021, respectively). Non-reactive nitrile-containing drugs and candidates were previously described in a set of high-quality, in-depth reviews in the literature.^{2-4,12}

Herein, we will explain the function of nitriles in covalent inhibitors, followed by how to tune the reactivity of this group and discuss whether it is possible to achieve selectivity only by replacing the warhead. Ultimately, we review some successful cases of drugs and drug candidates that take advantage of the cyano group to covalently bind to macromolecules and recent applications of covalent nitrile-based inhibitors targeting cysteine proteases.

The role of nitriles in covalent inhibitors

Small molecule drugs can modulate a specific target, wherein the compound will interact with the receptor in an equilibrium process, represented by k_1 and k_{-1} (Fig. 2). The inhibition constant (K_i) is determined using the initial equilibrium state.¹³ A compound that acts reversibly (without the formation of a covalent bond) is called a noncovalent inhibitor, and the main interactions with the receptor are noncovalent (*e.g.*, hydrogen bonding, hydrophobic contact, van der Waals interactions), thus the K_i can be determined.



Fig. 2 General mechanism representing the ability of an inhibitor to form a noncovalent complex in equilibrium with an enzyme ($E \cdots I$), followed by the formation of a covalently bound state ($E-I$).

On the other hand, some inhibitors can inactivate their biological targets by forming an irreversible covalent bond with a residue in the target of interest.^{13,14} This process is represented by the constant k_2 or k_{inact} (Fig. 2), and this class of compounds is dubbed irreversible covalent inhibitors. However, the covalent bond formed between the ligand and the target of interest can be reversible, depending on the stability of the formed adduct complex, as well as other possible mechanisms for the reverse reaction, and external factors, such as hydrolysis. The constant k_{-2} (Fig. 2) will guide the reversibility of the covalent complex, where $k_2 \approx k_{-2}$ indicates a reversible system, but when $k_2 \gg k_{-2}$, the complex turns out to be irreversible.^{14,15}

In this scenario, the nitrile group can play a significant role in both cases. Regarding noncovalent inhibitors, nitriles can participate in several different interactions with the target of interest. Recently, Wang and colleagues reviewed the role of nitrile groups in protein–ligand interactions.⁴ They highlighted the hydrogen bond interactions that the nitrile group can perform with hydrogen bond donors from the protein backbone or side chains. In addition, the cyano group can interact with bridging water molecules and also participate in hydrophobic interactions.⁴

Nonetheless, the recent success of many covalent drugs like kinase inhibitors (ibrutinib, afatinib, and many others),¹⁶ nirmatrelvir to treat COVID-19,¹⁷ and sotorasib (the first KRAS-blocking drug)¹⁸ is catching the attention of the medicinal chemistry community, with an increasing number of reports describing new structures over time.¹⁹

Covalent drugs offer considerable advantages relative to noncovalent ones, the most notable potential benefits being the prolongation of therapeutic response, higher potency, improved selectivity, lower dosage and toxicity, and reduced probability of resistance mechanisms. Furthermore, the covalent inhibition approach prompted the targeting of enzymes that were considered intractable, such as KRAS G12C.¹⁹⁻²¹

The electrophilic characteristic of the nitrile's carbon atom makes it a functional reactive moiety (warhead) in the chemical scaffold of an inhibitor.^{6,7} In this way, the warhead can react with an active site nucleophilic residue of the target macromolecule to form a covalent bond. Cysteine and serine are the most common side-chain nucleophilic residues that react with nitrile to form a covalent adduct.⁶ Occasionally, the lysine side chain amino group may also react with the nitrile.^{6,22}

Nevertheless, non-catalytic residues can also be targeted by covalent inhibitors, although reacting with residues other than an anionic thiolate from an active-site cysteine (CysS⁻) can be more challenging.²³ Since the reactivity of the amino acid side chain is a function of its $\text{p}K_a$, Cys residues have the advantage of presenting lower values of $\text{p}K_a$ in comparison with Ser and Lys. Moreover, Cys residues exposed to the surface may have an even lower $\text{p}K_a$ due to interactions that make them more polarized, such as hydrogen bonds with water or other polar amino acids.²⁴ Nonetheless, buried non-

catalytic Cys have also been extensively explored to discover new covalent inhibitors of kinases when a nucleophilic thiol is accessible in an allosteric pocket.⁹

Meanwhile, Ser is almost exclusively only nucleophilic at the active site of the protein, under a charge relay mechanism in a catalytic dyad or triad (*e.g.*, protease catalytic mechanism) to activate the residue (SerO⁻).²³ Finally, the Lys side chain residue is mostly protonated when exposed to the solvent, displaying a $pK_a \sim 10.8$ for the ϵ -amine. However, a buried Lys can have a pK_a value of down to around 5.7 depending on the protein microenvironment, thus making it nucleophilic.²⁵

The two-step process to form the covalent adduct when the warhead is a nitrile reacting with a cysteine protease is depicted in Fig. 3.²⁶ Initially, the formation of the ionic pair between the catalytic dyad (Cys⁻/His⁺) takes place before the association of the ligand, a crucial step to the formation of the covalent adduct since the thiolate from Cys can act as a strong nucleophile.²⁷ However, it is still an open question in the literature regarding the stage when the acid-base reaction between the cysteine and histidine residues occurs: (i) it may happen in the dissociated state; or (ii) in the presence of the ligand in the active site.^{27,28} In any case, the ligand is recognized by the enzyme's binding site in the association step (Fig. 3A), where the noncovalent state is formed in analogy to the Michaelis complex (therefore, enabling K_i determination represented in Fig. 2).

Subsequently, during the acid-base reaction, and with the warhead group positioned in a proper orientation to form the covalent bond, the thiolate from Cys can perform a nucleophilic attack on the carbon of the nitrile (Fig. 3B). Furthermore, the proton from the protonated histidine will be transferred to neutralize and stabilize the nitrogen atom from the emergent thioimidate. This reaction mechanism

can occur in a stepwise or concerted form, although computational studies made by our group with a nitrile-based cruzain inhibitor indicate that the concerted mechanism is most likely to happen.²⁹ Therefore, the attack from the thiolate co-occurs with the proton transfer from His to the inhibitor.

Finally, the covalently bound adduct is formed (Fig. 3C), and the energy barrier of the reverse step will define the reversibility of the reaction through the rate constant k_{-2} .^{13,14} Since the nitrile-bearing compound acts as a reversible covalent inhibitor, the barrier of the reverse reaction is not likely to exceed the value of around 23.5 kcal mol⁻¹, corresponding to a residence time of 10 hours, typical for reversible covalent inhibitors.^{9,30} However, cases of reversible covalent inhibitors with a residence time over 150 hours are known, and factors other than the stability of the covalent bond may also influence the rate of the reverse reaction, such as the conformation of the covalent complex, the acidity of the α -proton in the adduct (in the case of cyanoacrylamides; *vide infra*), and the noncovalent interactions in the bound state.³¹

In addition to the reversible reaction, hydrolysis is another possible pathway to cleave the thioimidate product (Scheme 1). This would be similar to the deacylation step in the catalytic mechanism for cruzain or other papain-like cysteine proteases.³² However, studies have shown that the thioimidate is not usually hydrolyzed by papain, due to conformation changes that make the adduct inaccessible to the solvent.^{33,34} Interestingly, the addition of an external reactive thiol such as β -mercaptoethanol (β ME) has been shown to result in a 100% yield for the hydrolysis of the nitrile to carboxylic acid by papain.³⁵ Nevertheless, serine proteases can perform the hydrolysis of the nitrile adduct without the presence of an external thiol, as will be discussed later.³⁵

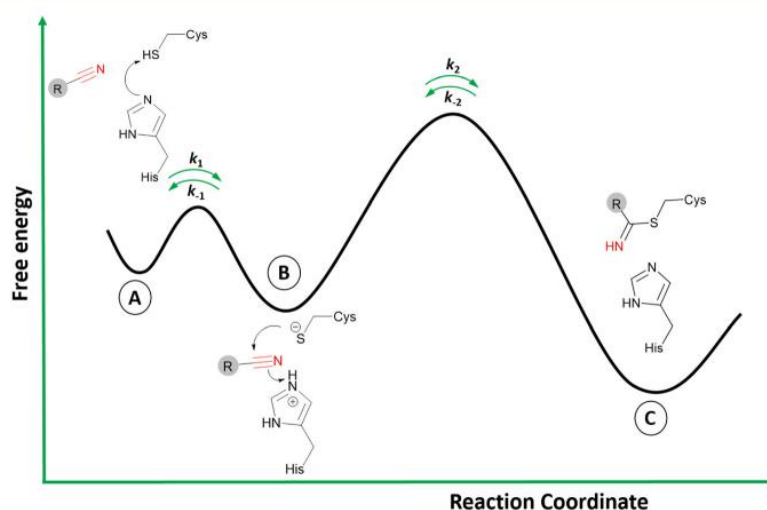
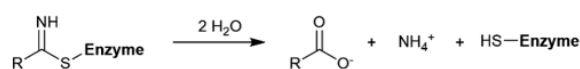


Fig. 3 Schematic reaction involving a reversible covalent inhibitor containing a nitrile warhead and a catalytic dyad Cys/His from a target enzyme. A) represents the dissociated state and the acid-base reaction; B) the noncovalent bound state; and C) the covalent state.

Review



Scheme 1 General hydrolysis reaction for the cleavage of the thioimide adduct.

The versatile use of nitriles also led to exciting novel reactive moieties when combined with acrylamides, obtaining α -cyanoacrylamide warheads. Even though the nitrile is not the group attacked by the nucleophilic residue, it is a compelling approach because an irreversible warhead is converted into a reversible one.³⁶

It is known that acrylamide-based inhibitors react irreversibly with their targets, and many approved kinase inhibitors have been designed to take advantage of this mechanism.^{8,9} However, the permanent inactivation of enzymes by covalent inhibition still raises concerns, mainly due to toxicity issues. Thus, replacing an irreversible warhead with a reversible one is, in most cases, a desirable approach.^{31,36}

In this context, Taunton and coworkers³⁶ pioneered the application of α -cyanoacrylamide warheads for kinase inhibition (Fig. 4). They observed that their reaction with β -mercaptoethanol (used to mimic the Cys residue) produced a reversible adduct. The reversibility of this reaction involved a β -elimination *via* an E1cB mechanism. Further studies³⁷ showed that the rate of the elimination had an inverse correlation with the calculated proton affinity of the corresponding carbanions, that is, increasing the acidity of the α -C-H in the adduct resulted in faster thiol elimination. Notably, the acidity of the proton at the α -position in the adduct can be modulated by diverse electron-withdrawing groups (EWGs) attached to the acrylonitrile.³⁷

Overall, the presence of the electron-withdrawing cyano group increases the reactivity of the acrylamide, and with it, the α -C-H acidity of the β -thioether adduct. Even though the cyanoacrylamide has a higher reactivity, it is also a reversible warhead, since the α -proton can be easily removed, which

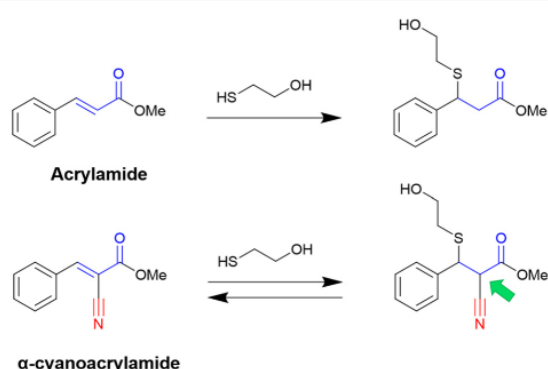


Fig. 4 Reactions using the acrylamide warhead result in an irreversible adduct, but a reversible product is formed for the α -cyanoacrylamide reagent. The α -C-H in the adduct form is indicated with a green arrow.

promotes the exit of the thiolate and thus the reverse reaction *via* an E1cB mechanism.^{36,37}

Reactivity and selectivity of nitrile-based warheads

Nitriles are weaker electrophiles than most employed warheads, such as aldehydes and azanitriles, which are almost four and ten times more reactive, respectively.³⁸ Despite that, nitriles are an excellent option for developing new covalent inhibitors, with the most recent successful case being the FDA-approved nirmatrelvir, as will be discussed further.

The reactivity of the nitrile group can be modulated by adding an adjacent electron-withdrawing group.³ Heteroaromatic nitriles, aminoacetonitriles, and cyanamides are examples of nitrile-derived warheads with improved reactivity.^{39–42} Among these, aminoacetonitriles are the weakest electrophilic groups because the electron-withdrawing amide is far from the reactive moiety. On the other hand, heteroaromatic nitriles (pyrimidine and triazine nitriles) and cyanamides are the most reactive warheads since they have electron-withdrawing atoms (*e.g.*, nitrogen) directly bonded to the CN group, resulting in a more electrophilic carbon atom.

Modulation of reactivity can also be explored in prodrugs, like the masking of the nitrile warhead in diacylfuroxan inhibitors. Under aqueous conditions, diacylfuroxans form a masked nitrile oxide, a very reactive organic functional group. These compounds are being investigated as GPX4 inhibitors, a therapeutic target for drug-resistant cancers.^{43,44}

The reactivity of nitriles can be investigated by computational methods. For example, the Fukui function and other DFT descriptors can be used for obtaining information about the reactivity site within a molecule.⁴⁵ QM/MM calculations can also be utilized to study the reaction mechanism of an inhibitor; hence, it is possible to extract information about the reactivity of the warhead in the transition state.¹⁴ Two interesting studies discussing the reactivity and how to tune the electrophilicity of nitrile-based inhibitors were done by Oballa *et al.*⁴⁶ and by Ehmke *et al.*⁴⁷ in which they proposed how to model them. Nevertheless, due to putative off-target effects, highly reactive inhibitors must be used carefully.⁴⁸

Experimentally, HPLC-based kinetic assays are commonly used to assess nitriles' intrinsic reactivity. These assays can be performed in the presence of a cysteine surrogate (*e.g.*, glutathione, β ME, and proteins) to quantify their half-life ($t_{1/2}$) for a nucleophilic attack from a soft thiol group.⁷ Correspondingly, the reactivity of nitrile-based warheads can be explored against *N*- α -acetyl-L-lysine (at pH 10.2) to mimic the attack from a deprotonated Lys residue.²⁵

In Fig. 5, the reactivity order for a wide range of different types of nitrile warheads with GSH is shown. Noteworthy, depending on the GSH assay, the values may vary sharply.

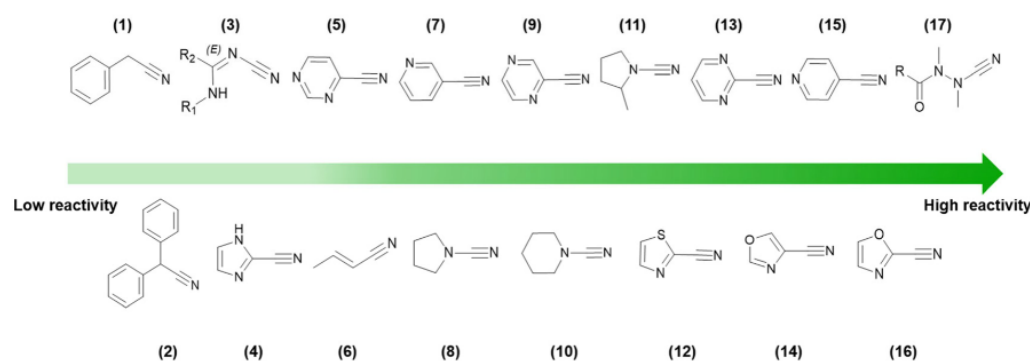


Fig. 5 HPLC-based kinetic assay using GSH to determine nitrile reactivity (decreasing from left to right) for different nitrile-based warheads. Values were retrieved from various academic publications based on comparable experimental assay conditions.^{38,49–51}

Still, the order of reactivity generally remains amidst studies.^{38,49–51}

The reactivity of different warheads (electrophiles) with thiol-based compounds (nucleophiles) can be rationalized in terms of their intrinsic electrophilicity, as well as the nucleophilicity and polarizability of reacting chemical species. Thus, the principle of hard and soft acids and bases (HSAB) defined by Pearson⁵² is of great value, since the anionic cysteine sulfhydryl is a highly polarizable group (soft base), due to the high energy of the $3sp^3$ orbitals and large ionic radius. This means that softer electrophiles should react more synergistically with cysteine.

As shown in Fig. 5, it is clear that being bonded to an electron-withdrawing heteroaromatic ring and the presence of electron-withdrawing groups in the vicinity increases the reactivity of the nitrile group. This is because these EWGs promote polarization *via* inductive and resonance effects that decrease the electron density at the carbon atom of the cyano moiety.

Outstandingly, azanitriles are among the most reactive warheads in the series, even being reported as irreversible inhibitors.⁵³ Several factors contribute to this, such as the inductive effect from the electronegative nitrogen, an extended electron density over the $N-C\equiv N$ moiety, and a resonance contribution from the lone pair of the nitrogen atom to the sp^2 carbon atom in the isothiosemicarbazide adduct, making it more stable.^{53,54}

Even though a warhead replacement can be an excellent strategy to improve an inhibitor's potency, the effect on selectivity might not be so straightforward,⁴⁷ since the new warhead will also affect potency over undesired targets if they present the same nucleophilic group. Nonetheless, this can be an interesting option when the target and its off-targets have different nucleophilic residues or a different chemical environment of the latter.^{14,55}

Keserü and coworkers⁴⁹ provided a protocol for prioritizing and designing warheads targeting the Cys residue. It was stated that the warhead chemistry could impact selectivity over undesirable targets. They noted that the selectivity of the electrophilic warhead group depended

on the targeted cysteine nucleophilicity. Therefore, they could efficiently propose selective warheads for target enzymes from a high-homology family (*e.g.*, human cysteine cathepsins) as long as these enzymes presented different cysteine nucleophilicity indexes. Thus, they concluded that modifications in the covalent fragment might be specific and conditional to the chemistry of the warhead, the cysteine reactivity, and the steric clashes that can happen between the ligand and the receptor.⁴⁹

Still, probably the best way to design selective covalent inhibitors over undesired targets involves exploring the noncovalent interactions in combination with a proper choice of the warhead.

Nitrile-containing pharmaceuticals in the reactive group

Vildagliptin and saxagliptin

Dipeptidyl peptidase 4 is a serine protease that inactivates incretin hormones and is a widely exploited target for treating type 2 diabetes mellitus (T2DM). Sandoz discovered one of the first hits for DPP-4: valine pyrrolidide (**18**), Fig. 6. However, it was observed that noncovalent compounds were less selective toward enzymes DPP-8 and DPP-9.⁵⁶ Therefore, further optimization of these noncovalent compounds involved the inclusion of warheads in their chemical scaffold. Phosphate diphenyl esters yielded less potent and irreversible inhibitors, and boronic acids were deemed too unstable. Nonetheless, it was noted that the nitrile warhead group could lead to inhibitors with nanomolar potency and

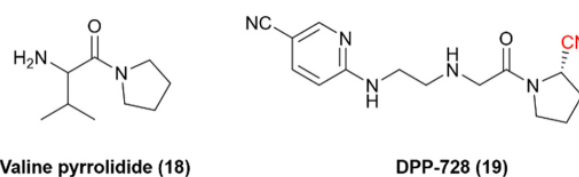


Fig. 6 Chemical structures of the pioneer DPP-4 inhibitor valine pyrrolidide and its optimized nitrile-containing analog DPP-728.

Review

adequate chemical stability to be administered orally. These studies eventually led to the development of the nitrile-based inhibitor DPP-728 (19), with improved glycemic control in patients with T2DM under clinical trials, providing the first proof-of-concept DPP-4 inhibitor.^{57,58}

It was determined that the nitrile warhead was responsible for forming a reversible covalent bond to the Ser630 residue. However, kinetic studies showed that DPP-728 was, in fact, a substrate for DPP-4 with a slow dissociation rate and not a true competitive inhibitor, with the nitrile being hydrolyzed to a carboxylic acid. Consequently, further optimization of this chemical series was necessary to improve the pharmacokinetic profile. These optimizations led to the development and approval of vildagliptin (20) and saxagliptin (21) for the treatment of T2DM (Fig. 7).^{57,59}

Vildagliptin and saxagliptin can form a reversible covalent bond with Ser630. Their binding modes were characterized by the crystal structure of these inhibitors bound to the enzyme in the covalent state (Fig. 7).^{60,61}

Even though the chemical structures and their binding modes are similar, QM/MM studies have suggested that only saxagliptin can be considered a true reversible covalent inhibitor, as it mainly dissociates intact from the active site. On the other hand, vildagliptin is more likely to be hydrolyzed by DPP-4 at the nitrile end, forming a carboxylic acid (also observed for DPP-728). These compounds have been termed “pseudo-irreversible inhibitors”, even though this terminology may not appropriately reflect the observed mechanism.⁶² Moreover, such findings are corroborated by the residence times of both drugs: vildagliptin has a relatively short residence time of 6.6 min, while saxagliptin shows an extended duration of 5.1 h.⁶³

The successful stories of vildagliptin and saxagliptin triggered many other research efforts in the field over the years, and many other DPP-4 nitrile-based inhibitors have

been approved worldwide (Fig. 8). Anagliptin (22) and trelagliptin (23) were approved in Japan in 2012 and 2015, respectively. Alogliptin (24) was approved in 2013 in the United States and European Union. Additionally, melogliptin (25) and bisegliptin (26) presented successful phase II clinical trials.

Nirmatrelvir

Researchers at Pfizer developed inhibitors for the SARS-CoV-1 main protease ($M^{Pro}/3CL^{Pro}$) during the SARS outbreak in 2003. Among them, PF-00835231 (27) was identified as a promising lead with an α -hydroxymethyl ketone as a warhead. The compound did not go through clinical trials due to the containment of the SARS-CoV-1 episode. However, with the start of the SARS-CoV-2 pandemic at the end of 2019, Pfizer's researchers looked back to SARS-CoV-1 inhibitors to transform an old drug candidate into a new one.

They decided to use PF-00835231 (Fig. 9) as a starting point for developing novel drug candidates to treat SARS-CoV-2 since the main proteases of SARS-CoV-1 and SARS-CoV-2 share a very similar structure and sequence.⁶⁴ As expected, PF-00835231 was also a potent inhibitor of the SARS-CoV-2 main protease (M^{Pro}), with a K_i of 0.27 ± 0.1 nM.⁶⁵

The major obstacle to the further development of PF-00835231 came from low gastrointestinal absorption. Therefore, the oral administration had to be replaced by the intravenous (IV) one. Since the goal of the pharmaceutical company was to develop a treatment to prevent the need for hospitalization, the chemical structure had to be optimized for oral bioavailability. One of the first strategies was the replacement of the α -hydroxymethyl ketone warhead with a chemical group that did not have hydrogen bond donors.

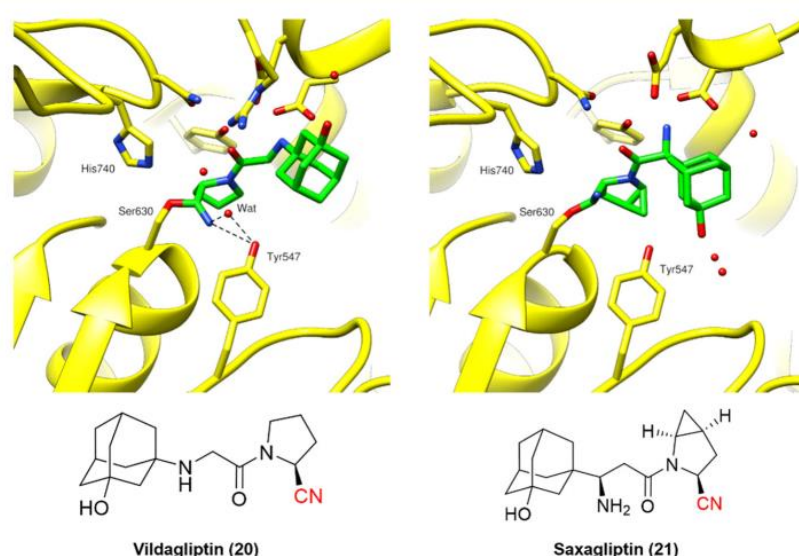


Fig. 7 Binding modes of DPP-4 inhibitors vildagliptin (PDB code: 6B1E)⁶⁰ and saxagliptin (PDB code: 3BJM).⁶¹

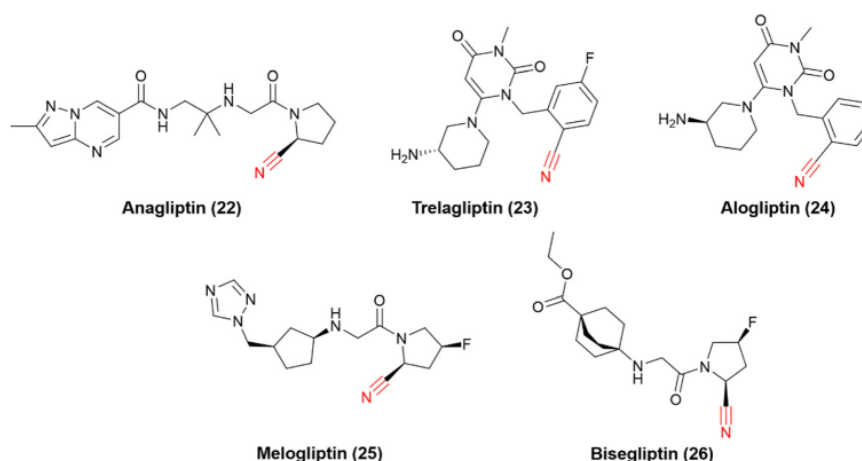


Fig. 8 Chemical representation of drugs and candidates that inhibit DPP-4 to treat T2DM.

Two series of compounds were produced, containing the nitrile (28) or the benzothiazol-2-yl ketone (29) warhead, with examples shown in Fig. 10.^{66,67}

Both series led to promising structures, but the team decided to proceed with the nitrile derivatives for three main reasons: (i) the nitriles were more soluble, (ii) less prone to epimerization, and (iii) the synthesis was easier to scale up. Besides, another study showed that the new lead compound was selective over human cathepsins (CatB, CatL, CatS),⁶⁸ reducing the probability of off-target effects.

Additional optimization efforts were made with the replacement of leucine by the cyclic P2 group of boceprevir to remove another hydrogen bond donor, also considered a positive point toward an oral drug. Finally, the chemical scaffold was modified to improve the potency and permeability by changing the P3 moiety and adding a P4 group.^{66,67}

The outcome of these optimization efforts was the inhibitor PF-07321332, or nirmatrelvir (30), represented in Fig. 11. Nirmatrelvir, in combination with the HIV antiviral ritonavir (used to decrease the CYP metabolism), was given

emergency use authorization in December 2021 by the FDA (followed by other regulatory agencies worldwide) for the treatment of COVID-19 under the trade name Paxlovid.

Following the publication of the chemical structure of nirmatrelvir, Vankadara and coworkers performed a warhead substitution study on this scaffold.⁶⁹ They found that four warheads (aldehyde, hydroxymethylketone, ketoamide and ketobenzothiazole) led to equally or more potent compounds towards M^{pro} . Two of these compounds (27 and 29) had already been explored by Pfizer, and 27 has been evaluated in clinical trials through IV administration. In line with the success of nirmatrelvir, novel nitrile-based M^{pro} inhibitors were developed.^{68,70} Interestingly, Breidenbach and coworkers employed the azanitrile warhead (31) and extended the chemical scaffold to interact with the S5 subsite, thereby improving the potency against M^{pro} ($K_i = 24.0$ nM), Fig. 12. In

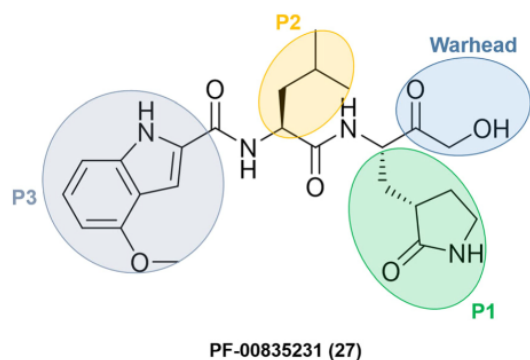


Fig. 9 Structure of SARS-CoV-1 and SARS-CoV-2 main protease inhibitor PF-00835231 emphasizing the P1–P3 substituents and the warhead moiety.

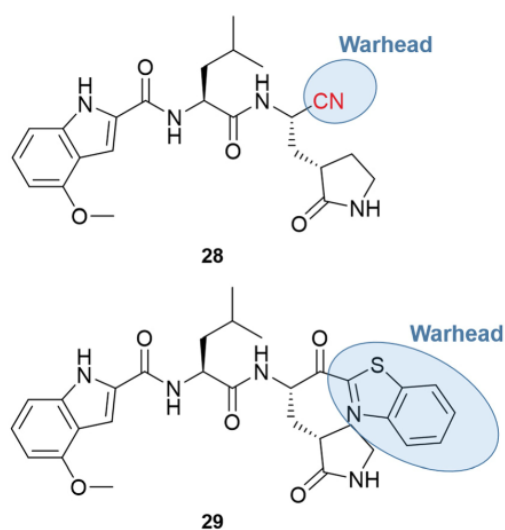


Fig. 10 Nitrile and benzothiazol-2-yl ketone derivatives of PF-00835231.

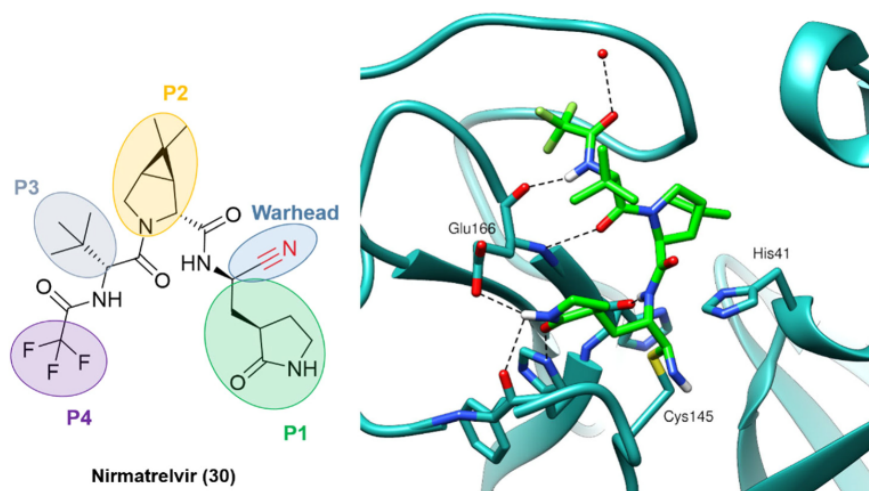


Fig. 11 Nirmatrelvir structure and the binding mode of its covalent form against SARS-CoV-2 M^{Pro}. PDB code: 7RFW.⁶⁶

addition, they noted that the compound bound irreversibly with the target with a k_{inact}/K_i of $37\,500\text{ m}^{-1}\text{ s}^{-1}$.⁵³

Rilzabrutinib and compounds targeting kinases

Bruton's tyrosine kinase (BTK), an enzyme involved in intracellular signaling and immune pathways, is a target of interest for cancer and autoimmune diseases.⁷¹ Targeting a cysteine residue (Cys481) proximal to the adenosine triphosphate (ATP) binding pocket has been a commonly explored mechanism for BTK inhibition. Notably, this residue is present in only ten kinases, thereby achieving excellent selectivity profiles.^{72,73} Using this strategy, several irreversible inhibitors of BTK have been discovered, such as the approved drugs ibrutinib and acalabrutinib.^{9,16}

Irreversible covalent inhibitors still raise safety concerns despite numerous successful clinical trials, especially when searching for a novel and selective autoimmune disease treatment. Therefore, reversible covalent inhibitors of BTK have been the subject of many research projects.⁷³

Acrylamide-based inhibitors were explored, based on the previous knowledge that acrylate/acrylamide-based kinase inhibitors would show an improved reactivity profile with the

addition of a cyano group to the electrophilic β -carbon, while avoiding the formation of irreversible adducts.^{36,74} However, previous studies with other kinases connected the β -carbon directly to a kinase-recognition scaffold.⁷⁴ Since the targeted cysteine group in BTK is distant from the ATP binding site, the authors postulated that an "inverted" cyanoacrylamide warhead, with the electrophilic β -carbon pointing away from the site, would have a more appropriate orientation for the formation of the covalent bond.³¹ It is worth mentioning that this "inversion" was relative to previously published cyanoacrylamide-based inhibitors, but it is in fact the "normal" orientation in classical acrylamide-based kinase inhibitors. Both warhead orientations are depicted in Fig. 13 to illustrate this design approach.

The "inverted" warhead positioning, combined with optimization *via* addition of a branched-alkyl capping group to the electrophilic carbon provided excellent results, such as

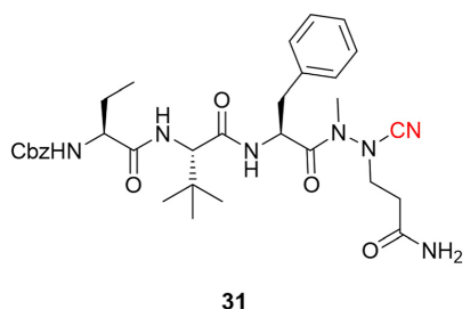


Fig. 12 Azanitrile-based inhibitor of the SARS-CoV-2 M^{Pro} target.

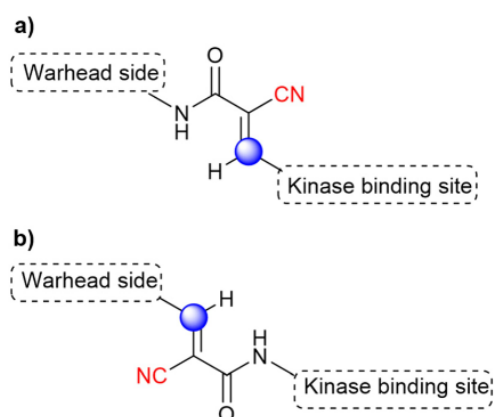
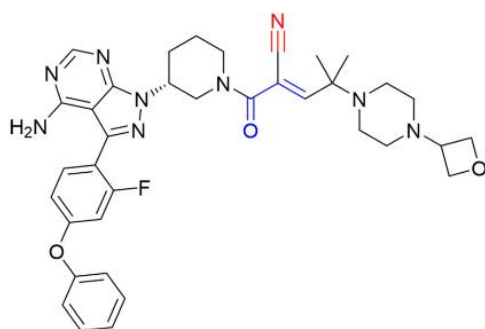


Fig. 13 The design orientation of the cyanoacrylamide warhead is based on (a) the traditional positioning of the investigated kinase inhibitors and (b) the "inverted" approach. The electrophilic carbon is marked in blue, while the warhead is shown in red.



Rilzabrutinib (32)

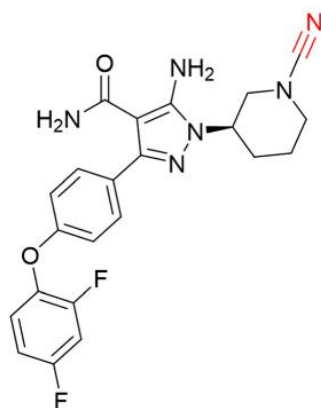
Fig. 14 Structure of the BTK inhibitor rilzabrutinib.

a residence time of 167 ± 21 h for one of the compounds in a kinetic-competition assay against BTK.³¹ The improvement in residence time was also observed in *in vivo* assays. Hence, further optimization efforts led to the development of PRN1008, or rilzabrutinib (32), a very potent covalent reversible inhibitor of BTK ($IC_{50} = 1.3 \pm 0.5$ nM) with long BTK occupancy and increased efficacy in rat models. Still, the compound is highly selective against more than 200 kinases (Fig. 14).⁷³

Despite failing on a phase 3 clinical trial for pemphigus,⁷⁵ rilzabrutinib was granted FDA Fast Track Designation in 2020 for immune thrombocytopenia. In addition, a phase 2 study is being conducted for the autoimmune condition known as IgG4-related disease.⁷⁶

Cyanamides have also been explored in this context of targeting noncatalytic cysteines in kinases. This electrophile has the cyano group bonded to a nitrogen atom and is more reactive than its carbon-bonded counterparts.

Schnute and colleagues⁷⁷ synthesized a series of compounds containing cyanamide as a warhead to replace the acrylamide-based irreversible BTK inhibitors. The authors managed to obtain reversible inhibitors with improved



33

Fig. 15 Potent and selective inhibitor of BTK by targeting noncatalytic Cys481.

selectivity against a panel of 51 kinases and excellent rat pharmacokinetic profiles by targeting Cys481. Compound 33 exhibited excellent results for inhibiting wild-type BTK ($IC_{50} = 1.5$ nM) with high oral bioavailability in rats, Fig. 15.

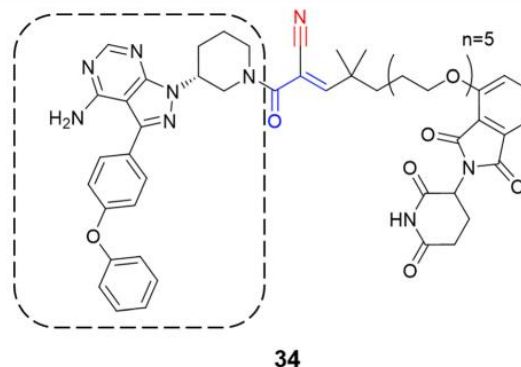
Janus kinase 3 (JAK3) is another kinase that has a noncatalytic cysteine (Cys909) to be targeted. It has been the subject of studies involving the development of covalent reversible inhibitors based on both cyanoacrylamide⁷⁸ and cyanamide.⁷⁹ Both warheads showed promising results for JAK3 inhibition, and selectivity over the kinome, also having good PK profiles.

Recently, one more interesting option for targeting kinases has been the use of proteolysis targeting chimeras (PROTACs).⁸⁰ In a prominent study, Gabizon *et al.* employed the cyanoacrylamide as the electrophilic moiety (34) targeting BTK, wherein it achieved high degradation potency ($DC_{50} = 6$ nM) and maximal degradation of the protein ($D_{max} = 85\%$).⁸¹ The chemical structure of the compound was based on ibrutinib with the replacement of the acrylamide warhead by cyanoacrylamide, thereby creating a reversible covalent PROTAC (Fig. 16). A reversible warhead is preferable, as irreversible PROTACs could not take advantage of the catalytic mechanism of reversible PROTACs. However, mixed results were achieved for PROTACs that exploit the covalent mechanism of action.^{81–83}

The recent development of nitrile-based inhibitors against cysteine proteases

Parasitic cysteine proteases

Cysteine protease inhibitors containing a nitrile warhead have been investigated for many years, especially for human cathepsins and related enzymes from viruses and parasites,^{41,84–88} yet no human or parasitic cysteine protease inhibitor has ever been approved to treat any disease. Nevertheless, this area of research is still prolific and will likely gain popularity after the approval of nirmatrelvir. Therefore, we will review recent developments for parasitic



34

Fig. 16 Structure of the PROTAC containing the cyanoacrylamide warhead with the ibrutinib scaffold highlighted by the dashed line.

Review

targets involved in several neglected diseases and advances in targeting human cathepsins.

Malaria is an infectious disease caused by protozoa of the genus *Plasmodium*. Despite the existing drugs available for its treatment, more than 600 000 people die each year from the disease.⁸⁹ The parasite variability among the species, infection stage, drug resistance and many other factors hinder the identification of efficacious drugs.⁹⁰

Nitrile-containing inhibitors of the parasitic cysteine protease falcipain-2 have been developed and evaluated as potential new treatments for malaria, such as pyrimidine nitrile-pyrazolines (35, IC_{50} = 1.63 μ M),⁹¹ and peptide-like compounds (36, IC_{50} = 2.7 nM), Fig. 17.⁹² Nevertheless, the diversity and complexity of *Plasmodium* parasites make drug resistance a significant problem, with single target inhibitors being especially prone to the emergence of resistant strains. To overcome this issue, researchers modified endoperoxide-containing compounds by appending to them nitrile moieties from known falcipain-2 inhibitors, thereby creating hybrid structures with two mechanisms of action: oxidative stress inducer coupled to falcipain-2 inhibition. The resulting compound (37, Fig. 17) showed excellent enzyme inhibition values (IC_{50} = 3.4 nM) while also presenting activity against three strains of *P. falciparum* resistant to commonly used drugs.⁹³

Another parasitic illness, Chagas disease, caused by the protozoa *Trypanosoma cruzi*, has only two nitro-heterocyclic approved drugs (nifurtimox and benznidazole). However, the use of these drugs is related to toxic effects, bringing the need to find potential new drugs to treat this disease.⁹⁴ Accordingly, inhibitors of the main protease of *T. cruzi*, cruzain, have been explored.

Dipeptidyl nitrile derivatives (38, K_i = 0.5 μ M) were investigated as cruzain inhibitors by our research group, Fig. 18.⁹⁵ The results showed that increasing the

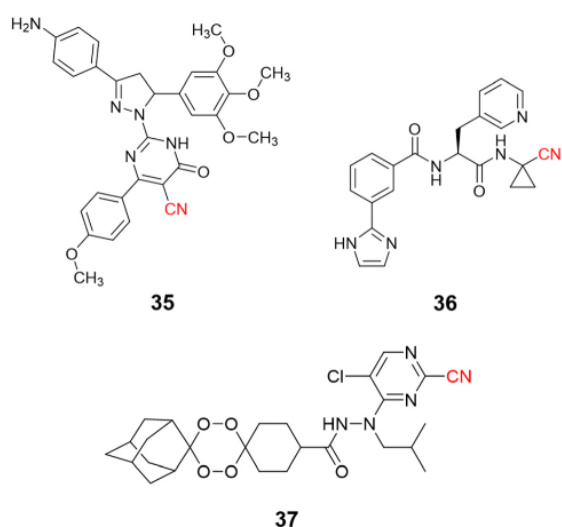


Fig. 17 Examples of nitrile-based inhibitors of cysteine protease falcipain-2.

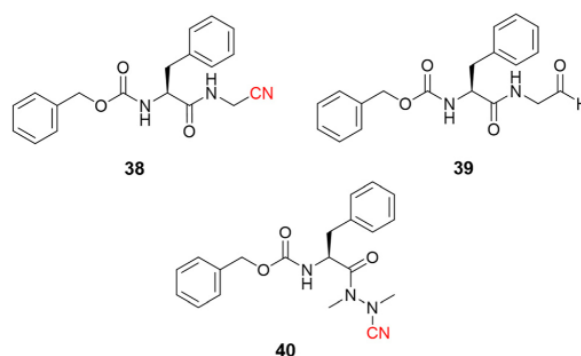


Fig. 18 The potency of *T. cruzi* cysteine protease cruzain inhibitors depicted here increases from nitrile to aldehyde and to azanitrile, following the reactivity profile of the warhead.

electrophilicity of the warhead by replacing the nitrile with an aldehyde (39, K_i = 0.005 μ M, Fig. 18) or oxime (K_i = 0.1 μ M) increased the potency against the enzyme, however, at the cost of increasing the toxicity according to the follow-up cell assays. Of particular interest, compounds bearing the azanitrile (40, Fig. 18)⁵⁴ were the most potent against cruzain, with two orders of magnitude higher than the corresponding nitrile counterparts. Presumably, the higher potency of the azanitrile warhead comes from its intrinsic reactivity (Fig. 5).⁹⁵

Further studies with azanitrile derivatives reinforced the higher potency against cruzain than other warheads.³⁸ Computer simulations and experimental assays were coupled to analyze the compounds' half-life after the incubation with GSH or cysteine. These results linked the intrinsic reactivity of the warhead to the inhibition against cruzain.³⁸

The use of nitrile as a warhead for reversible covalent inhibitors of cysteine proteases is not limited to dipeptidyl scaffolds. A series of bioisosteric replacements of the P2–P3 amide bond by a trifluoromethylamine moiety (41) has been explored, yielding compounds with excellent potency (K_i = 1.58 nM) and selectivity over human cathepsins. Moreover, it is less likely to be hydrolyzed, improving metabolic stability, a known weak spot of peptide-like structures.^{96,97} Another interesting approach involved the preparation of peptoids (42) based on the peptide-like scaffold. Using peptoid-based compounds led to an inhibitor with a K_i of 0.16 μ M against

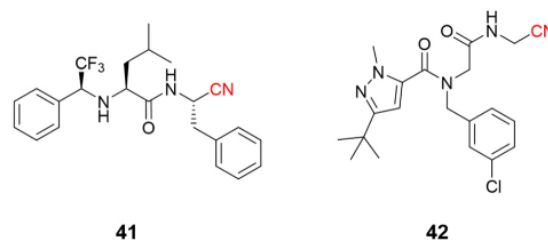


Fig. 19 Further expansion of the chemical diversity of cruzain inhibitors was made for nitrile-based dipeptidyl (41) and peptoid (42) derivatives.

cruzain. Even though this is a reasonably good affinity for a hit compound, the inhibition was lower than for the corresponding peptide. This difference can be rationalized according to the displacement of the P2 position in the S2 recognition site of the protein, as suggested by docking studies (Fig. 19).⁹⁸

Azanitrile-based inhibitors were also explored for *Leishmania mexicana* cysteine protease B (*LmCPB*), another parasitic cysteine protease target found in the causative agent of leishmaniasis. A wide potency range was achieved with nitrile-based inhibitors (K_i ranging from 50 μM to 5 nM) coming from the modulation of noncovalent interactions with the amino acid residues in the active site.⁹⁹ Azanitrile-based compounds were more potent than their nitrile counterparts. Compound (43) is the most potent *LmCPB* inhibitor reported to date ($K_i = 0.2$ nM) and it was used to obtain the first X-ray crystal structure of *LmCPB* (Fig. 20).⁹⁹ Although azanitriles present promising results in terms of potency for several targets containing catalytic cysteine residues, more studies will be required to verify if these compounds can be appropriately modulated for selectivity and toxicity, given the high reactivity of the warhead and its irreversible mechanism of inhibition.

Another disease tackled by targeting a cysteine protease is African human trypanosomiasis (HAT, sleeping sickness), an endemic disease in sub-Saharan Africa caused by *Trypanosoma brucei*.^{100,101} Currently, four drugs are approved to treat HAT (suramin, pentamidine, melarsoprol, and eflornithine). However, these drugs have severe toxicity issues that limit their uses, as for Chagas disease treatments. Hence, developing new therapies against HAT is crucial, with the cysteine protease rhodesain being a widely explored target for developing new inhibitors since it is vital for the parasite's life cycle.¹⁰²

Recently, di Chio and coworkers developed a series of dipeptidyl nitriles as rhodesain inhibitors,¹⁰³ once again

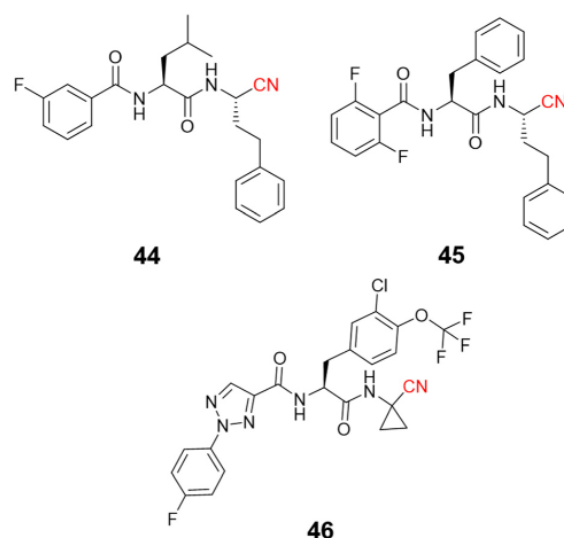


Fig. 21 Promising compounds to treat HAT by the inhibition of rhodesain.

highlighting the value of peptidomimetic nitriles targeting cysteine proteases. The structure-based drug design was used to design compounds that achieved nanomolar inhibition potency against rhodesain and micromolar potency in *in vitro* assays against *T. brucei brucei*. Fig. 21 depicts the most potent rhodesain inhibitor (44, $K_i = 14.1$ μM) and the best antiparasitic compound (45, $\text{EC}_{50} = 8.8$ μM) developed in this work.

Another impressive potential candidate for treating HAT is compound 46, with a $K_i = 7.4$ nM against rhodesain and $\text{IC}_{50} = 18.8$ nM against *T. brucei rhodesiense*, also displaying a good PK profile and selectivity over other human cysteine proteases. In addition, derivatives of this compound exhibited promising *in vivo* results when administered

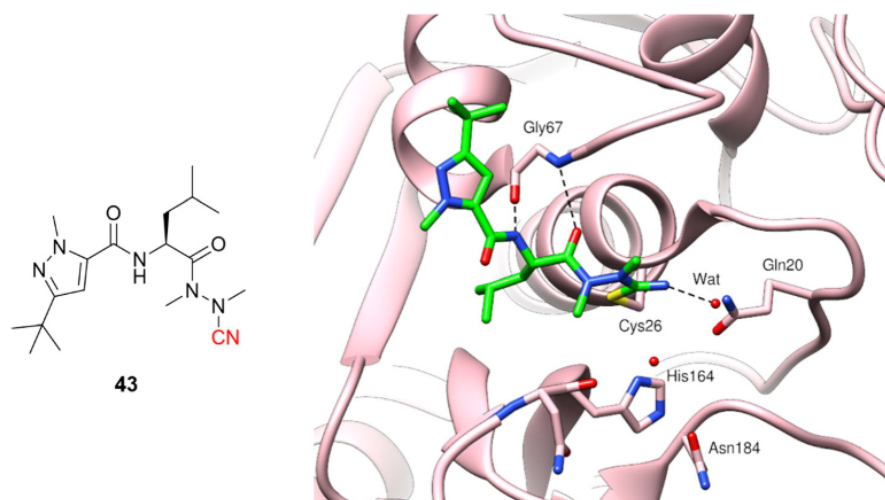
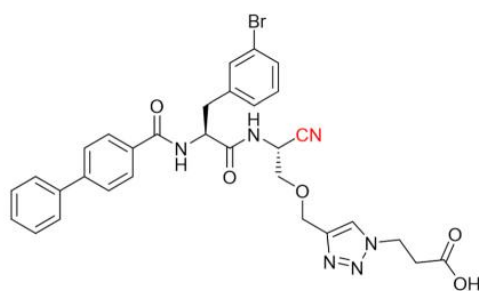


Fig. 20 Azanitrile-based inhibitor of *LmCPB* and its cocrystallized structure (PDB code: 6P4E).⁹⁹



47

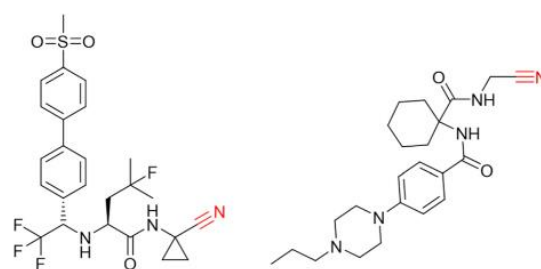
Fig. 22 The nitrile-based inhibitor of CatB targeting the occluding loop region of the enzyme.

orally, reducing almost 50% parasitemia compared to untreated mice.¹⁰⁴

Human cysteine proteases

Eleven human cysteine proteases constitute the family of enzymes called cathepsins (B, C, F, H, K, L, O, S, V, W, and X). These enzymes are essential regulators of physiological processes.^{105,106} However, many pathological conditions observed in humans are related to the dysfunction of these enzymes, making them attractive targets for developing new drugs.^{106,107} There is still no approved drug targeting human cathepsins, due to side effects observed in clinical trials.

The most studied cathepsins are CatB, CatK, and CatL, since these three enzymes have an increased expression in cancer cells, indicating that they may be involved with neoplastic progression.^{108,109} CatB is also related to neurodegenerative diseases such as Parkinson's and Alzheimer's, where the cells responsible for the defense of



Odanacatib (53)

Balicatib (54)

Fig. 24 Structures of CatK inhibitors odanacatib and balicatib that evolved to clinical trials.

the central nervous system secrete this enzyme due to upregulation, causing apoptosis.^{110,111} CatK is widely known to be an interesting target in bone-related diseases, as it is highly expressed in osteoclasts. When secreted, it is responsible for degrading the bone matrix.^{112,113} CatL has become an interesting target to impair the viral replication of SARS-CoV-2 since it is a crucial enzyme related to the virus entry and replication in the host cells.^{114,115}

Schmitz and colleagues employed click chemistry to synthesize a series of CatB inhibitors containing the nitrile group as the warhead, achieving nanomolar potency.¹¹⁶ Compound 47 (Fig. 22) and its analogs were designed to interact with the occluding loop of the CatB in the S1'. This improved the selectivity more than 10-fold over CatK, CatL, and CatS.

Hardegger and coworkers described for the first time the importance of the halogen bond in the context of CatL inhibition to achieve selectivity over other cathepsins and increase potency.¹¹⁷ The series of nitrile-containing compounds presented outstanding CatL inhibition values,

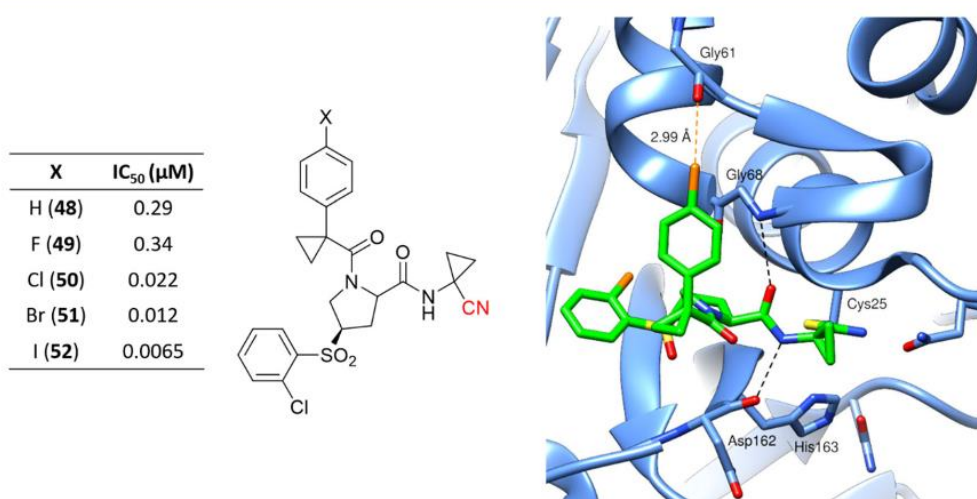


Fig. 23 Series of CatL inhibitors developed emphasizes the importance of the halogen bond to enzyme recognition. The cocrystallized structure of 50 with CatL (PDB code: 2XU1)¹¹⁷ is also represented. Cl and the dashed lines representing the halogen bond between the Cl atom and the main chain of the Gly61 oxygen atom are shown in orange.

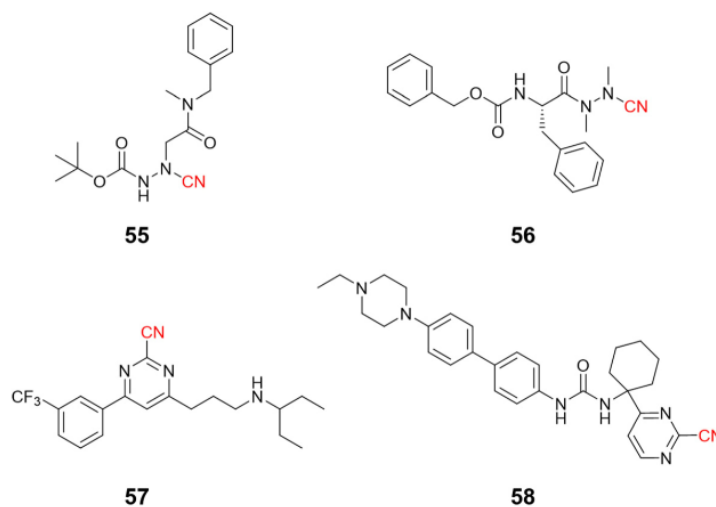


Fig. 25 Highly potent nitrile-containing CatK inhibitors described in the literature.

and the potency increased with the use of heavier halogen atoms (Fig. 23). The sigma-hole, which is a region of positive electrostatic potential in the halogen's surface due to the anisotropic distribution of the charges, is responsible for the halogen bond interaction with Gly61. The electrostatic potential in the sigma-hole will be more positive in the order $I > Br > Cl$, with Fig. 23 illustrating the halogen bond for compound 50 in the enzyme active site. Further studies have been done exploiting the halogen bond to model new CatL inhibitors.^{26,118}

Recently, an article was published showing strategies to achieve selectivity for CatB over CatL (or *vice versa*) for nitrile-based inhibitors.⁴¹ In addition, a review discussing patents of CatB and CatL inhibitors¹¹⁹ highlighted that some of the most promising compounds presented nitrile as the reactive group.

Ultimately, the most notable inhibitors of CatK are the nitrile-containing compounds odanacatib (53) and balicatib (54), Fig. 24. Both reached clinical trials (phases III and II, respectively); however, the studies were discontinued due to undesirable side effects. Nevertheless, these compounds were crucial for developing the next generation of nitrile-based inhibitors for CatK. Like CatB and CatL inhibitors, the most attractive reversible covalent inhibitors patented for CatK present the nitrile group as their reactive center.¹²⁰

Compounds targeting CatK may feature a variety of groups in their chemical scaffold.¹²¹ For instance, Benýšek and coworkers developed azanitriles to reach picomolar potency against CatK (55, $K_i = 13$ pM), Fig. 25.¹²² In the same work, the authors developed a structurally distinct azanitrile-based compound (56) to inhibit CatK with a K_i of 0.91 nM. Rankovic and colleagues employed a 2-cyanopyrimidine-based compound to obtain highly selective and potent compounds for CatK. In addition, compound 57 (Fig. 25) showed an excellent pharmacokinetic profile with the possibility to be administrated orally.¹²³ Researchers from Hanlim

Pharmaceutical Co., LTD. designed a series of compounds employing 2-cyanopyrimidine as the reactive group in urea-based compounds (58), reaching a nanomolar range of inhibition; the company patented the resulting series containing more than 190 compounds.¹²⁴

Conclusions

Nitriles are present in many approved drugs, most of which take advantage of their ability to perform different types of noncovalent interactions with their respective targets. In addition, incorporating a cyano group into a drug molecule candidate can also improve pharmacokinetic properties.

Besides these characteristics, introducing the nitrile group in a compound can also result in covalent interactions with the macromolecule. This is an attractive approach for designing new molecules since forming a covalent bond between an inhibitor and its target presents remarkable benefits. In particular, using nitriles as the reactive center within a compound may result in the formation of a reversible covalent adduct, which is more compelling than the irreversible counterpart in terms of toxicity and selectivity.

Another great use of nitriles comes from the exploration of its electron-withdrawing property, for instance, by attaching it to an acrylamide, converting an irreversible warhead into a reversible covalent reactive center. This is a valuable strategy for developing reversible covalent inhibitors for kinases. Likewise, the use of α -cyanoacrylamides in PROTACs seems to have a bright future for treatments exploring protein degradation.

The reactivity of the nitrile warhead can be modulated by changing the atoms or groups close to the reactive center, which allows for the synthesis of a wide range of groups with increasing reactivity that can be explored to achieve a desired potency and/or selectivity. Furthermore, even less reactive

Review

warheads, such as aminoacetonitriles, have been observed to make covalent bonds with less reactive groups, such as lysines,¹²⁵ showing that, with an appropriate orientation of the inhibitor on the target enzyme, covalent inhibition can be achieved while avoiding the use of more reactive warheads.

We have presented several examples of covalent inhibitors and approved drugs that contain nitriles and derivatives as their electrophilic moiety. We have shown that including nitriles in inhibitors is a versatile strategy that can be used for a wide range of biological targets and in different contexts, such as the conversion of non-covalent to covalent inhibitors, from irreversible to reversible inhibitors, or by modulating the reactivity profile of other warheads. Therefore, we expect to observe more nitrile-based or nitrile-modified drug candidates being disclosed in the coming years, given the wide interest of the medicinal chemistry community in kinase enzymes and also in the recent interest in cysteine proteases, especially after the successful development of nirmatrelvir.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We acknowledge the Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP grants #2021/01633-3, #2020/04653-2, and 2018/15904-6. The authors also thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) #306708/2020-5, #304030/2018-0 and #305182/2021-8.

References

- 1 DrugBank, <https://go.drugbank.com/>, (accessed May 2022).
- 2 X. Wang, Y. Wang, X. Li, Z. Yu, C. Song and Y. Du, *RSC Med. Chem.*, 2021, **12**, 1650–1671.
- 3 F. F. Fleming, L. Yao, P. C. Ravikumar, L. Funk and B. C. Shook, *J. Med. Chem.*, 2010, **53**, 7902–7917.
- 4 Y. Wang, Y. Du and N. Huang, *Future Med. Chem.*, 2018, **10**, 2713–2728.
- 5 J.-Y. le Questel, M. Berthelot and C. Laurence, *J. Phys. Org. Chem.*, 2000, **13**, 347–358.
- 6 P. Ábrányi-Balogh and G. M. Keserű, in *Advances in Chemical Proteomics*, ed. X. Yao, Elsevier, 2022, pp. 47–73.
- 7 M. Gehringer and S. A. Laufer, *J. Med. Chem.*, 2019, **62**, 5673–5724.
- 8 F. M. Ferguson and N. S. Gray, *Nat. Rev. Drug Discovery*, 2018, **17**, 353–377.
- 9 A. Abdeldayem, Y. S. Raouf, S. N. Constantinescu, R. Moriggl and P. T. Gunning, *Chem. Soc. Rev.*, 2020, **49**, 2617–2687.
- 10 J. T. Hunt, C. Z. Ding, R. Batorsky, M. Bednarz, R. Bhide, Y. Cho, S. Chong, S. Chao, J. Gullo-Brown, P. Guo, S. H. Kim, F. Y. F. Lee, K. Leftheris, A. Miller, T. Mitt, M. Patel, B. A. Penhallow, C. Ricca, W. C. Rose, R. Schmidt, W. A. Slusarchyk, G. Vite and V. Manne, *J. Med. Chem.*, 2000, **43**, 3587–3595.
- 11 P. D. A. Singh, J. R. Jackson and S. P. James, *Biochem. Pharmacol.*, 1985, **34**, 2207–2209.
- 12 S. Brogi, R. Ibba, S. Rossi, S. Butini, V. Calderone, S. Gemma and G. Campiani, *Molecules*, 2022, **27**, 2561.
- 13 C. McWhirter, in *Annual Reports in Medicinal Chemistry*, ed. R. A. Ward and N. P. Grimster, Academic Press, 2021, vol. 56, pp. 1–31.
- 14 E. Awoonor-Williams, A. G. Walsh and C. N. Rowley, *Biochim. Biophys. Acta, Proteins Proteomics*, 2017, **1865**, 1664–1675.
- 15 J. Singh, R. C. Petter, T. A. Baillie and A. Whitty, *Nat. Rev. Drug Discovery*, 2011, **10**, 307–317.
- 16 P. Cohen, D. Cross and P. A. Jänne, *Nat. Rev. Drug Discovery*, 2021, **20**, 551–569.
- 17 Y. N. Lamb, *Drugs*, 2022, **82**, 585–591.
- 18 H. A. Blair, *Drugs*, 2021, **81**, 1573–1579.
- 19 F. Sutanto, M. Konstantinidou and A. Dömling, *RSC Med. Chem.*, 2020, **11**, 876–884.
- 20 S. de Cesco, J. Kurian, C. Dufresne, A. K. Mittermaier and N. Moitessier, *Eur. J. Med. Chem.*, 2017, **138**, 96–114.
- 21 J. Singh, *J. Med. Chem.*, 2022, **65**, 5886–5901.
- 22 J. Pettinger, K. Jones and M. D. Cheeseman, *Angew. Chem., Int. Ed.*, 2017, **56**, 15200–15209.
- 23 L. H. Jones, in *Annual Reports in Medicinal Chemistry*, ed. R. A. Ward and N. P. Grimster, Academic Press, 2021, vol. 56, pp. 95–134.
- 24 S. M. Marino and V. N. Gladyshev, *J. Biol. Chem.*, 2012, **287**, 4419–4425.
- 25 U. P. Dahal, A. M. Gilbert, R. S. Obach, M. E. Flanagan, J. M. Chen, C. Garcia-Irizary, J. T. Starr, B. Schuff, D. P. Uccello and J. A. Young, *MedChemComm*, 2016, **7**, 864–872.
- 26 V. Bonatto, A. Shamim, F. R. Rocho, A. Leitão, F. J. Luque, J. Lameira and C. A. Montanari, *J. Chem. Inf. Model.*, 2021, **61**, 4733–4744.
- 27 G. Oanca, M. Asadi, A. Saha, B. Ramachandran and A. Warshel, *J. Phys. Chem. B*, 2020, **124**, 11349–11356.
- 28 D. W. Kneller, G. Phillips, K. L. Weiss, Q. Zhang, L. Coates and A. Kovalevsky, *J. Med. Chem.*, 2021, **64**, 4991–5000.
- 29 A. M. Santos, L. Cianni, D. Vita, F. Rosini, A. Leitão, C. A. Loughton, J. Lameira and C. A. Montanari, *Phys. Chem. Chem. Phys.*, 2018, **20**, 24317–24328.
- 30 D. Mondal and A. Warshel, *Biochemistry*, 2020, **59**, 4601–4608.
- 31 J. M. Bradshaw, J. M. McFarland, V. O. Paaivilainen, A. Bisconte, D. Tam, V. T. Phan, S. Romanov, D. Finkle, J. Shu, V. Patel, T. Ton, X. Li, D. G. Loughhead, P. A. Nunn, D. E. Karr, M. E. Gerritsen, J. O. Funk, T. D. Owens, E. Verner, K. A. Brameld, R. J. Hill, D. M. Goldstein and J. Taunton, *Nat. Chem. Biol.*, 2015, **11**, 525–531.
- 32 X. Zhai and T. D. Meek, *Biochemistry*, 2018, **57**, 3176–3190.
- 33 J. B. Moon, R. S. Coleman and R. P. Hanzlik, *J. Am. Chem. Soc.*, 1986, **108**, 1350–1351.

- 34 T.-C. Liang and R. H. Abeles, *Arch. Biochem. Biophys.*, 1987, **252**, 626–634.
- 35 B. J. Gour-Salin, A. C. Storer, A. Castelhana, A. Krantz and V. Robinson, *Enzyme Microb. Technol.*, 1991, **13**, 408–411.
- 36 I. M. Serafimova, M. A. Pufall, S. Krishnan, K. Duda, M. S. Cohen, R. L. Maglathlin, J. M. McFarland, R. M. Miller, M. Frödin and J. Taunton, *Nat. Chem. Biol.*, 2012, **8**, 471–476.
- 37 S. Krishnan, R. M. Miller, B. Tian, R. D. Mullins, M. P. Jacobson and J. Taunton, *J. Am. Chem. Soc.*, 2014, **136**, 12624–12630.
- 38 V. Bonatto, P. H. J. Batista, L. Cianni, D. Vita, D. G. Silva, R. Cedron, D. Y. Tezuka, S. Albuquerque, C. B. Moraes, C. H. Franco, J. Lameira, A. Leitão and C. A. Montanari, *RSC Med. Chem.*, 2020, **11**, 1275–1284.
- 39 I. Podgorski, *Future Med. Chem.*, 2009, **1**, 21–34.
- 40 L. H. Jones, *RSC Chem. Biol.*, 2020, **1**, 298–304.
- 41 L. Cianni, F. R. Rocho, V. Bonatto, F. C. P. Martins, J. Lameira, A. Leitão, C. A. Montanari and A. Shamim, *Bioorg. Med. Chem.*, 2021, **29**, 115827.
- 42 L. Cianni, F. R. Rocho, F. Rosini, V. Bonatto, J. F. R. Ribeiro, J. Lameira, A. Leitão, A. Shamim and C. A. Montanari, *Bioorg. Chem.*, 2020, **101**, 104039.
- 43 J. K. Eaton, R. A. Ruberto, A. Kramm, V. S. Viswanathan and S. L. Schreiber, *J. Am. Chem. Soc.*, 2019, **141**, 20407–20415.
- 44 J. K. Eaton, L. Furst, R. A. Ruberto, D. Moosmayer, A. Hilpmann, M. J. Ryan, K. Zimmermann, L. L. Cai, M. Niehues, V. Baddock, A. Kramm, S. Chen, R. C. Hillig, P. A. Clemons, S. Gradl, C. Montagnon, K. E. Lazarski, S. Christian, B. Bajrami, R. Neuhaus, A. L. Eheim, V. S. Viswanathan and S. L. Schreiber, *Nat. Chem. Biol.*, 2020, **16**, 497–506.
- 45 P. K. Chattaraj, S. Giri and S. Duley, *Chem. Rev.*, 2011, **111**, PR43–PR75.
- 46 R. M. Oballa, J.-F. Truchon, C. I. Bayly, N. Chauret, S. Day, S. Crane and C. Berthelette, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 998–1002.
- 47 V. Ehmke, J. E. Q. Quinsaat, P. Rivera-Fuentes, C. Heindl, C. Freymond, M. Rottmann, R. Brun, T. Schirmeister and F. Diederich, *Org. Biomol. Chem.*, 2012, **10**, 5764–5768.
- 48 R. A. Ward, in *Reference Module in Biomedical Sciences*, ed. O. Wolkenhauer, Academic Press, Oxford, 2021, pp. 174–189.
- 49 P. Ábrányi-Balogh, L. Petri, T. Imre, P. Szijj, A. Scarpino, M. Hrast, A. Mitrović, U. P. Fonovič, K. Németh, H. Barreateau, D. I. Roper, K. Horváti, G. G. Ferenczy, J. Kos, J. Ilaš, S. Gobec and G. M. Keserű, *Eur. J. Med. Chem.*, 2018, **160**, 94–107.
- 50 A. Berteotti, F. Vacondio, A. Lodola, M. Bassi, C. Silva, M. Mor and A. Cavalli, *ACS Med. Chem. Lett.*, 2014, **5**, 501–505.
- 51 A. Keeley, P. Ábrányi-Balogh and G. M. Keserű, *MedChemComm*, 2019, **10**, 263–267.
- 52 R. G. Pearson, *J. Am. Chem. Soc.*, 1963, **85**, 3533–3539.
- 53 J. Breidenbach, C. Lemke, T. Pillaiyar, L. Schäkel, G. al Hamwi, M. Dieltz, R. Gedschold, N. Geiger, V. Lopez, S. Mirza, V. Namasivayam, A. C. Schiedel, K. Sylvester, D. Thimm, C. Vielmuth, L. Phuong Vu, M. Zyulina, J. Bodem, M. Gütschow and C. E. Müller, *Angew. Chem., Int. Ed.*, 2021, **60**, 10423–10429.
- 54 R. Löser, M. Frizler, K. Schilling and M. Gütschow, *Angew. Chem., Int. Ed.*, 2008, **47**, 4331–4334.
- 55 J. S. Martin, C. J. MacKenzie, D. Fletcher and I. H. Gilbert, *Bioorg. Med. Chem.*, 2019, **27**, 2066–2074.
- 56 P. Veken, A. Haemers and K. Augustyns, *Curr. Top. Med. Chem.*, 2007, **7**, 621–635.
- 57 J. E. Foley and B. Ahrén, *Eur. Endocrinol.*, 2017, **13**, 56–61.
- 58 T. E. Hughes, M. D. Mone, M. E. Russell, S. C. Weldon and E. B. Villhauer, *Biochemistry*, 1999, **38**, 11597–11603.
- 59 E. B. Villhauer, J. A. Brinkman, G. B. Naderi, B. F. Burkey, B. E. Dunning, K. Prasad, B. L. Mangold, M. E. Russell and T. E. Hughes, *J. Med. Chem.*, 2003, **46**, 2774–2789.
- 60 J. P. Berger, R. SinhaRoy, A. Poci, T. M. Kelly, G. Scapin, Y.-D. Gao, K. A. D. Pryor, J. K. Wu, G. J. Eiermann, S. S. Xu, X. Zhang, D. A. Tatosian, A. E. Weber, N. A. Thornberry and R. D. Carr, *Endocrinol., Diabetes Metab.*, 2018, **1**, e00002.
- 61 W. J. Metzler, J. Yanchunas, C. Weigelt, K. Kish, H. E. Klei, D. Xie, Y. Zhang, M. Corbett, J. K. Tamura, B. He, L. G. Hamann, M. S. Kirby and J. Marcinkeviciene, *Protein Sci.*, 2008, **17**, 240–250.
- 62 Y.-H. Wang, F. Zhang, H. Diao and R. Wu, *ACS Catal.*, 2019, **9**, 2292–2302.
- 63 P. J. Tummino and R. A. Copeland, *Biochemistry*, 2008, **47**, 5481–5492.
- 64 A. A. T. Naqvi, K. Fatima, T. Mohammad, U. Fatima, I. K. Singh, A. Singh, S. M. Atif, G. Hariprasad, G. M. Hasan and Md. I. Hassan, *Biochim. Biophys. Acta, Mol. Basis Dis.*, 2020, **1866**, 165878.
- 65 R. L. Hoffman, R. S. Kania, M. A. Brothers, J. F. Davies, R. A. Ferre, K. S. Gajiwala, M. He, R. J. Hogan, K. Kozminski, L. Y. Li, J. W. Lockner, J. Lou, M. T. Marra, L. J. Mitchell, B. W. Murray, J. A. Nieman, S. Noell, S. P. Planken, T. Rowe, K. Ryan, G. J. Smith, J. E. Solowiej, C. M. Steppan and B. Taggart, *J. Med. Chem.*, 2020, **63**, 12725–12747.
- 66 D. R. Owen, C. M. N. Allerton, A. S. Anderson, L. Aschenbrenner, M. Avery, S. Berritt, B. Boras, R. D. Cardin, A. Carlo, K. J. Coffman, A. Dantonio, L. Di, H. Eng, R. Ferre, K. S. Gajiwala, S. A. Gibson, S. E. Greasley, B. L. Hurst, E. P. Kadar, A. t. S. Kalgutkar, J. C. Lee, J. Lee, W. Liu, S. W. Mason, S. Noell, J. J. Novak, R. S. Obach, K. Ogilvie, N. C. Patel, M. Pettersson, D. K. Rai, M. R. Reese, M. F. Sammons, J. G. Sathish, R. S. P. Singh, C. M. Steppan, A. E. Stewart, J. B. Tuttle, L. Updyke, P. R. Verhoest, L. Wei, Q. Yang and Y. Zhu, *Science*, 2021, **374**, 1586–1593.
- 67 B. Halford, *ACS Cent. Sci.*, 2022, **8**, 405–407.
- 68 B. Bai, E. Arutyunova, M. B. Khan, J. Lu, M. A. Joyce, H. A. Saffran, J. A. Shields, A. S. Kandadai, A. Belovodskiy, M. Hena, W. Vuong, T. Lamer, H. S. Young, J. C. Vederas, D. L. Tyrrell, M. J. Lemieux and J. A. Nieman, *RSC Med. Chem.*, 2021, **12**, 1722–1730.
- 69 S. Vankadara, M. D. Dawson, J. Y. Fong, Q. Y. Oh, Q. A. Ang, B. Liu, H. Y. Chang, J. Koh, X. Koh, Q. W. Tan, J. Joy and C. S. B. Chia, *ACS Med. Chem. Lett.*, 2022, **13**, 1345–1350.

- 70 D. W. Kneller, H. Li, G. Phillips, K. L. Weiss, Q. Zhang, M. A. Arnould, C. B. Jonsson, S. Surendranathan, J. Parvathareddy, M. P. Blakeley, L. Coates, J. M. Louis, P. Bonnesen and A. Kovalevsky, *Nat. Commun.*, 2022, **13**, 2268.
- 71 L. J. Crofford, L. E. Nyhoff, J. H. Sheehan and P. L. Kendall, *Expert Rev. Clin. Immunol.*, 2016, **12**, 763–773.
- 72 L. Sung, in *Annual Reports in Medicinal Chemistry*, ed. R. A. Ward and N. P. Grimster, Academic Press, 2021, vol. 56, pp. 33–74.
- 73 T. D. Owens, K. A. Brameld, E. J. Verner, T. Ton, X. Li, J. Zhu, M. R. Masjedizadeh, J. M. Bradshaw, R. J. Hill, D. Tam, A. Bisconte, E. O. Kim, M. Francesco, Y. Xing, J. Shu, D. Karr, J. LaStant, D. Finkle, N. Loewenstein, H. Haberstock-Debic, M. J. Taylor, P. Nunn, C. L. Langrish and D. M. Goldstein, *J. Med. Chem.*, 2022, **65**, 5300–5316.
- 74 R. M. Miller, V. O. Paavilainen, S. Krishnan, I. M. Serafimova and J. Taunton, *J. Am. Chem. Soc.*, 2013, **135**, 5298–5301.
- 75 L. Parsons, Sanofi's oral BTK inhibitor rilzabrutinib fails Phase III trial, https://www.pharmatimes.com/news/sanofis_oral_btk_inhibitor_rilzabrutinib_fails_phase_iii_trial_1376448, (accessed May 2022).
- 76 Sanofi, Rilzabrutinib granted FDA Fast Track Designation for treatment of immune thrombocytopenia, <https://www.sanofi.com/en/media-room/press-releases/2020/2020-11-18-07-15-00>, (accessed May 2022).
- 77 M. E. Schnute, S. E. Benoit, I. P. Buchler, N. Caspers, M. L. Grapperhaus, S. Han, R. Hotchandani, N. Huang, R. O. Hughes, B. M. Juba, K.-H. Kim, E. Liu, E. McCarthy, D. Messing, J. S. Miyashiro, S. Mohan, T. N. O'Connell, J. F. Ohren, M. D. Parikh, M. Schmidt, S. R. Selness, J. R. Springer, V. Thanabal, J. I. Trujillo, D. P. Walker, Z.-K. Wan, J. M. Withka, A. J. Wittwer, N. L. Wood, L. Xing, C. W. Zapf and J. Douhan, *ACS Med. Chem. Lett.*, 2019, **10**, 80–85.
- 78 M. Forster, A. Chaikuad, S. M. Bauer, J. Holstein, M. B. Robers, C. R. Corona, M. Gehringer, E. Pfaffenrot, K. Ghoreschi, S. Knapp and S. A. Laufer, *Cell Chem. Biol.*, 2016, **23**, 1335–1340.
- 79 A. Casimiro-Garcia, J. I. Trujillo, F. Vajdos, B. Juba, M. E. Banker, A. Aulabaugh, P. Balbo, J. Bauman, J. Chrencik, J. W. Coe, R. Czerwinski, M. Dowty, J. D. Knafels, S. Kwon, L. Leung, S. Liang, R. P. Robinson, J.-B. Telliez, R. Unwalla, X. Yang and A. Thorarensen, *J. Med. Chem.*, 2018, **61**, 10665–10699.
- 80 M. Békés, D. R. Langley and C. M. Crews, *Nat. Rev. Drug Discovery*, 2022, **21**, 181–200.
- 81 R. Gabizon, A. Shraga, P. Gehrtz, E. Livnah, Y. Shorer, N. Gurwicz, L. Avram, T. Unger, H. Aharoni, S. Albeck, A. Brandis, Z. Shulman, B.-Z. Katz, Y. Herishanu and N. London, *J. Am. Chem. Soc.*, 2020, **142**, 11734–11742.
- 82 G. Xue, J. Chen, L. Liu, D. Zhou, Y. Zuo, T. Fu and Z. Pan, *Chem. Commun.*, 2020, **56**, 1521–1524.
- 83 C. P. Tinworth, H. Lithgow, L. Dittus, Z. I. Bassi, S. E. Hughes, M. Muelbaier, H. Dai, I. E. D. Smith, W. J. Kerr, G. A. Burley, M. Bantscheff and J. D. Harling, *ACS Chem. Biol.*, 2019, **14**, 342–347.
- 84 H.-H. Otto and T. Schirmeister, *Chem. Rev.*, 1997, **97**, 133–172.
- 85 P. D. Greenspan, K. L. Clark, R. A. Tommasi, S. D. Cowen, L. W. McQuire, D. L. Farley, J. H. van Duzer, R. L. Goldberg, H. Zhou, Z. Du, J. J. Fitt, D. E. Coppa, Z. Fang, W. Macchia, L. Zhu, M. P. Capparelli, R. Goldstein, A. M. Wigg, J. R. Doughty, R. S. Bohacek and A. K. Knap, *J. Med. Chem.*, 2001, **44**, 4524–4534.
- 86 R. Löser, K. Schilling, E. Dimmig and M. Gütschow, *J. Med. Chem.*, 2005, **48**, 7688–7707.
- 87 L. Cianni, G. Sartori, F. Rosini, D. de Vita, G. Pires, B. R. Lopes, A. Leitão, A. C. B. Burtoloso and C. A. Montanari, *Bioorg. Chem.*, 2018, **79**, 285–292.
- 88 N. Asaad, P. A. Bethel, M. D. Coulson, J. E. Dawson, S. J. Ford, S. Gerhardt, M. Grist, G. A. Hamlin, M. J. James, E. v. Jones, G. I. Karoutchi, P. W. Kenny, A. D. Morley, K. Oldham, N. Rankine, D. Ryan, S. L. Wells, L. Wood, M. Augustin, S. Krapp, H. Simader and S. Steinbacher, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 4280–4283.
- 89 World Health Organization, Treating Malaria, <https://www.who.int/activities/treating-malaria>, (accessed May 2022).
- 90 E. G. Tse, M. Korsik and M. H. Todd, *Malar. J.*, 2019, **18**, 93.
- 91 A. Marella, M. Akhter, M. Shaquiquzzaman, O. Tanwar, G. Verma and M. M. Alam, *Med. Chem. Res.*, 2015, **24**, 1018–1037.
- 92 E. Nizi, A. Sferrazza, D. Fabbrini, V. Nardi, M. Andreini, R. Graziani, N. Gennari, A. Bresciani, G. Paonessa and S. Harper, *Bioorg. Med. Chem. Lett.*, 2018, **28**, 1540–1544.
- 93 R. Oliveira, R. C. Guedes, P. Meireles, I. S. Albuquerque, L. M. Gonçalves, E. Pires, M. R. Bronze, J. Gut, P. J. Rosenthal, M. Prudêncio, R. Moreira, P. M. O'Neill and F. Lopes, *J. Med. Chem.*, 2014, **57**, 4916–4923.
- 94 J. Bermudez, C. Davies, A. Simonazzi, J. Pablo Real and S. Palma, *Acta Trop.*, 2016, **156**, 1–16.
- 95 D. G. Silva, J. F. R. Ribeiro, D. Vita, L. Cianni, C. H. Franco, L. H. Freitas-Junior, C. B. Moraes, J. R. Rocha, A. C. B. Burtoloso, P. W. Kenny, A. Leitão and C. A. Montanari, *Bioorg. Med. Chem. Lett.*, 2017, **27**, 5031–5035.
- 96 J. C. Gomes, L. Cianni, J. Ribeiro, F. R. Rocho, S. C. M. Silva, P. H. J. Batista, C. B. Moraes, C. H. Franco, L. H. G. Freitas-Junior, P. W. Kenny, A. Leitão, A. C. B. Burtoloso, D. Vita and C. A. Montanari, *Bioorg. Med. Chem.*, 2019, **27**, 115083.
- 97 R. F. Lameiro, A. Shamim, F. Rosini, R. Cendron, P. H. J. Batista and C. A. Montanari, *Future Med. Chem.*, 2020, **13**, 25–43.
- 98 L. Alves, D. A. Santos, R. Cendron, F. R. Rocho, T. K. B. Matos, A. Leitão and C. A. Montanari, *Bioorg. Med. Chem.*, 2021, **41**, 116211.
- 99 J. F. R. Ribeiro, L. Cianni, C. Li, T. G. Warwick, D. Vita, F. Rosini, F. R. Rocho, F. C. P. Martins, P. W. Kenny, J. Lameira, A. Leitão, J. Emsley and C. A. Montanari, *Bioorg. Med. Chem.*, 2020, **28**, 115743.
- 100 P. Büscher, G. Cecchi, V. Jamonneau and G. Priotto, *Lancet*, 2017, **390**, 2397–2409.

- 101 R. Ettari, S. Previti, L. Tamborini, G. Cullia, S. Grasso and M. Zappalà, *Mini-Rev. Med. Chem.*, 2016, **16**, 1374–1391.
- 102 R. Ettari, L. Tamborini, I. C. Angelo, N. Micale, A. Pinto, C. de Micheli and P. Conti, *J. Med. Chem.*, 2013, **56**, 5637–5658.
- 103 C. di Chio, S. Previti, G. Amendola, R. Ravichandran, A. Wagner, S. Cosconati, U. A. Hellmich, T. Schirmeister, M. Zappalà and R. Ettari, *Eur. J. Med. Chem.*, 2022, **236**, 114328.
- 104 M. Giroud, B. Kuhn, S. Saint-Auret, C. Kuratli, R. E. Martin, F. Schuler, F. Diederich, M. Kaiser, R. Brun, T. Schirmeister and W. Haap, *J. Med. Chem.*, 2018, **61**, 3370–3388.
- 105 L. Cianni, C. W. Feldmann, E. Gilberg, M. Gütschow, L. Juliano, A. Leitão, J. Bajorath and C. A. Montanari, *J. Med. Chem.*, 2019, **62**, 10497–10525.
- 106 S. Patel, A. Homaei, H. R. El-Seedi and N. Akhtar, *Biomed. Pharmacother.*, 2018, **105**, 526–532.
- 107 B. Turk, *Nat. Rev. Drug Discovery*, 2006, **5**, 785–799.
- 108 M. M. Mohamed and B. F. Sloane, *Nat. Rev. Cancer*, 2006, **6**, 764–775.
- 109 M. Siklos, M. BenAissa and G. R. J. Thatcher, *Acta Pharm. Sin. B*, 2015, **5**, 506–519.
- 110 H.-G. Bernstein and G. Keilhoff, *Neural Regener. Res.*, 2018, **13**, 2100–2101.
- 111 H. Nakanishi, *Neural Regener. Res.*, 2020, **15**, 25–29.
- 112 M. T. Drake, B. L. Clarke, M. J. Oursler and S. Khosla, *Endocr. Rev.*, 2017, **38**, 325–350.
- 113 A. Littlewood-Evans, T. Kokubo, O. Ishibashi, T. Inaoka, B. Wlodarski, J. A. Gallagher and G. Bilbe, *Bone*, 1997, **20**, 81–86.
- 114 T. Liu, S. Luo, P. Libby and G.-P. Shi, *Pharmacol. Ther.*, 2020, **213**, 107587.
- 115 X. Ou, Y. Liu, X. Lei, P. Li, D. Mi, L. Ren, L. Guo, R. Guo, T. Chen, J. Hu, Z. Xiang, Z. Mu, X. Chen, J. Chen, K. Hu, Q. Jin, J. Wang and Z. Qian, *Nat. Commun.*, 2020, **11**, 1620.
- 116 J. Schmitz, T. Li, U. Bartz and M. Gütschow, *ACS Med. Chem. Lett.*, 2016, **7**, 211–216.
- 117 L. A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, J.-M. Plancher, G. Hartmann, D. W. Banner, W. Haap and F. Diederich, *Angew. Chem., Int. Ed.*, 2011, **50**, 314–318.
- 118 J. Lameira, V. Bonatto, L. Cianni, F. R. Rocho, A. Leitão and C. A. Montanari, *Phys. Chem. Chem. Phys.*, 2019, **21**, 24723–24730.
- 119 Y.-Y. Li, J. Fang and G.-Z. Ao, *Expert Opin. Ther. Pat.*, 2017, **27**, 643–656.
- 120 F. R. Rocho, V. Bonatto, R. F. Lameiro, J. Lameira, A. Leitão and C. A. Montanari, *Expert Opin. Ther. Pat.*, 2022, **32**, 561–573.
- 121 J. Lu, M. Wang, Z. Wang, Z. Fu, A. Lu and G. Zhang, *J. Enzyme Inhib. Med. Chem.*, 2018, **33**, 890–904.
- 122 J. Benýšek, M. Buša, P. Rubešová, J. Fanfrlík, M. Lepšík, J. Brynda, Z. Matoušková, U. Bartz, M. Horn, M. Gütschow and M. Mareš, *J. Enzyme Inhib. Med. Chem.*, 2022, **37**, 515–526.
- 123 Z. Rankovic, J. Cai, J. Kerr, X. Fradera, J. Robinson, A. Mistry, E. Hamilton, G. McGarry, F. Andrews, W. Caulfield, I. Cumming, M. Dempster, J. Waller, P. Scullion, I. Martin, A. Mitchell, C. Long, M. Baugh, P. Westwood, E. Kinghorn, J. Bruin, W. Hamilton, J. Uitdehaag, M. van Zeeland, D. Potin, L. Saniere, A. Fouquet, F. Chevallier, H. Deronzier, C. Dorleans and E. Nicolai, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 1524–1527.
- 124 Hanlim Pharmaceutical Co., LTD., *US Pat.*, 20210047301, 2021.
- 125 P. A. Smith, M. F. T. Koehler, H. S. Girgis, D. Yan, Y. Chen, Y. Chen, J. J. Crawford, M. R. Durk, R. I. Higuchi, J. Kang, J. Murray, P. Paraselli, S. Park, W. Phung, J. G. Quinn, T. C. Roberts, L. Rougé, J. B. Schwarz, E. Skippington, J. Wai, M. Xu, Z. Yu, H. Zhang, M.-W. Tan and C. E. Heise, *Nature*, 2018, **561**, 189–194.

4. On the Intrinsic Reactivity of Highly Potent Trypanocidal Cruzain Inhibitors

Reproduced from BONATTO, V.; BATISTA, P. H. J.; CIANNI, L.; DE VITA, D. ; SILVA, D. G. ; CEDRON, R.; TEZUKA, D. Y.; DE ALBUQUERQUE, S.; MORAES, C. B.; FRANCO, C. H.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A. . On the intrinsic reactivity of highly potent trypanocidal cruzain inhibitors. **RSC Medicinal Chemistry**, Cambridge, v. 11, p. 1275-1284, 2020 with permission from the Royal Society of Chemistry.

Cite this: <https://pubs.rsc.org/en/content/articlelanding/2020/md/d0md00097c/unauth>

The supporting information for this work can be found in Appendix A section.

Cite this: *RSC Med. Chem.*, 2020, **11**, 1275

On the intrinsic reactivity of highly potent trypanocidal cruzain inhibitors†

Vinicius Bonatto,^{‡,a} Pedro Henrique Jatai Batista,^{‡,a} Lorenzo Cianni,^a Daniela De Vita,^a Daniel G. Silva,^a Rodrigo Cedron,^a Daiane Y. Tezuka,^{ab} Sérgio de Albuquerque,^b Carolina Borsoi Moraes,^c Caio Haddad Franco,^{id c} Jerônimo Lameira,^{id ad} Andrei Leitão^{id *a} and Carlos A. Montanari^{id *a}

The cysteine protease cruzipain is considered to be a validated target for therapeutic intervention in the treatment of Chagas disease. Hence, peptidomimetic cruzipain inhibitors having a reactive group (known as warhead) are subject to continuous studies to discover novel antichagasic compounds. Here, we evaluated how different warheads for a set of structurally similar related compounds could inhibit the activity of cruzipain and, ultimately, their trypanocidal effect. We first investigated *in silico* the intrinsic reactivity of these compounds by applying the Fukui index to correlate it with the enzymatic affinity. Then, we evaluated their potency against *T. cruzi* (Y and Tulahuen strains), which revealed the reversible cruzain inhibitor **Neq0656** as a better trypanocidal agent ($EC_{50}^{Ystrain} = 0.1 \mu\text{M}$; $SI = 58.4$) than the current drug benznidazole ($EC_{50}^{Ystrain} = 5.1 \mu\text{M}$; $SI > 19.6$). We also measured the half-life time by HPLC analysis of three lead compounds in the presence of glutathione and cysteine to experimentally assess their intrinsic reactivity. Results clearly illustrated the reactivity trend for the warheads (azanitrile > aldehyde > nitrile), where the aldehyde displayed an intermediate intrinsic reactivity. Therefore, the aldehyde bearing peptidomimetic compounds should be subject for in-depth evaluation in the drug discovery process.

Received 28th March 2020,
Accepted 3rd August 2020

DOI: 10.1039/d0md00097c

rsc.li/medchem

1. Introduction

Chagas disease, whose etiological agent is the protozoan parasite *Trypanosoma cruzi*, is a serious health and social problem for people living in Latin America and areas previously considered non-endemic such as Japan, East Europe and North America.¹ More than 300 000 new cases are reported every year in 21 countries around the world, with an average of one million people currently infected with *T. cruzi*.^{2–4} The only two existing drugs available for the treatment of this unmet medical need, benznidazole and nifurtimox, show many side effects and high inefficiency in the chronic stage of the disease.⁵ Despite this, benznidazole

has recently been approved for therapeutic use in children under 12 in the USA. Nonetheless and beyond doubt, new drugs that are safe and efficacious are therefore critically needed. One approach consists in the discovery and development of cruzain (Cz) inhibitors, which is the major *T. cruzi* cysteine protease responsible for the survival and propagation of the protozoan parasite.⁶ Recently, we have reported different covalent reversible inhibitors of cruzain as potent trypanocidal agents.^{7–10} Moreover, we explored the effects on the affinity of cruzain inhibition by replacing a nitrile group using alternative warheads.¹¹

The active site of cruzain is V-shaped containing a Cys25, His162, and Asn182 catalytic triad. In general, the catalytic cysteine is deprotonated by the histidine, which is stabilized by the Asn175 and a Trp184 residue shielding the thiolate–imidazolium ion-pair from the solvent. The stabilized negative charge renders the active site cysteine capable of attacking the warhead of certain types of covalent inhibitors.¹² The same mechanism is present for different mammalian and protozoan cysteine proteases. In general, 19 different types of warheads have been applied for cysteine protease inhibition. However, just a few of them inhibit cruzain.^{11,12} K777, the leader of the first generation of irreversible Cz inhibitors, was withdrawn from the preclinical phase due to substantial side effects caused by its irreversible

^a Medicinal Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, Avenue Trabalhador Sancarlene, 400, 23566-590, São Carlos/SP, Brazil. E-mail: andreileitao@iqsc.usp.br, carlos.montanari@usp.br

^b Ribeirão Preto School of Pharmaceutical Sciences, University of São Paulo, Ribeirão Preto, São Paulo, Brazil

^c Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais (CNPEN), Campinas, São Paulo, Brazil

^d Laboratório de Planejamento e Desenvolvimento de Fármacos, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Rua Augusto Corrêa 01, CP 66075-110, Belém-PA, Brazil

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0md00097c

‡ These authors contributed equally to the work.

mode of action.^{12–14} The main concern is to consider that warheads require a balanced reactivity profile. They should be sufficiently reactive to form a covalent bond with the cysteine protease in the active site. Their inherent reactivity should be reduced to a necessary minimum to prevent nonspecific off-target labeling.¹⁵ As a result of the time-dependent nature of covalent inhibition, even optimized compounds with sufficient selectivity in screening panels may exert significant off-target reactivity after extended exposure times.¹² The quantification of the half-life ($t_{1/2}$) for the reaction of compounds with glutathione (GSH) or cysteine provides information about warhead electrophilicity and liability toward the putative off-target reactivity.¹⁶ A comparison of $t_{1/2}$ data across a range of compounds with different warheads provides useful information for the design of new compounds within a desired reactivity range.¹⁷ Recently, Balogh and coworkers have used 137 chemical probes with 36 different warheads in order to investigate the impact of the warheads in the reactivity and specificity of a given covalent fragment.¹⁸ In another contribution, Martin and coworkers have demonstrated through an NMR based assay that the reactivity of a covalent modifier is dependent on the amino acid residue such as cysteine, serine, tyrosine, and threonine as a nucleophile in aqueous solution.¹⁹ McGregor and coworkers have used experimental techniques such as crystallography, hydrogen/deuterium exchange, and differential scanning fluorimetry to study electrophiles that are useful for targeting oncogenic K-Ras mutant proteins.²⁰ Computational tools can also be employed to understand the

reactivity of compounds, through quantum mechanics (QM) reactivity indices.^{21,22} These indices can help to reduce the synthetic effort required during a drug discovery endeavor while increasing knowledge of how substitutions affect warhead reactivity. Herein, we have studied the intrinsic reactivity of a set of Cz inhibitors and their trypanocidal activities on two different *T. cruzi* strains resulting in four trypanocidal agents equipotent to benznidazole and one nanomolar *T. cruzi* killer acting on the Y strain. Also, we have measured the $t_{1/2}$ by HPLC for key compounds to evaluate their intrinsic reactivity in water in the presence of the cysteine nucleophilic thiol. We have found that the local electrophilicity index correlates with the Cz affinity for a set of selected pairs of compounds.

2. Results

2.1. Design, synthesis and kinetic characterization

In our previous study, we reported a dipeptidyl nitrile compound, **Neq0409**, as a reversible covalent inhibitor of Cz. The crystal structure (PDB: 4QH6) displayed the covalent bond between the residue Cys25 at the nitrile group. Kinetic characterization reveals a fast reversible mode of binding.⁷ In the second step, we explored the effects on potency of cruzain inhibition by replacing a nitrile group with alternative warheads.¹¹ Thanks to the knowledge gained in our recent SARs, we have herein designed and synthesized compounds **Neq0673** and **Neq0646**. The other compounds present in this study have already been reported in our previous work.¹¹

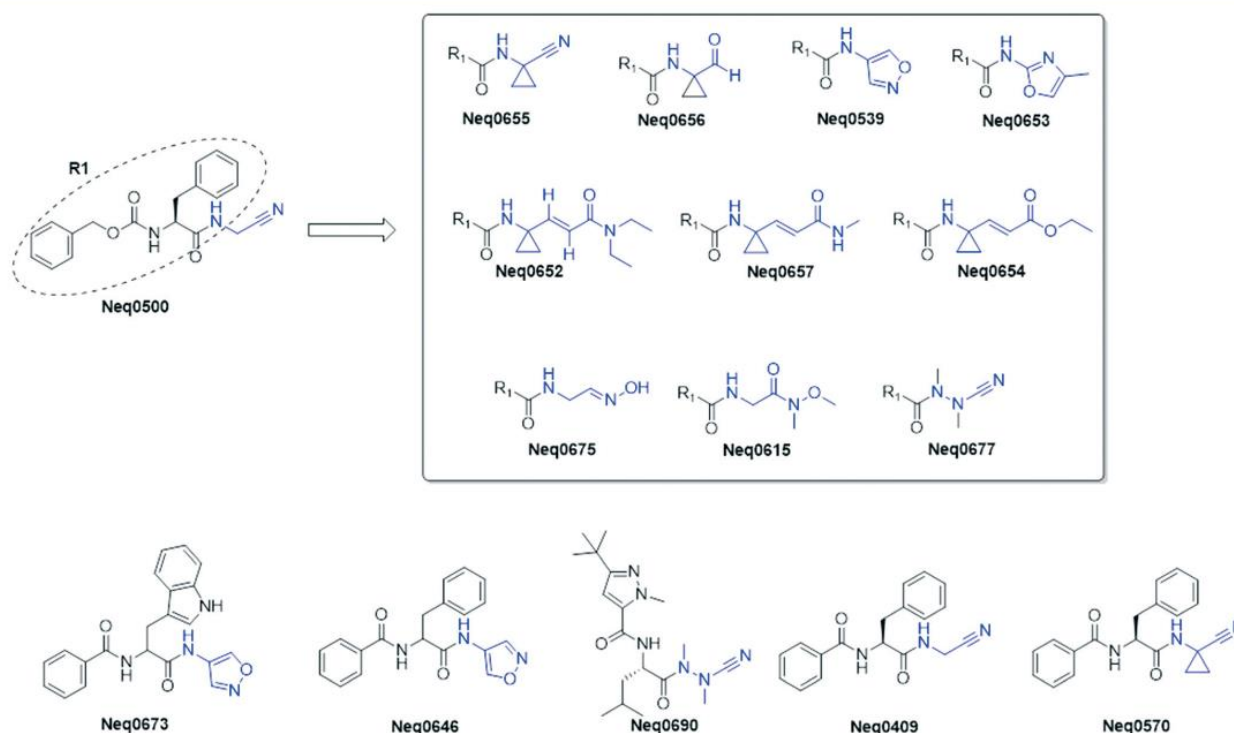


Fig. 1 2D schematic representation of the compounds presented in this study. Blue color depicts the difference in warhead structures.

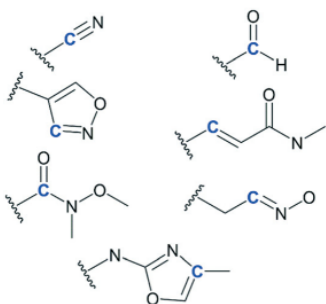


Fig. 2 Schematic representation of the warheads investigated using electrophilicity indices. Carbon (in blue) reacts with the thiolate nucleophile of cysteine. The Fukui index of electrophilicity was calculated for the carbon represented in blue.

Neq0673 and **Neq0646** bear isoxazole as a warhead, which can be seen as prodrugs for the oxime group, which has displayed an improvement in reactivity toward Cz compared with nitrile-based inhibitors. **Neq0673** and **Neq0646** contain Trp and Phe moieties in P2, respectively, which are privileged substructures for inhibition of Cz and trypanocidal activity.^{7,9} Modifications in P3 were accomplished by exchanging a carbamate (Cbz protecting group) for an amide bond with benzoic acid to increase their metabolic stability (Fig. 1).

An overview of the synthesis of compounds **Neq0646** and **Neq0673** is presented in Fig. 2. First, the 4-amine-isoxazole (**2**) was synthesized by selective nitration of isoxazole followed by reduction with SnCl₂ under acidic conditions to give the desired product. Then, the warhead (**2**) was coupled to the corresponding compound (**3** or **4**) with EDC and HOBT to afford the desired dipeptidyl isoxazole compounds. The final compounds were then purified using a HPLC equipped with a chiral column to achieve a purity higher than 95% (Fig. 1 and Scheme 1).

Cz inhibition was evaluated by fluorometric assays. The results reported in Table 1 (see below) show a wide range of affinity, spanning from high micromolar to one-digit nanomolar inhibition. As already described in our previous work,¹¹ substitution of the nitrile warhead for the azanitrile leads to a strong increase in potency ($\Delta pK_i > 2.0$) while all

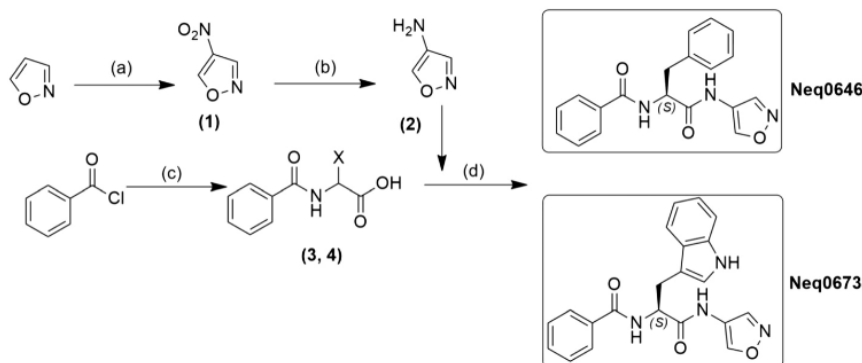
the other modifications resulted in a decrement in the reactivity profile for Cz inhibition.

2.2. Local electrophilicity index

One of the key aspects of the optimization process of covalent drugs is the modulation of reactivity of the warhead. Recently, Palazzesi and co-workers have used the electrophilicity index to estimate the absolute covalent warhead reactivity of acrylamides in aqueous solution.²³ In another contribution, Berteotti and coworkers have investigated the mechanistic cysteine nucleophilic attack on nitrile-carrying compounds using DFT calculations and kinetic measurements in a water environment. Note that the reactivity of the inhibitor can be influenced by the presence of specific interactions in the active site of the protein. However, in this work, the reactivity of compounds is explored in a water environment following the same approach of previous studies.^{23,24} The advantage of this approach is that QM calculations can estimate the reactivity only based on the electrophilicity, without including protein-inhibitor specific interaction effects, which would require long computational times. Herein, the local electrophilicity was used to predict the reactivity of warheads. We have used the Fukui function, in order to estimate the electrophilic character of the carbon (see Fig. 2) involved in the nucleophilic attack of the Cys25. Overall, we make use of the electrophilic Fukui function (f_c^+), global electrophilicity (ω) and local electrophilicity (ω_c^+) to estimate the covalent inhibitor reactivity (see Table 1). Then, we have evaluated the correlation of this parameter with the ligand binding affinity. Other electronic parameters calculated for the compounds can be found in the ESI.†

From the data shown in Table 3, we have found a strong correlation between the local electrophilicity (ω_c^+) and the ligand binding affinity, as shown in Fig. 3. The correlation between the other parameters and pK_i can be found in the ESI.†

Fig. 3 shows that a higher value of the pK_i against Cz is related to a higher value of ω_c^+ . This result is interesting because a single descriptor obtained in the aqueous phase



Scheme 1 Synthetic scheme for the synthesis of compounds **Neq0646** and **Neq0673**. Conditions: (a) TFAA, NH₄NO₃, rt, 2 h; (b) 6 M HCl, SnCl₂, 1.5 h, rt; (c) Trp-OH or Phe-OH, 1 M NaOH, 0 °C, 0.5 h; (d) EDC, HOBT, THF, rt, 18 h.

Table 1 The Mulliken charge (Q_C) of neutral and anionic forms, electrophilic Fukui function (f_c^+), global electrophilicity (ω) and local electrophilicity (ω_c^+) were computed for the compound set from which the respective pK_i against Cz was measured

Compound (Neq)	$Q_{C,Neutral}$	$Q_{C,anion}$	f_c^+	ω (eV)	ω_c^+ (eV)	pK_i^{Cz}
500	0.321	0.332	0.0108	0.183	0.0020	6.3
539	0.045	0.050	0.0048	0.136	0.0006	5
615	0.835	0.833	-0.0015	0.199	-0.0003	4
646	0.045	0.049	0.0040	0.130	0.0005	4.7
652	-0.027	-0.044	-0.0179	0.131	-0.0023	3.7
653	0.081	0.071	-0.0105	0.135	-0.0014	4.6
654	-0.023	-0.026	-0.0025	0.148	-0.0004	3.9
655	0.315	0.333	0.0183	0.087	0.0016	5.2
656	0.574	0.580	0.0051	0.177	0.0009	5.4
657	-0.036	-0.063	-0.0270	0.164	-0.0044	3.4
673	0.045	0.049	0.0040	0.128	0.0005	5
675	0.125	0.127	0.0030	0.238	0.0007	5
677	0.515	0.537	0.0221	0.183	0.0041	8.7
690	0.518	0.536	0.0181	0.189	0.0034	8.8

presented a good correlation with the inhibition of the enzyme, as other authors have done in order to obtain a descriptor correlated with the inhibition of cysteine proteases,^{24,25} even knowing that it is difficult to find a single property to describe the biochemical activity.^{21,26–28}

Compounds **Neq0539** ($pK_i = 5.0$) and **Neq0653** ($pK_i = 4.6$) are over an order of magnitude less potent than the reference nitrile **Neq0500** ($pK_i = 6.3$). A ΔpK_i value of 1.3 for the transformation is sufficiently small to be consistent with covalent bond formation. In the same direction, we observed a loss in potency ($\Delta pK_i = 2.3$) when exchanging the nitrile warhead to an amide group (**Neq0615**) which can be considered as a negative control for the covalent bond formation. The difference in potency for these compounds is nicely related to the local electrophilicity. On the other hand, the functional behaviour of warheads is sometimes assumed to be determined entirely by electrophilicity; however, the product of

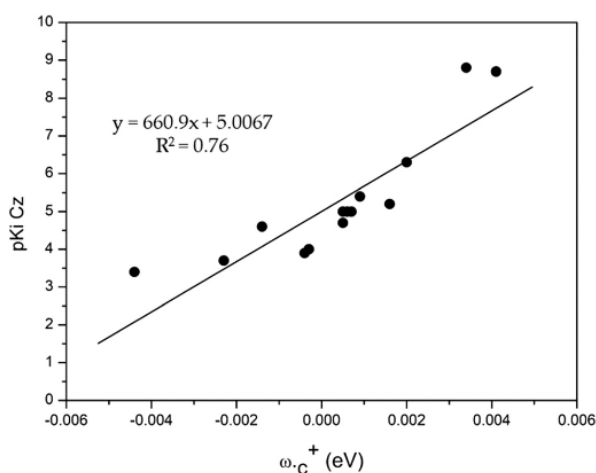


Fig. 3 The putative linear equation and the coefficient of determination obtained through the linear correlation between ω_c^+ and pK_i^{Cz} .

the covalent bond can form different non-covalent interactions with the target protein that also modulate the affinity (e.g. difference between nitrile and ketoamide dipeptidyl nitriles for cathepsin inhibition, PDB IDs: 3HHA and 1TU6).

Overall, the higher the value of local electrophilicity index for the carbon to be attacked by cysteine, the greater the stabilization of the addition of electrons; thus, this model can also be used for choosing warheads or groups that can influence the electronic properties of warheads in the design of new compounds.

2.3. Biological activities

The trypanocidal activity was then assessed for all compounds using two *T. cruzi* strains (Y and Tulahuen) as reported in Table 2, along with the cytotoxicity against the mammalian host cells (U2OS and LLC-MK2) and the respective selectivity indices (SIs). In general, biological activities show that Cz inhibitors are more potent against the *T. cruzi* Y strain than the Tulahuen one. This effect could arise from a different level of expression of Cz and its isoforms for the different strains.²⁹ The difference in potency for different strains is already known, so it is vital to emphasize the importance of testing the same set of compounds in different strains.³⁰ **Neq0500**, **Neq0539**, and **Neq0646**, the dipeptidyl nitrile based compounds, appear to be trypanocidal agents equipotent to Bz, showing a low micromolar potency against the parasite (Y strain) and no cytotoxicity to the host cell. On the other hand, these dipeptidyl nitriles are ineffective against the Tulahuen strain, as also underlined in our previous work.⁷

Compound **Neq0655** has a higher cytotoxic profile and low selectivity index when tested on the U2OS cell line. Azadipeptidyl nitriles display the same trend being equipotent to Bz as trypanocidal agents for both strains but with a selectivity index lower than 10. **Neq0675** bearing an isoxazole as warhead is inactive for the Y strain of *T. cruzi*, while its analogue (**Neq0646**) with a Phe in P2 shows mild potency and no cytotoxicity to the same strain. Interestingly, compound **Neq0656** bearing an aldehyde as a warhead is a potent trypanocidal agent for the Y strain with a high selectivity in relation to that for the mammalian cell. Although the activity of **Neq0656** is not transferable to the Tulahuen strain, we believe that this result is fundamental in the ongoing research for small molecules as potent trypanocidal agents due to the high infectivity of the Y strain.³¹ It is noteworthy that because most of the reported compounds are covalent inhibitors of Cz, their potency against the parasite is also time-dependent; hence, future studies should be focused on the time-dependency of the provided activity.

2.4. Study of the intrinsic reactivity

Covalent inhibitors binding normally involve a two-step process in which an initial reversible binding event takes place, followed by the covalent bond-forming reaction. In the first stage, enzyme E and inhibitor I form an enzyme

Table 2 Trypanocidal activity (EC₅₀), cytotoxicity (CC₅₀) and selectivity index (SI) for the set of compounds using the amastigote form of two strains (Y and Tulahuén) of *T. cruzi* in the mammalian host cells (U2OS and LLC-MK2)

Neq ID	EC ₅₀ Tc Y (μM)	CC ₅₀ U2OS (μM)	SI	EC ₅₀ Tc Tula (μM)	CC ₅₀ LLC-MK2 (μM)	SI
409 ^a	—	—	—	Inact	>100	ND
500	3.6	>100	>27.8	Inact	>100	1.1
539	5.1	>100	>19.6	25.8	>100	>3.9
570 ^a	—	—	—	Inact	>100	ND
615	ND	ND	ND	Inact	>100	ND
646	6.0	>100	>16.6	ND	ND	ND
652	ND	ND	ND	Inact	>100	ND
653	ND	ND	ND	Inact	>100	ND
654	ND	ND	ND	79.6	>100	>1.2
655	1.0	6.0	4.4	Inact	>100	ND
656	0.1	3.0	58.4	Inact	>100	ND
657	12.1	>100	>8.2	ND	ND	ND
673	22.4	>100	>8.2	ND	ND	ND
675	Inact	>100	ND	ND	ND	ND
677	3.2	10.2	3.2	21.7	42.0	1.9
690	8.4	26.9	3.2	8.4	26.9	3.2
Bz	5.1	>100	>19.6	4.3	>100	>23.3

Standard deviations are lower than 15%. TC Y: *T. cruzi* Y strain. Tc Tula: *T. cruzi* Tulahuén strain. ND: not determined. Inact: inactive compound (EC₅₀ > 100 μM). ^a Values retrieved from the recent literature.⁷

inhibitor complex (E⋯I) whose binding energy (ΔG_{bind}) depends only on non-covalent interactions and is related to the inhibition constant K_i . The subsequent binding step (k_{inact}) involves the formation of a covalent bond between the inhibitor and target. The latter step will be partly governed by the intrinsic reactivity of the warhead.^{15,32} In our case, the measured K_i values mostly reflect the binding free energy for non-covalent interactions and the covalent bond formation. It is difficult to define the contribution of the covalent bond formation without the k_{inact} values. Indeed, these values are influenced by several factors other than just the chemical nature of the warhead. Also, depending on the magnitude of the non-covalent interactions with respect to the total potency outcome, we should not expect a direct correlation between the intrinsic reactivity, measured in the absence of the target enzyme, and the K_i .

However, to better evaluate the promiscuity of the different warheads independent of their biological target, the local electrophilicity (ω_c^+) and intrinsic reactivity are still pivotal parameters in drug discovery.

The correlation achieved between the local electrophilicity (ω_c^+) and the enzyme inhibition led us to investigate whether

these compounds could be promiscuous electrophiles due to the high reactivity. Hence, the intrinsic reactivity (given as the half-life – $t_{1/2}$) was assessed using a method described recently,¹⁷ where cysteine and glutathione were the non-specific nucleophiles. Glutathione is present in high concentration (1–10 mM) in the cell.³³ It works as an important component of the intracellular redox machinery, while cysteine is a building block amino acid for many proteins, and it is the reactive amino acid in the active pocket of cysteine proteases. We measured the half-life of four compounds by HPLC using two different aqueous systems, with the results displayed in Table 3. As a model for nitrile warheads, we used **Neq0409**, whose crystallographic structure was previously resolved, and **Neq0570**, for which the reversible mode of action has also been investigated.⁷ In general, we can identify a trend for the intrinsic reactivity for both nucleophiles where the reactivity was azanitrile \gg aldehyde > nitrile. In particular, the azanitrile compound **Neq0690** displayed such strong reactivity with both glutathione and cysteine that it was not possible to quantify it. This high intrinsic reactivity may explain the low selectivity between the parasite and the host cell given by the low SI

Table 3 Half-life measurement and decay constant by HPLC with glutathione and cysteine

Neq ID	pK _i	Warhead	Half-life with glutathione	Half-life with cysteine	Decay constant with glutathione	Decay constant with cysteine
409	6.3 ^a	Methylene-nitrile	>6000 min	74.3 min	—	$9.33 \times 10^{-3} \text{ min}^{-1}$
570	6.6 ^a	Cyclopropane-nitrile	>6000 min	159.0 min	—	$4.36 \times 10^{-3} \text{ min}^{-1}$
656	5.4	Aldehyde	40.0 min	22.2 min	$1.73 \times 10^{-2} \text{ min}^{-1}$	$3.12 \times 10^{-2} \text{ min}^{-1}$
690 ^b	8.8	Aza-nitrile	<5 min	<5 min	—	—
Nilvadipine	—	—	>5000 min	49.5 min	—	$1.41 \times 10^{-2} \text{ min}^{-1}$

^a Values retrieved from the recent literature.⁷ ^b When the concentrations of the compound **Neq0690** and the thiol are half of the initial values, the k_{obs} is $1.01 \times 10^{-2} \text{ min}^{-1}$ for the glutathione assay and $3.15 \times 10^{-2} \text{ min}^{-1}$ for the cysteine assay ($t_{1/2} = 68.45$ and 21.98 min respectively). The standard deviation is lower than 10% for all experiments.

values (Table 2). Nitrile based compounds have low intrinsic reactivity, which was possible to measure when cysteine is present in solution as the nucleophile. As expected, the cyclopropane moiety (as in compound **Neq0570**) led to a small decrease in reactivity when compared with the methylene-nitrile pair (**Neq0409**). On the other hand, for the Cz inhibition, we observed an opposite effect thanks to the ability of the cyclopropane moiety to drive the warhead in the active site.³⁴ Also, dipeptidyl nitriles (**Neq0409** and **Neq0570**) are ten-times more potent inhibitors of Cz than the aldehyde (**Neq0656**) but less reactive to nucleophilic attack in water solution. This difference in reactivity/potency can arise from the non-covalent interaction formed by the thiomidate with the catalytic site of the enzyme.^{7,32}

The aldehyde **Neq0656** displays a balanced reactivity for both systems under study. According to Macfaul,¹⁷ compounds should have $t_{1/2}$ at least equal to that of nilvadipine to be used in further assays. Nonetheless, **Neq0656**, with half of the $t_{1/2}$ value of nilvadipine, was still selective toward the parasite, making it a good lead structure. These results are in agreement with the cytotoxicity observed in the biological studies. Therefore, we can assume that the substitution of the nitrile warhead for the aldehyde can provide a balanced reactivity, which leads to an increase in the trypanocidal activity with a high selectivity index against the host cell.

3. Material and methods

3.1. Computational details

Herein, all calculations were performed with Gaussian09.³⁵ Geometry optimization and vibrational analysis were performed at the density functional theory (DFT) level, using the B3LYP functional^{36,37} and 6-311G(d,p) basis set³⁸ in the gas phase. After confirming that the structure was at the minimum through frequency analysis, a single point calculation at the perturbation theory of Møller–Plesset (MP2)³⁹ level and the diffuse basis set 6-311++G(d,p)³⁸ was performed in the water phase using the PCM⁴⁰ for neutral, anionic and cationic systems, keeping the external potential constant. Then, the global electrophilicity index⁴¹ (ω) was obtained by:

$$\omega = \frac{\mu^2}{2\eta} \quad (1)$$

where μ and η are the chemical potential and chemical hardness, respectively.

In the present study, we have used the Fukui index^{42–45} (f_k^+) as a parameter for the evaluation of electrophilicity of the inhibitor. The Fukui index was obtained from the population analysis⁴⁶ and charges through NBO analysis⁴⁷ at the MP2 level using the diffuse basis set 6-311++G(d,p). The f_k^+ was calculated through finite difference, as represented below:

$$f_k^+ = q_k(N+1) - q_k(N) \quad (2)$$

where $q_k(N)$ represents the Mulliken charge of atom k in the neutral molecule, and $q_k(N+1)$ corresponds to the Mulliken

charge of atom k in the same geometry of the neutral molecule, but in the anionic form. Since the covalent inhibition of Cz involves the nucleophilic attack of the negatively charged Cys25 (S atom) on the carbon atom of the warhead (see Fig. 2), it is important to have a parameter capable of estimating the reactivity of the warhead. Therefore, using electrophilicity and the Fukui index, it is possible to describe the electrophilic character of a reactive site within a molecule, through the local electrophilicity index⁴¹ $\omega(r)$:

$$\omega(r) = \omega f_k^+(r) \quad (3)$$

3.2. Synthetic chemistry

Melting points were determined on a Büchi 510 oil bath apparatus and are uncorrected. Infrared spectra were obtained from a FT-IR Thermo Scientific Nicolet 380. The reagents, starting materials and solvents were of commercial quality and were used without further purification unless otherwise stated. All syntheses started with enantiopure amino acids. TLC analysis was carried out on Merck 60 F₂₅₄ silica gel plates and visualized under UV light at 254 nm and 365 nm or by using a ninhydrin staining solution.

Purity was determined with an LC-MS instrument (AmaZon SL ESI-MS, Shimadzu LC) with a cellulose-2 Phenomenex column (250 × 4.6 mm, 5 μm) or a Daicel column (IC-chiralpak, 250 × 4.6 mm, 5 μm). Isocratic elution with MeCN and water was applied as specified (stop time 60 min, flow 0.5 mL min⁻¹). NMR spectra were recorded on Bruker Avance 400 MHz and Bruker Avance DRX 500 MHz NMR spectrometers. Chemical shifts are reported in ppm relative to TMS or the residual proton peak of the re-protonated deuterated solvent, and the spectra were calibrated against the residual proton peak of the used deuterated solvent. The following symbols indicate spin multiplicities: s (singlet), s br (broad singlet), d (doublet), dd (doublet of doublet), t (triplet), tt (triplet of triplet), q (quartet), sept (septet), and m (multiplet). HRMS spectra were recorded on a Thermo Scientific LTQ Velos Orbitrap, in electrospray ionization (ESI) mode by direct injection.

Synthesis of 4-nitroisoxazole (**1**): isoxazole (15 mmol, 960 μL) was dissolved in TFAA (7.3 mL); then, NH₄NO₃ (22.5 mmol, 1.81 g) was added in 0.3 g portions each 15 min, keeping the reaction mixture at 25–30 °C. After complete addition, the mixture was kept at room temperature for 2 h. After that ice water (30 mL) was poured and this aqueous washing was extracted with CHCl₃ (3 × 15 mL); the combined organic extracts were dried over Na₂SO₄ and evaporated (bath at room temperature) to give an oil that was triturated with *n*-hexane to give a yellow solid (50% yield). ¹H-NMR (CDCl₃) δ = 9.29 (s, 1H), 8.83 (s, 1H) ppm.⁴⁸

Synthesis of 4-aminoisoxazole (**2**): to a yellow solution of 4-nitroisoxazole (**1**, 160 mg, 1.4 mmol) in 6 M HCl (7 mL), SnCl₂ (1.327 g, 7 mmol) was added in one portion. After 1.5 h at room temperature, the resulting orange solution was treated with a saturated solution of Na₂CO₃ until the pH was

9. The formed solid was removed by filtration, and the aqueous solution was extracted with ethylacetate (5 × 50 mL); the organic phase, dried over MgSO₄, was evaporated to give a brown oil (*R_f* = 0.64, ethylacetate 100%/silica) and stored at 4 °C in an inert atmosphere (65% yield). ¹H-NMR (DMSO-d₆) δ = 8.16 (s, 1H), 8.13 (s, 1H), 4.26 (s br, 2H) ppm.

Synthesis of 2-benzamido-3-phenylpropanoic acid (**3**) and 3-(1*H*-indol-3-yl)-2-(phenylformamido)propanoic acid (**4**): 2.75 mmol of the corresponding amino acid was dissolved in 1 M NaOH (6 mL) in an ice-bath. Benzoyl chloride (261 μL, 2.25 mmol) was added. After 5 min, the reaction mixture was allowed to stand at room temperature. After 20 min, the solution was cooled in ice and 1 M KHSO₄ (16 mL) was added slowly. The obtained white solid was washed with 1 M KHSO₄ (3 × 5 mL), H₂O (10 × 3 mL), and 9:1 EtOH:H₂O (3 × 3 mL) and dried under vacuum on P₂O₅ (yield 88%). For compound **3**: ¹H NMR (500 MHz, DMSO-d₆): 12.62 (s br, 1H, D₂O exchange), 10.75 (s, 1H, D₂O exchange), 8.56 (m, 1H, D₂O exchange), 7.28 (m, 8H), 7.13 (m, 2H), 4.54 (m, 1H), 3.29 (m, 1H), 3.19 (m, 1H). For compound **4**: ¹H NMR (500 MHz, DMSO-d₆): 12.69 (s br, 1H, D₂O exchange), 10.80 (s, 1H, D₂O exchange), 8.61 (d, *J* = 8.0 Hz, 1H, D₂O exchange), 7.81 (d, *J* = 8.5 Hz, 2H), 7.59 (d, *J* = 8 Hz, 1H), 7.51 (t, *J* = 7.0 Hz, 1H), 7.44 (t, *J* = 7.5 Hz, 2H), 7.31 (d, *J* = 8.0 Hz, 1H), 7.20 (s, 1H), 7.05 (t, *J* = 7.5 Hz, 1H), 6.98 (t, *J* = 7.5 Hz, 1H), 4.65 (m, 1H), 3.30 (m, 1H, H₂O overlapping), 3.19 (m, 1H).

Synthesis of *N*-(1-(isoxazol-4-ylamino)-1-oxo-3-phenylpropan-2-yl)benzamide (**Neq0646**): to a suspension of (±)-2-benzamido-3-phenylpropanoic acid (**3**, 216 mg, 0.70 mmol), HOBt (123 mg, 0.91 mmol) and EDC (175 mg, 0.91 mmol) in CH₂Cl₂ (8 mL) were added under argon at 0 °C. After stirring for 1 hour at room temperature, the mixture was kept in an ice-bath, and a solution of 4-aminoisoxazole (235 mg, 2.80 mmol) in dry CH₂Cl₂ (2 mL) was added. The resulting mixture was kept overnight at room temperature, then the solvent was evaporated, and the residue was treated with AcOEt (30 mL) and washed with H₂O (2 × 20 mL) and brine (2 × 20 mL). The organic phase was dried over MgSO₄ and evaporated to give a crude residue which was purified by silica column chromatography using CHCl₃/AcOEt (1:1) as the mobile phase, to give a solid (*R_f* = 0.4) crystallized from AcOEt (36% yield).

Secondary purification was carried out on a cellulose-2 Phenomenex column, in isocratic elution mode with a flow rate of 2.36 mL min⁻¹, at 32 °C; the mobile phase composition was *n*-hexane/ethanol (70:30) (v/v) to give **Neq0646**. ¹H-NMR (500 MHz, DMSO-d₆) δ = 10.80 (s br, 1H), 10.51 (s, 1H), 9.13 (s, 1H), 8.72 (d, *J* = 8.0 Hz, 1H), 8.63 (s, 1H), 7.84 (m, 2H), 7.69 (d, *J* = 8.0 Hz, 1H), 7.52 (tt, *J* = 7.5 Hz, *J* = 1.5 Hz, 1H), 7.44 (m, 2H), 7.31 (dt, *J* = 8.0 Hz, *J* = 1.0 Hz, 1H), 7.23 (d, *J* = 2.0 Hz, 1H), 7.05 (m, 1H), 6.98 (m, 1H), 4.83 (qd, *J* = 9.5 Hz, *J* = 8.0 Hz, *J* = 5.0 Hz, 1H), 3.30 (dd, *J* = 14.5 Hz, *J* = 5.0 Hz, 1H), 3.22 (dd, *J* = 14.5 Hz, *J* = 9.5 Hz, 1H) ppm. ¹³C NMR (DMSO-d₆): 170.72, 166.90, 147.56, 144.85, 136.52, 134.23, 131.86, 128.63, 127.97, 127.59, 124.23, 121.41, 120.17, 118.89, 118.73, 111.83, 110.55, 54.89, 27.74 ppm.

HRMS (+) calc. for [C₁₉H₁₇N₃O₃]⁺ 335.12699, found: 336.12663 [M + H]⁺.

Synthesis of *N*-(3-(1*H*-indol-3-yl)-1-(isoxazol-4-ylamino)-1-oxopropan-2-yl)benzamide (**Neq0673**): to a suspension of (±)-2-benzamido-3-phenylpropanoic acid (216 mg, 0.70 mmol), HOBt (124 mg, 0.91 mmol) and EDC (175 mg, 0.91 mmol) in CH₂Cl₂ (8 mL) were added under argon at 0 °C. After stirring for 1 hour at room temperature, the mixture was kept in an ice-bath and a solution of 4-aminoisoxazole (235 mg, 2.80 mmol) in dry CH₂Cl₂ (2 mL) was added. The resulting mixture was kept overnight at RT, then the solvent was evaporated, and the residue was treated with AcOEt (30 mL) and washed with H₂O (2 × 20 mL) and brine (2 × 20 mL). The organic phase was dried over MgSO₄ and evaporated to give a crude residue that was purified by column chromatography on silica using CHCl₃/AcOEt (1:1) as the mobile phase, to give a solid (*R_f* = 0.4) crystallized from AcOEt (36% yield).

Secondary purification was carried out on a cellulose-2 Phenomenex column, in isocratic elution mode with a flow rate of 2.36 mL min⁻¹, at 32 °C; the mobile phase composition was *n*-hexane/ethanol (70:30) (v/v) to give **Neq0673**. ¹H-NMR (500 MHz, DMSO-d₆) δ = 10.80 (s br, 1H), 10.51 (s, 1H), 9.13 (s, 1H), 8.72 (d, *J* = 8.0 Hz, 1H), 8.63 (s, 1H), 7.84 (m, 2H), 7.69 (d, *J* = 8.0 Hz, 1H), 7.52 (tt, *J* = 7.5 Hz, *J* = 1.5 Hz, 1H), 7.44 (m, 2H), 7.31 (dt, *J* = 8.0 Hz, *J* = 1.0 Hz, 1H), 7.23 (d, *J* = 2.0 Hz, 1H), 7.05 (m, 1H), 6.98 (m, 1H), 4.83 (qd, *J* = 9.5 Hz, *J* = 8.0 Hz, *J* = 5.0 Hz, 1H), 3.30 (dd, *J* = 14.5 Hz, *J* = 5.0 Hz, 1H), 3.22 (dd, *J* = 14.5 Hz, *J* = 9.5 Hz, 1H) ppm. ¹³C NMR (DMSO-d₆): 170.72, 166.90, 147.56, 144.85, 136.52, 134.23, 131.86, 128.63, 127.97, 127.59, 124.23, 121.41, 120.17, 118.89, 118.73, 111.83, 110.55, 54.89, 27.74 ppm. HRMS (+) calc. for [C₂₁H₁₈N₄O₃]⁺ 374.13789, found: 375.13895 [M + H]⁺.

3.3. Enzyme inhibition study

Recombinant cruzain, consisting of the catalytic domain of cruzipain but excluding the carboxy-terminal extension, was expressed and purified as previously described.¹¹ The inhibitors were assayed fluorometrically (Biotek® Synergy™ HT), monitoring the hydrolysis rate of the fluorogenic substrate Z-Phe-Arg-7-amido-4-methylcoumarin (Z-FR-AMC, Sigma-Aldrich) by the enzyme cruzain with fluorescence emission at 460 nm (excitation at 355 nm) and at 37 °C. The reactions were followed over 5 min for all compounds (fast-binding, irreversible and non-covalent inhibitors) excluding **Neq0690** and **Neq0677** (slow-binding behaviour) for which the reaction was followed over 30 min. Enzyme kinetic assays were carried out in 200 μL of a solution containing 100 mM acetate buffer pH 5.5, 300 mM NaCl, 5 mM DTT (dithiothreitol), 5% v/v DMSO (dimethyl sulfoxide), 0.01% v/v Triton X-100 and 0.15 nM cruzain, using Corning® 96-well black flat bottom microplates. The enzyme stock aliquot was rapidly thawed at 37 °C and kept on ice until activation, in which it was incubated for 20 min in the assay buffer (100 mM acetate pH 5.5, 300 mM NaCl, 5 mM DTT) followed by

Research Article

an additional 2 min with inhibitors before the reaction was started by the addition of the substrate.

Visual inspection and a pre-reading of plate wells were performed to check respectively for possible precipitation and background fluorescence. None of the compounds displayed a significant fluorescence signal around 460 nm (the emission wavelength used to monitor reaction kinetics).

3.4. Mammalian cytotoxicity assay

LLC-MK2 cells were cultured in 96-well plates at a concentration of 5×10^4 cells per mL. After 48 h, the plates were washed twice with PBS, and 200 μ L RPMI medium was added with serial dilutions of the compounds and benznidazole (1.95 μ M to 250 μ M) in triplicate. After 72 h at 37 °C, the cytotoxic activity of the compounds was determined by the classical MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] method. Briefly, 50 μ L MTT dissolved in PBS (2.0 mg mL⁻¹) was added to each well, and the plates were incubated for 4 h at 37 °C. The formed formazan crystals were dissolved with DMSO (50 μ L per well), and the absorbance of the samples was measured using a spectrophotometer at 570 nm after 30 min. The cytotoxicity results (CC₅₀) were calculated as a percentage by the formula $\{[(\text{ABS}_{\text{control}} - \text{ABS}_{\text{sample}})/\text{ABS}_{\text{control}}] \times 100\}$, where ABS_{control} represents the mean absorbance of the untreated control (viable cells) and ABS_{sample}, the absorbance in each cellular treatment.

The U2OS cells were kept in high glucose DMEM media, and the culture conditions and assays were the same for LLC-MK2.

3.5. *In vitro* trypanocidal activity evaluation on intracellular amastigote forms (Tulahuen strain)

Cells were evaluated in 96-well plates. LLC-MK2 cells were plated at a concentration of 5×10^4 cells per mL. Trypomastigote forms of the Tulahuen LacZ strain were added at a concentration of 5×10^5 parasites per mL and placed in an incubator at 37 °C with 5% CO₂ for 24 h. After the incubation period, the trypomastigote forms present were removed by successive washes with PBS, remaining only as intracellular amastigote forms. The compounds were added at different concentrations (1.95 μ M to 250 μ M serial dilutions) and incubated for 72 h. At the end of this period, the substrate CPRG (chlorophenol red β -D-galactopyranoside, 400 μ M in 0.3% Triton X-100, pH 7.4) was added. After 4 h of incubation at 37 °C, the plates were analyzed on a spectrophotometer at 570 nm to obtain the effective concentration (EC₅₀) to reduce the parasitemia inside the host cell. Benznidazole was used as a positive control at the same concentrations as the substances, and DMSO as a negative control. The compounds were solubilized in DMSO. The selectivity index (SI) was calculated using the formula: $\text{SI} = \text{EC}_{50}/\text{CC}_{50}$. All statistical analyses were done with the program GraphPad Prism v.5.

3.6. *In vitro* trypanocidal activity evaluation on intracellular amastigote forms (Y strain)

The *T. cruzi* Y strain was donated by A. Avila (Instituto Carlos Chagas, Fiocruz, Curitiba, Brazil). Trypomastigote forms were obtained from the supernatant of LLC-MK2 tissue cultures infected with the *Trypanosoma cruzi* Y. Infected cultures were maintained in low glucose DMEM media (Vitrocell) supplemented with 2% FBS, 100 U mL⁻¹ penicillin and 100 μ g mL⁻¹ streptomycin (all from Life Technologies), henceforth described as "Low DMEM Media". The experiment was performed as previously described.¹¹

3.7. Intrinsic reactivity

Liquid chromatography assay measured the half-life in a gradient mode (5–100% of B in 10 minutes) with multi-channel PDA detection (210–400 nm). The assay medium was 0.05 M phosphate buffer (pH 7.4), with 1.0 mM EDTA, 5.0 mM cysteine solution and 5% acetonitrile. Initially, 80 μ L of inhibitor solution (2.5 mM) was added in the assay medium and injected in the HPLC. After that, the final solution was kept in thermal equilibrium at 37 °C. Each aliquot was taken only at the time of injection. The half-life and *K* were determined with GraphPad® software for a pseudo-first-order reaction. The chromatographic analysis was performed with a Shimadzu Prominence HPLC system. The HPLC system was equipped with an LC-20AT/AD ternary pumping system, SIL-20A autosampler, and CTO-20A column oven (Shimadzu Corp). The column used was a Phenomenex Gemini® 5.0 μ m (4.6 \times 150 mm, Phenomenex, Torrance, CA). The gradient elution, used for separation, was performed with a mobile phase composed of water as solvent A, and acetonitrile, as solvent B, at a flow rate of 1.0 mL min⁻¹. The gradient set is as follows: 0.0–10.00 min phase B increased from 5% to 100%, 10.01–15.0 min phase B remained at 100%, 15.01 min phase B decreased to 5%, and 15.01–25.0 min phase B remained at 5%. The column oven was set at 37 °C, with the injection of 5 μ L aliquots. The cysteine employed was of the levogyre form (L-cysteine), MW: 121.16 g mol⁻¹ (C₃H₁₇NO₂S) sold by the brand Sigma-Aldrich (code C7352-25G), \geq 98% from a non-animal source. The glutathione used was purchased from Sigma-Aldrich (PHR1359-690MG).

The glutathione and cysteine solutions were prepared with a 50 mM final concentration in phosphate buffer with pH 7.4 and 50 mM concentration. Data have been processed with GraphPad Prism® using the one phase decay model: $Y = (Y_0 - \text{plateau}) \times \exp(-K \times X) + \text{plateau}$. For each experiment, a negative control without glutathione or cysteine was assessed to measure the stability of the compounds and their warheads in the time frame of the experiment.

4. Conclusions

In this study, we evaluated how the intrinsic reactivity of a set of structurally similar related compounds influenced the biological outcome of trypanocidal agents. Currently,

computational methods are vital in the design of new chemical entities, but there is still little effort to estimate the intrinsic reactivity of warheads. Here, we report a successful approach to using the Fukui index to correlate anti-*T. cruzi* action of trypanocidal agents with different warheads. It is worth mentioning that we have discovered a new compound, **Neq0656**, which is 10 times more potent than the control drug, benzimidazole, as a trypanocidal agent. **Neq0656** also shows a higher selectivity index than benzimidazole. This result underlines that warhead replacement of nitrile to aldehyde is an innovative strategy in the research of potent trypanocidal agents against the Y strain of *T. cruzi*. Indeed, most of the compounds here tested are also selective toward the Y strain, in comparison with the Tulahuén strain.³⁰ This discrepancy may arise from the difference in the expression of Cz and its isoforms. We also investigated the intrinsic reactivity by HPLC to clearly illustrate that azanitrile warheads are highly reactive moieties. The higher reactivity with respect to that of nitrile can be explained by the presence of nitrogen in the alpha position, as already mentioned by Gütschow and co-workers.⁴⁹ Most of them, aldehydes, which have an intermediate intrinsic reactivity between nitriles (low reactivity) and azanitriles (high reactivity), display the best biological outcome against the parasite. Although certain questions concerning the toxicity and metabolic stability of the aldehyde moiety remained unclear,⁵⁰ this compound class was selected in the optimization studies of different covalent inhibitors due to its excellent potency and selectivity profile^{12,15} with low cytotoxicity herein observed. Outstandingly, this warhead has been selected in the structure-based design optimization processes of covalent inhibitors of the SARS-CoV-2 main protease (M^{Pro}); the dipeptidyl aldehyde inhibitors displayed good pharmacokinetic properties and low toxicity *in vivo*, leading to promising drug candidates for the treatment of COVID-19.⁵¹ However, aldehydes remain relatively unpopular in drug discovery for their metabolic liability, as these chemical functionalities are prone to reduction by aldo-keto reductases, and oxidation to the corresponding acids by aldehyde dehydrogenases.⁵²

Overall, we can elect a bivalent behaviour of compound **Neq0656**. On the one side, based merely on the enzymatic activity and the intrinsic reactivity, the nitrile based-compound is more efficient as it shows less intrinsic reactivity while presenting roughly the same apparent potency. On the other side, the exchange of the nitrile warhead for an aldehyde upraises the trypanocidal activity/selectivity.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

The authors acknowledge Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (grant #2013/18009-4, #2018/15904-6 and #2018/21749-3) for financing this project, and

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for their financial support. JL also gives thanks for the access to the computational resources of the Supercomputer Santos Dumont (SDumont) provided by the Laboratório de Computação Científica (LNCC).

References

- 1 J. R. Coura and P. A. Viñas, *Nature*, 2010, **465**, S6–S7.
- 2 C. Bern, S. Kjos, M. J. Yabsley and S. P. Montgomery, *Clin. Microbiol. Rev.*, 2011, **24**, 655–681.
- 3 E. Chatelain, *J. Biomol. Screening*, 2015, **20**, 22–35.
- 4 Y. Jackson, A. Pinto and S. Pett, *Trop. Med. Int. Health*, 2014, **19**, 212–218.
- 5 P. C. Pereira and E. Navarro, *J. Venomous Anim. Toxins Incl. Trop. Dis.*, 2013, **19**, 34.
- 6 M. E. McGrath, A. E. Eakin, J. C. Engel, J. H. McKerrow, C. S. Craik and R. J. Fletterick, *J. Mol. Biol.*, 1995, **247**, 251–259.
- 7 L. A. A. Avelar, C. D. Camilo, S. de Albuquerque, W. B. Fernandes, C. Gonçalves, P. W. Kenny, A. Leitão, J. H. McKerrow, C. A. Montanari, E. V. M. Orozco, J. F. R. Ribeiro, J. R. Rocha, F. Rosini and M. E. Sidel, *PLoS Neglected Trop. Dis.*, 2015, **9**, e0003916.
- 8 A. C. B. Burtoloso, S. de Albuquerque, M. Furber, J. C. Gomes, C. Gonçalves, P. W. Kenny, A. Leitão, C. A. Montanari, J. C. Quilles, J. F. R. Ribeiro and J. R. Rocha, *PLoS Neglected Trop. Dis.*, 2017, **11**, e0005343.
- 9 L. Cianni, G. Sartori, F. Rosini, D. V. De, G. Pires, B. R. Lopes, A. Leitão, A. C. B. Burtoloso and C. A. Montanari, *Bioorg. Chem.*, 2018, **79**, 285–292.
- 10 J. C. Gomes, L. Cianni, J. Ribeiro, F. dos Reis Rocho, S. da Costa Martins Silva, P. H. J. Batista, C. B. Moraes, C. H. Franco, L. H. G. Freitas-Junior, P. W. Kenny, A. Leitão, A. C. B. Burtoloso, D. de Vita and C. A. Montanari, *Bioorg. Med. Chem.*, 2019, **27**, 115083.
- 11 D. G. Silva, J. F. R. Ribeiro, D. De Vita, L. Cianni, C. H. Franco, L. H. Freitas-Junior, C. B. Moraes, J. R. Rocha, A. C. B. Burtoloso, P. W. Kenny, A. Leitão and C. A. Montanari, *Bioorg. Med. Chem. Lett.*, 2017, **27**, 5031–5035.
- 12 L. Cianni, C. W. Feldmann, E. Gilberg, M. Gütschow, L. Juliano, A. Leitão, J. Bajorath and C. A. Montanari, *J. Med. Chem.*, 2019, **62**, 10497–10525.
- 13 K777 (Chagas)|DNDi, <https://www.dndi.org/diseases-projects/portfolio/completed-projects/k777/>, (accessed 17 July 2019).
- 14 B. D. Jones, A. Tochowicz, Y. Tang, M. D. Cameron, L.-I. McCall, K. Hirata, J. L. Siqueira-Neto, S. L. Reed, J. H. McKerrow and W. R. Roush, *ACS Med. Chem. Lett.*, 2016, **7**, 77–82.
- 15 M. Gehringer and S. A. Laufer, *J. Med. Chem.*, 2019, **62**, 5673–5724.
- 16 R. Lonsdale, J. Burgess, N. Colclough, N. L. Davies, E. M. Lenz, A. L. Orton and R. A. Ward, *J. Chem. Inf. Model.*, 2017, **57**, 3124–3137.
- 17 P. A. MacFaul, A. D. Morley and J. J. Crawford, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 1136–1138.

- 18 P. Ábrányi-Balogh, L. Petri, T. Imre, P. Szijj, A. Scarpino, M. Hrast, A. Mitrović, U. P. Fonovič, K. Németh, H. Barreteau, D. I. Roper, K. Horváti, G. G. Ferenczy, J. Kos, J. Ilaš, S. Gobec and G. M. Keserű, *Eur. J. Med. Chem.*, 2018, **160**, 94–107.
- 19 J. S. Martin, C. J. MacKenzie, D. Fletcher and I. H. Gilbert, *Bioorg. Med. Chem.*, 2019, **27**, 2066–2074.
- 20 L. M. McGregor, M. L. Jenkins, C. Kerwin, J. E. Burke and K. M. Shokat, *Biochemistry*, 2017, **56**, 3178–3183.
- 21 K. Sayin and D. Karakaş, *Spectrochim. Acta, Part A*, 2018, **188**, 537–546.
- 22 M. Gonzalez-Suarez, A. Aizman, J. Soto-Delgado and R. Contreras, *J. Org. Chem.*, 2012, **77**, 90–95.
- 23 F. Palazzesi, M. A. Grundl, A. Pautsch, A. Weber and C. S. Tautermann, *J. Chem. Inf. Model.*, 2019, **59**, 3565–3571.
- 24 V. Ehmke, J. E. Q. Quinsaas, P. Rivera-Fuentes, C. Heindl, C. Freymond, M. Rottmann, R. Brun, T. Schirmeister and F. Diederich, *Org. Biomol. Chem.*, 2012, **10**, 5764.
- 25 R. M. Oballa, J.-F. Truchon, C. I. Bayly, N. Chauvet, S. Day, S. Crane and C. Berthelette, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 998–1002.
- 26 A. Cherkasov, E. N. Muratov, D. Fourches, A. Varnek, I. I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y. C. Martin, R. Todeschini, V. Consonni, V. E. Kuz'min, R. Cramer, R. Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard and A. Tropsha, *J. Med. Chem.*, 2014, **57**, 4977–5010.
- 27 S. Mondal Roy, *Comput. Biol. Chem.*, 2018, **75**, 91–100.
- 28 R. Parthasarathi, V. Subramanian, D. R. Roy and P. K. Chattaraj, *Bioorg. Med. Chem.*, 2004, **12**, 5533–5543.
- 29 A. P. C. A. Lima, F. C. G. dos Reis, C. Serveau, G. Lalmanach, L. Juliano, R. Ménard, T. Vernet, D. Y. Thomas, A. C. Storer and J. Scharfstein, *Mol. Biochem. Parasitol.*, 2001, **114**, 41–52.
- 30 B. Zingales, M. A. Miles, C. B. Moraes, A. Luquetti, F. Guhl, A. G. Schijman and I. Ribeiro, *Mem. Inst. Oswaldo Cruz*, 2014, **109**, 828–833.
- 31 H. O. Rodriguez, N. A. Guerrero, A. Fortes, J. Santi-Rocca, N. Gironès and M. Fresno, *Acta Trop.*, 2014, **139**, 57–66.
- 32 J. R. A. Silva, L. Cianni, D. Araujo, P. H. J. Batista, D. de Vita, F. Rosini, A. Leitão, J. Lameira and C. A. Montanari, *J. Chem. Inf. Model.*, 2020, **60**, 1666–1677.
- 33 A. Meister, *J. Biol. Chem.*, 1988, **263**, 17205–17208.
- 34 A. M. Dos Santos, L. Cianni, D. De Vita, F. Rosini, A. Leitão, C. A. Laughton, J. Lameira and C. A. Montanari, *Phys. Chem. Chem. Phys.*, 2018, **20**, 24317–24328.
- 35 M. Gaus, Q. Cui and M. Elstner, *J. Chem. Theory Comput.*, 2011, **7**, 931–948.
- 36 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- 37 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.
- 38 M. J. Frisch, J. A. Pople and J. S. Binkley, *J. Chem. Phys.*, 1984, **80**, 3265–3269.
- 39 Chr. Möller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.
- 40 S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117–129.
- 41 P. K. Chattaraj, U. Sarkar and D. R. Roy, *Chem. Rev.*, 2006, **106**, 2065–2091.
- 42 P. W. Ayers and R. G. Parr, *J. Am. Chem. Soc.*, 2000, **122**, 2010–2018.
- 43 C. Lee, W. Yang and R. G. Parr, *J. Mol. Struct.: THEOCHEM*, 1988, **163**, 305–313.
- 44 S. Liu and R. G. Parr, *J. Chem. Phys.*, 1997, **106**, 5578–5586.
- 45 P. Pérez, *J. Org. Chem.*, 2003, **68**, 5886–5889.
- 46 W. Yang and W. J. Mortier, *J. Am. Chem. Soc.*, 1986, **108**, 5708–5711.
- 47 A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899–926.
- 48 U. Wriede, M. Fernandez, K. F. West, D. Harcour and H. W. Moore, *J. Org. Chem.*, 1987, **52**, 4485–4489.
- 49 R. Löser, M. Frizler, K. Schilling and M. Gütschow, *Angew. Chem., Int. Ed.*, 2008, **47**, 4331–4334.
- 50 R. M. LoPachin and T. Gavin, *Chem. Res. Toxicol.*, 2014, **27**, 1081–1091.
- 51 W. Dai, B. Zhang, X.-M. Jiang, H. Su, J. Li, Y. Zhao, X. Xie, Z. Jin, J. Peng, F. Liu, C. Li, Y. Li, F. Bai, H. Wang, X. Cheng, X. Cen, S. Hu, X. Yang, J. Wang, X. Liu, G. Xiao, H. Jiang, Z. Rao, L.-K. Zhang, Y. Xu, H. Yang and H. Liu, *Science*, 2020, **368**, 1331–1335.
- 52 P. Malátková and V. Wsól, *Drug Metab. Rev.*, 2014, **46**, 96–123.

5. QM-cluster Approach to Model the Warhead of Reversible Covalent Inhibitors

As shown in the previous chapter, QM methods and descriptors have proven to be effective in linking the intrinsic reactivity of the warhead to the experimentally observed potency. This approach can be valuable in developing QSAR models or employing machine learning techniques, where specific descriptors are calculated and utilized to establish correlations with affinity or other relevant properties. However, it is essential to note that in this context, the focus is solely on the ligand structure, similar to LBDD methods, without considering the influence of the macromolecular target.

Another application of QM methodology involves the use of simplified systems to assess the contribution of chemical bond formation (i.e., the importance of the warhead) to the binding free energy of the complex. In this approach, the catalytic residue(s) of the protein are considered in a simplified manner. Diederich and co-workers have explored this approach by designing a reaction model where ligands with similar warheads react with SHCH_3 as the reagents.¹ The energy of the resulting adduct is then calculated to represent the products, providing insight into the reaction energy. This approach, known as the electrophilicity index, offers an exciting and simplified alternative for such calculations.

Prof. Warshel has recently conducted similar studies, as mentioned in the introduction.² In his approach, the reagents were defined as the ligand, the catalytic residue (e.g., Cys), and a few amino acids involved in hydrogen bonding with the ligand. The product, on the other hand, included the same residues and the resulting covalent adduct. This simplified methodology provides a more accessible way to estimate the reaction energy, incorporating a limited representation of the protein microenvironment by including additional protein residues in the reaction model. However, these simplified models fail to capture the anisotropic effects of the binding site environment.

Although it is possible to use more robust QM methodologies, the main limitation of these simplifications is that proteins of the same family with identical or similar residues may have the same energy values. Therefore, these simplifications restrict the applicability of the methodology when selectivity-related aspects are desired.

An effective alternative is to employ multiscale calculations, where the active site, protein, and explicit solvent are represented, allowing for the accurate capture of the energy involved in the chemical reaction.³ However, if the use of QM/MM is not preferable due to computational expenses or the challenges in system preparation and obtaining a feasible reaction pathway, there is another promising alternative that relies solely on QM methodology.

This methodology is the QM-cluster approach, which can be valuable in obtaining the free energy related to the formation of the covalent adduct and can overcome the problems mentioned earlier for other methods.^{4,5} This method involves constructing limited models of the protein binding site, typically consisting of 300 to 400 atoms.^{5,6} By utilizing powerful supercomputers, it becomes possible to employ relatively accurate QM methods for studying the system.

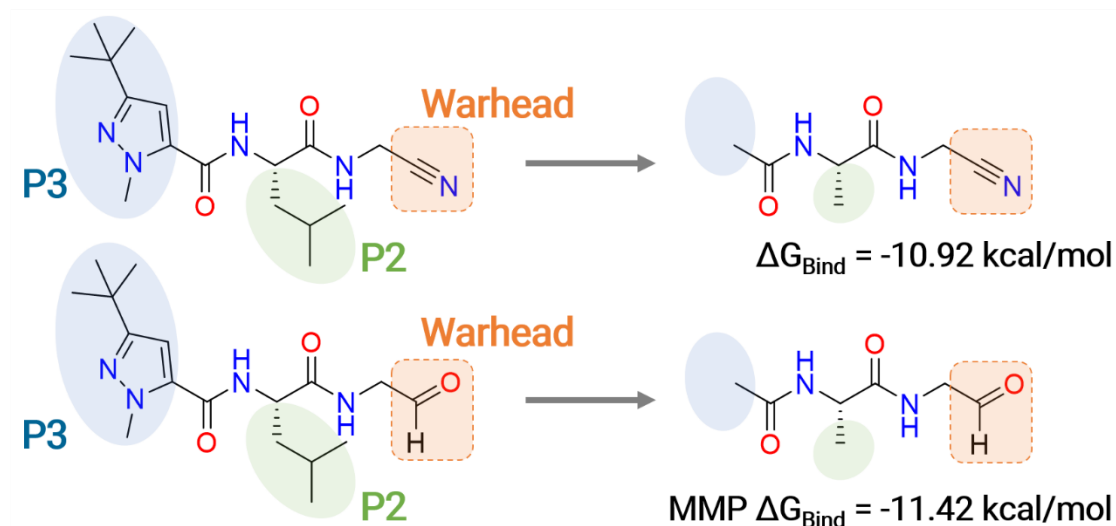
In order to create the cluster model, certain constraints need to be applied to account for the steric influence imposed by the enzyme matrix on the active site. This involves fixing specific atoms in the active site model to preserve this influence.^{4,5,7} Typically, alpha carbons of amino acids and selected hydrogen atoms are held fixed. Furthermore, implicit solvation models are employed to represent the remaining protein region that was removed from the active site model.^{4,5,7} This approach utilizes a dielectric constant value to mimic a more nonpolar environment, consistent with the protein surroundings. Typically, a dielectric constant value of 4 is chosen for this purpose.

Nevertheless, it is important to note that excessively strict constraints should be avoided when constructing the cluster model.^{4,5,7} A highly rigid model may lead to inaccurate estimates of the free energy. Additionally, the approximation tends to improve as the size of the model increases, allowing for greater flexibility among the groups.^{4,5,7} Therefore, the inclusion of more atoms in the system enhances the accuracy of the calculations.

The same principle applies to the treatment of electrostatic influence using the dielectric constant. While the value $\epsilon=4$ is commonly employed in enzymatic catalysis studies, it is important to note that as the model size increases, the impact of this influence tends to diminish.^{4,5,7} In fact, it has been observed that when the model reaches approximately 150-200 atoms, the solvation effect saturates and becomes almost negligible.

In this study, we aim to explore a novel approach by applying this methodology, which, to our knowledge, has not been previously utilized in reversible covalent inhibitor development. Specifically, we will focus on the molecular modeling of different warheads to investigate their impact on inhibitor potency. Two ligands will be examined, one featuring a nitrile-based warhead and the other an aldehyde-based warhead (Figure 1). In order to simplify the system, all substituent groups (P2 and P3) were removed from both ligands, as they did not possess a P1 group (only CH₂). Since we expect this effect to be nullified due to the identical groups in both ligands, removing the substituent groups simplifies the system by reducing the total number of atoms. Thus, only the warhead and the dipeptide scaffold of the inhibitors are considered (Figure 1).

Figure 1 – The nitrile- and aldehyde-based reversible covalent inhibitors used to explore the potential use of the QM-cluster approach in TCIs modeling.



Source: Image created by the author.

The nitrile-based compound is a known inhibitor of CatL with a pK_i value of 8.1 ($\Delta G_{\text{Bind}} = -10.92 \text{ kcal/mol}$). On the other hand, the aldehyde-based inhibitor is a novel compound that has not yet been tested against CatL, although its synthesis has already been performed. However, based on internal data from our research group at Nequimed/IQSC/USP, utilizing matched molecular pair (MMP) analysis with other inhibitors of CatL where only the Nitrile \rightarrow Aldehyde warhead is modified, it is expected that the aldehyde-containing inhibitor will be

approximately 0.5 kcal/mol more stable compared to the nitrile-based inhibitor ($\Delta G_{\text{Bind}} = -11.42$ kcal/mol).

As the only distinction between the inhibitors lies in the warhead, we attribute this discrepancy in potency to the warhead modification. To capture this variation accurately, it is crucial to employ quantum methods to calculate the free energy difference between these two compounds. Therefore, the objective is to test the applicability of the QM-cluster approach method to predict energy differences in the substitution of warheads in pair ligands targeting a specific protein.

The binding site model was created using the enzyme from PDB code 3HHA,⁸ and the ligands were manually placed within the active site for both the covalent and non-covalent states. Subsequently, 100 ns molecular dynamics simulations were performed to equilibrate the systems. The simulations were performed using the Amber20⁹ program, utilizing the ff19SB force field for the protein and the OPC model¹⁰ for water molecules. The ligand charges were obtained using the RESP approach,¹¹ and parameterization of the ligand was performed with the GAFF2 force field.¹² Following the 100 ns simulations, the final frame was selected for model generation, removing water molecules, counterions, and residues more than 8 Å away from the catalytic Cys sulfur atom in order to focus on the active site region.

Therefore, the active site models comprised 309 and 310 atoms for the nitrile-based and aldehyde-based inhibitor systems, respectively. Two models were created for each inhibitor, representing the covalent and non-covalent states. This enabled the estimation of the binding free energy difference associated with the formation of the covalent adduct ($E_{\text{covalent}} - E_{\text{noncovalent}}$).

It is worth noting that for the non-covalent state, the ionic pair of the catalytic residues (Cys⁻/His⁺) was considered, as commonly used in QM/MM calculations to describe the reaction pathways of covalent inhibitors of cysteine proteases.¹³⁻¹⁷ Since the reaction involves the nucleophilic attack of the Cys thiolate, we considered this configuration as the non-covalent state, representing the reactants of the chemical reaction. Thus, it is possible to estimate the contribution of the covalent bond formation.

The residues presented in the catalytic site model are: Gln19, Gly23, Ser24, Cys25 (catalytic), Trp26, Ala27, Ser29, Gly67, Gly68, Leu69, Met70,

Ile136, Ala138, Phe143, Met161, Asp162, His163 (catalytic), Asn187, Ser188, Trp189 e Trp193. In the end, 96 atoms were frozen, mainly the alpha carbons of most of the residues, but the catalytic ones. Also, several hydrogens atoms were kept fixed in order to speed up the calculations, making the system more rigid. However, not so rigid as to lead to erroneous energy values.

As the binding site clusters models were created for both the non-covalent and covalent states for each of the inhibitors, the next step is to perform quantum calculations using Gaussian 16.¹⁸ The choice of the methodology is a crucial factor for calculations that take advantage of the use of QM-cluster. In this context, we employed the B3LYP¹⁹ level of theory with empirical dispersion correction, B3LYP-D3(BJ),^{20,21} which offers a favorable balance between accuracy and computational efficiency.⁶ For the basis set, the 6-31G(d,p) was chosen for the optimization of the complexes, which will be referred to as BS1. This choice of level of theory has been widely used in other QM-clusters works to obtain reliable optimized structures in the gas phase with good values of energy.

After optimizing the system, it becomes necessary to obtain accurate energy values. To improve the accuracy, the first step is to include an implicit solvent model using the SMD approach with a dielectric constant of 4, considering the BS1 basis set for single-point (SP) calculations. Moreover, with the BS1, the zero-point energy (ZPE) was obtained through frequency calculations to be included as a correction to the total energy. Finally, larger basis sets are required to achieve better energy values. Therefore, SP calculations were performed with the basis set changed to 6-311+G(2d,2p), denoted as BS2.

Thus, for each complex, the energy is obtained using the following equation:

$$E_{total} = E_{BS2} + ZPE_{BS1} + (E_{solv,BS1} - E_{BS1}) \quad (1)$$

BS1: B3LYP-D3(BJ) | 6-31G(d,p)

solv, BS1: SP in SMD | $\epsilon = 4$ | BS1

BS2: SP B3LYP-D3(BJ) | 6-311+G(2d,2p)

ZPE: Freq calculation | BS1

With the E_{total} for the non-covalent and covalent states, it is possible to obtain the reaction energy for the formation of the covalent adduct ($E_{\text{cov}} - E_{\text{noncov}}$) for both inhibitors. This allows for the comparison of the energetic difference between compounds with different warheads.

The cartesian coordinates of each system optimized with BS1 can be found in the Appendix B section. The values obtained for each system are presented in Table 1 for nitrile- and aldehyde-based inhibitors and in both the non-covalent and covalent states.

Table 1 – Calculated absolute energies and energy corrections for the ligands in each state (non-covalent and covalent). The relative difference between both ligands are also shown. All the values are in hartree, but the ΔE_{total} for both inhibitors and the relative difference energy are in kcal.mol⁻¹.

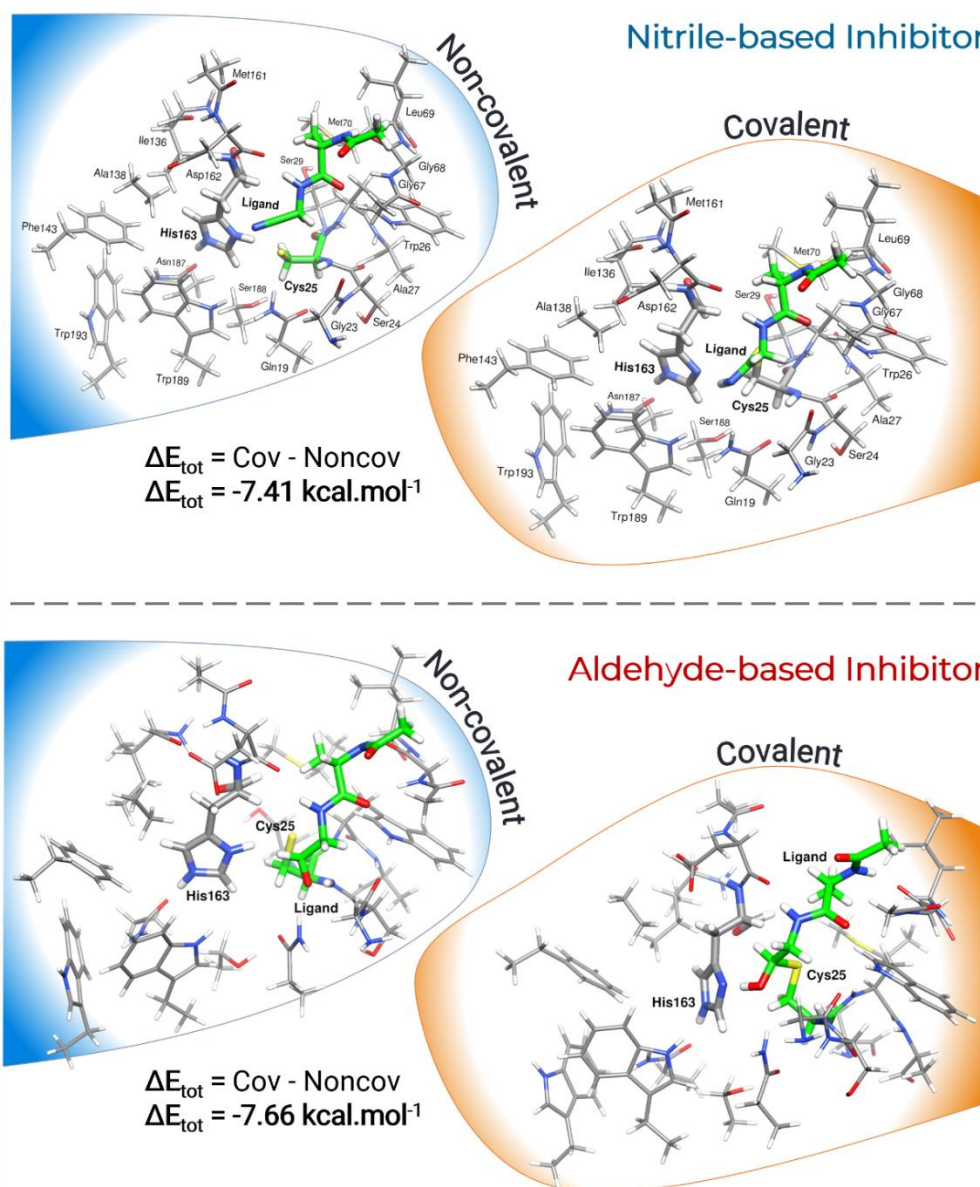
		E_{BS2}	ZPE_{BS1}	$E_{\text{solv,BS1}}$	E_{BS1}	E_{total}
NITRILE	Non-covalent	-7588.73	2.09	-7586.82	-7586.72	-7586.746
	Covalent	-7588.75	2.09	-7586.84	-7586.74	-7586.758
	$\Delta E_{\text{total_Nitrile}}$					-7.41 kcal.mol ⁻¹
ALDEHYDE	Non-covalent	-7609.83	2.10	-7607.92	-7607.81	-7607.83
	Covalent	-7609.84	2.10	-7607.93	-7607.83	-7607.85
	$\Delta E_{\text{total_Aldehyde}}$					-7.66 kcal.mol ⁻¹
RELATIVE	$\Delta \Delta G_{\text{Nitrile - Aldehyde}}$					0.25 kcal.mol ⁻¹

Source: Table created by the author.

Hence, it can be observed that the energy values for both the nitrile and aldehyde are pretty similar. However, the aldehyde-based inhibitor system is slightly more stable than the system containing nitrile. Therefore, the reaction free energy for the aldehyde-based system is 0.25 kcal/mol lower than that of the nitrile-based system. This also indicates that the aldehyde is more reactive than the nitrile and achieves better stabilization after the formation of the adduct. Compared to experimental data, the trend is maintained, with the aldehyde being

more stable than the nitrile, and the expected experimental difference of 0.5 kcal/mol is almost reached, with an error of only 0.25 kcal/mol. The representations of the systems in both states are shown in Figure 2.

Figure 2 – Optimized structures of cluster models for non-covalent and covalent states for nitrile- (309 atoms) and aldehyde-based (310 atoms) inhibitors. The ligands are highlighted with green carbons.



Source: Image created by the author.

Recently, Lameira and colleagues obtained the reaction energy value for a nitrile-based inhibitor against CatK using QM/MM calculations.¹⁶ The values obtained by them were -3.5 kcal/mol and -11.8 kcal/mol using the M06-2X/6-31++G(d,p)/MM and B3LYP-D3/6-31+G(d)/MM potentials, respectively. In

comparison, our nitrile-based inhibitor showed a reaction energy value of -7.41 kcal/mol, which falls within an acceptable range and is consistent with the values obtained in the literature using QM/MM methodologies, which can be considered the gold standard for comparison. Of course, both the enzyme and the inhibitor are different. However, the enzyme in Lameira's work is CatK,¹⁶ from the same family as CatL, where an RMSD of only 0.5 Å is observed near the active site residues.²² Also, the nucleophilicity of the catalytic Cys is likely to be similar in these two proteins. Moreover, the intrinsic reactivity of nitrile-based inhibitors is expected to be comparable, with minimal influence from substituent groups. Still, both nitrile- and aldehyde-based inhibitors used in this work exhibited reaction energies below 23 kcal/mol, which aligns with their reversible inhibitor nature.^{23,24}

In this way, the results obtained with the QM-cluster approach are consistent with expectations and comparable to QM/MM for a very similar system. We obtained the reaction free energy for two paired inhibitors, changing only the warhead (nitrile and aldehyde), and it appears that we captured the difference in potency between these inhibitors. For this purpose, we considered only a simplified portion of the enzyme's active site and the inhibitor. To our knowledge, this is the first time that the QM-cluster approach has been used for covalent ligands.

The advantage of the QM-cluster approach lies in its ability to consider the protein environment that directly participates in enzymatic catalysis/inhibition, thus overcoming the limitation of using only the catalytic residue or a few amino acids that interact with the inhibitor, neglecting the anisotropic nature of the binding site.^{4-6,25} Additionally, the QM-cluster method offers more straightforward system preparation and faster results compared to the gold standard for studying enzyme reactions computationally, QM/MM.

Therefore, for the first time, it has been demonstrated that obtaining the free energy of reaction for reversible covalent inhibitors is possible using this methodology, with great potential for application in future drug discovery projects focused on covalent inhibitors. Furthermore, the QM-cluster approach can also be helpful in the molecular modeling of irreversible inhibitors, allowing for the determination of the inhibitor's transition state in the active site environment and the associated energy. Thus, this methodology can be highly valuable in designing new reversible and irreversible covalent inhibitors.

Finally, by obtaining the free energy of the reaction, it is also possible to combine these results with those obtained from other methodologies, such as FEP, as mentioned in the INTRODUCTION. To illustrate this potential combination of methodologies, Himo used a combination of QM-cluster and FEP results to explain the enzymatic catalysis mechanisms of an acyl transferase.²⁶ Herein, with the determination of the reaction energy using the QM-cluster approach, the energy of ΔG_{dc}^{core} is obtained.²⁷ Therefore, with appropriate FEP calculations, it is possible to obtain the ABFE for the reversible covalent inhibitor, as describe in equation 5 in the INTRODUCTION chapter, and this can be an outstanding contribution to the field.

REFERENCES

- 1 EHMKE, V.; QUINSAAT, J. E. Q.; RIVERA-FUENTES, P.; HEINDL, C.; FREYMOND, C.; ROTTMANN, M.; BRUN, R.; SCHIRMEISTER, T.; DIEDERICH, F. Tuning and predicting biological affinity: aryl nitriles as cysteine protease inhibitors. **Organic & Biomolecular Chemistry**, Cambridge, v. 10, n. 30, p. 5764–5768, 2012.
- 2 ZHOU, J.; SAHA, A.; HUANG, Z.; WARSHEL, A. Fast and effective prediction of the absolute binding free energies of covalent inhibitors of SARS-CoV-2 main protease and 20S proteasome. **Journal of the American Chemical Society**, Washington, v. 144, n. 17, p. 7568–7572, 2022.
- 3 GRIGORENKO, B. L.; POLYAKOV, I. V.; KHRENOVA, M. G.; GIUDETTI, G.; FARAJI, S.; KRYLOV, A. I.; NEMUKHIN, A. V. Multiscale simulations of the covalent inhibition of the SARS-CoV-2 main protease: Four compounds and three reaction mechanisms. **Journal of the American Chemical Society**, Washington, v. 145, n. 24, p. 13204–13214, 2023.
- 4 SIEGBAHN, P. E. M.; HIMO, F. The quantum chemical cluster approach for modeling enzyme reactions. **Wiley Interdisciplinary Reviews: computational molecular science**, Hoboken, v. 1, n. 3, p. 323–336, 2011.
- 5 SHENG, X.; HIMO, F. The quantum chemical cluster approach in biocatalysis. **Accounts of Chemical Research**, Washington, v. 56, n. 8, p. 938–947, 2023.
- 6 SHENG, X.; KAZEMI, M.; PLANAS, F.; HIMO, F. Modeling enzymatic enantioselectivity using quantum chemical methodology. **ACS Catalysis**, Washington, v. 10, n. 11, p. 6430–6449, 2020.
- 7 SIEGBAHN, P. E. M. A quantum chemical approach for the mechanisms of redox-active metalloenzymes. **RSC Advances**, Cambridge, v. 11, n. 6, p. 3495–3508, 2021
- 8 ASAAD, N.; BETHEL, P. A.; COULSON, M. D.; DAWSON, J. E.; FORD, S. J.; GERHARDT, 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, 2009.
- 9 CASE, D. A.; BELFON, K.; BEN-SHALOM, I. Y.; BROZELL, S. R.; CERUTTI, D. S.; T.E. CHEATHAM, I.; CRUZEIRO, V. W. D.; DARDEN, T. A.; DUKE, R. E.; GIAMBASU, G.; GILSON, M. K.; GOHLKE, H.; GOETZ, A. W.; HARRIS, R.; IZADI, S.; IZMAILOV, S. A.; KASAVAJHALA, K.; KOVALENKO, A.; KRASNY, R.; KURTZMAN, T.; LEE, T. S.; LEGRAND, S.; P. LI, C. L.; LIU, J.; LUCHKO, T.; LUO, R.; MAN, V.; MERZ, K. M.; MIAO, Y.; MIKHAILOVSKII, O.; MONARD, G.; NGUYEN, H.; ONUFRIEV, A.; F.PAN; PANTANO, S.; QI, R.; ROE, D. R.; ROITBERG, A.; SAGUI, C.; SCHOTT-VERDUGO, S.; SHEN, J.; SIMMERLING, C. L.; N.R.SKRYNNIKOV; SMITH, J.; SWAILS, J.; WALKER, R. C.; WANG, J.; WILSON, L.; WOLF, R. M.; WU, X.; XIONG, Y.; XUE, Y.; YORK, D. M.; KOLLMAN, P. A. **AMBER 2020**. Version 2020. San Francisco: University of California, 2020.
- 10 IZADI, S.; ANANDAKRISHNAN, R.; ONUFRIEV, A. V. Building water models: a different approach. **The Journal of Physical Chemistry Letters**, Washington, v. 5, n. 21, p. 3863–3871, 2014.

- 11 BAYLY, C. I.; CIEPLAK, P.; CORNELL, W.; KOLLMAN, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. **The Journal of Physical Chemistry**, Washington, v. 97, n. 40, p. 10269–10280, 1993.
- 12 WANG, J.; WOLF, R. M.; CALDWELL, J. W.; KOLLMAN, P. A.; CASE, D. A. Development and testing of a general amber force field. **Journal of Computational Chemistry**, Hoboken, v. 25, n. 9, p. 1157–1174, 2004.
- 13 MARTÍ, S.; ARAFET, K.; LODOLA, A.; MULHOLLAND, A. J.; ŚWIDEREK, K.; MOLINER, V. Impact of warhead modulations on the covalent inhibition of SARS-CoV-2 M^{pro} explored by QM/MM simulations. **ACS Catalysis**, Washington, v. 12, n. 1, p. 698–708, 2022.
- 14 ARAFET, K.; SERRANO-APARICIO, N.; LODOLA, A.; MULHOLLAND, A. J.; GONZÁLEZ, F. V.; ŚWIDEREK, K.; MOLINER, V. Mechanism of inhibition of SARS-CoV-2 M^{pro} by N3 peptidyl Michael acceptor explained by QM/MM simulations and design of new derivatives with tunable chemical reactivity. **Chemical Science**, Cambridge, v. 12, n. 4, p. 1433–1444, 2021.
- 15 CHAN, H. T. H.; MOESSER, M. A.; WALTERS, R. K.; MALLA, T. R.; TWIDALE, R. M.; JOHN, T.; DEEKS, H. M.; JOHNSTON-WOOD, T.; MIKHAILOV, V.; SESSIONS, R. B.; DAWSON, W.; SALAH, E.; LUKACIK, P.; STRAIN-DAMERELL, C.; OWEN, C. D.; NAKAJIMA, T.; ŚWIDEREK, K.; LODOLA, A.; MOLINER, V.; GLOWACKI, D. R.; SPENCER, J.; WALSH, M. A.; SCHOFIELD, C. J.; GENOVESE, L.; SHOEMARK, D. K.; MULHOLLAND, A. J.; DUARTE, F.; MORRIS, G. M. Discovery of SARS-CoV-2 M^{pro} peptide inhibitors from modelling substrate and ligand binding. **Chemical Science**, Cambridge, v. 12, n. 41, p. 13686–13703, 2021.
- 16 DOS SANTOS, A. M.; OLIVEIRA, A. R. S.; DA COSTA, C. H. S.; KENNY, P. W.; MONTANARI, C. A.; VARELA, J. de J. G.; LAMEIRA, J. Assessment of reversibility for covalent cysteine protease inhibitors using quantum mechanics/molecular mechanics free energy surfaces. **Journal of Chemical Information and Modeling**, Washington, v. 62, n. 17, p. 4083–4094, 2022.
- 17 DA COSTA, C.; BONATTO, V.; DOS SANTOS, A.; LAMEIRA, J.; LEITAO, A.; MONTANARI, C. Evaluating QM/MM free energy surfaces for ranking cysteine protease covalent inhibitors. **Journal of Chemical Information and Modeling**, Washington, v. 60, n. 2, p. 880–889, 2020.
- 18 FRISCH, M. J.; TRUCKS, G. W.; SCHLEGEL, H. B.; SCUSERIA, G. E.; ROBB, M. A.; CHEESEMAN, J. R.; SCALMANI, G.; BARONE, V.; PETERSSON, G. A.; NAKATSUJI, H.; LI, X.; CARICATO, M.; MARENICH, A. V.; BLOINO, J.; JANESKO, B. G.; GOMPERTS, R.; MENNUCCI, B.; HRATCHIAN, H. P.; ORTIZ, J. V.; IZMAYLOV, A. F.; SONNENBERG, J. L.; WILLIAMS-YOUNG, D.; DING, F.; LIPPARINI, F.; EGIDI, F.; GOINGS, J.; PENG, B.; PETRONE, A.; HENDERSON, T.; RANASINGHE, D.; ZAKRZEWSKI, V. G.; GAO, J.; REGA, N.; ZHENG, G.; LIANG, W.; HADA, M.; EHARA, M.; TOYOTA, K.; FUKUDA, R.; HASEGAWA, J.; ISHIDA, M.; NAKAJIMA, T.; HONDA, Y.; KITAO, O.; NAKAI, H.; VREVEN, T.; THROSSELL, K.; MONTGOMERY JR., J. A.; PERALTA, J. E.; OGLIARO, F.; BEARPARK, M. J.; HEYD, J. J.; BROTHERS, E. N.; KUDIN, K. N.; STAROVEROV, V. N.; KEITH, T. A.; KOBAYASHI, R.; NORMAND, J.; RAGHAVACHARI, K.; RENDELL, A. P.; BURANT, J. C.; IYENGAR, S. S.; TOMASI, J.; COSSI, M.; MILLAM, J. M.; KLENE, M.; ADAMO, C.; CAMMI, R.; OCHTERSKI, J. W.; MARTIN, R. L.; MOROKUMA, K.; FARKAS, O.; FORESMAN, J. B.; FOX, D. J. **Gaussian 16**. Version 16. Wallingford: Gaussian Inc., 2016.

- 19 BECKE, A. D. Density-functional thermochemistry. III. the role of exact exchange. **The Journal of Chemical Physics**, Melville, v. 98, p. 5648–5652, 1993.
- 20 BURSCH, M.; CALDEWEYHER, E.; HANSEN, A.; NEUGEBAUER, H.; EHLERT, S.; GRIMME, S. Understanding and quantifying London dispersion effects in organometallic complexes. **Accounts of Chemical Research**, Washington, v. 52, n. 1, p. 258–266, 2019.
- 21 GRIMME, S. Density functional theory with London dispersion corrections. **Wiley Interdisciplinary Reviews: computational molecular science**, Hoboken, v. 1, n. 2, p. 211–228, 2011.
- 22 CIANNI, L.; FELDMANN, C. W.; GILBERG, E.; GÜTSCHOW, M.; JULIANO, L.; LEITÃO, A.; BAJORATH, J.; MONTANARI, C. A. Can cysteine protease cross-class inhibitors achieve selectivity? **Journal of Medicinal Chemistry**, Washington, v. 62, n. 23, p. 10497–10525, 2019.
- 23 TUMMINO, P. J.; COPELAND, R. A. Residence time of receptor–ligand complexes and its effect on biological function. **Biochemistry**, Washington, v. 47, n. 20, p. 5481–5492, 2008.
- 24 MONDAL, D.; WARSHEL, A. Exploring the mechanism of covalent inhibition: simulating the binding free energy of α -ketoamide inhibitors of the main protease of SARS-CoV-2. **Biochemistry**, Washington, v. 59, n. 48, p. 4601–4608, 2020.
- 25 PREJANÒ, M.; ŠKERLOVÁ, J.; STENMARK, P.; HIMO, F. Reaction mechanism of human PAICS elucidated by quantum chemical calculations. **Journal of the American Chemical Society**, Washington, v. 144, n. 31, p. 14258–14268, 2022.
- 26 KAZEMI, M.; SHENG, X.; KROUTIL, W.; HIMO, F. computational study of *Mycobacterium smegmatis* acyl transferase reaction mechanism and specificity. **ACS Catalysis**, Washington, v. 8, n. 11, p. 10698–10706, 2018.
- 27 CHATTERJEE, P.; BOTELLO-SMITH, W. M.; ZHANG, H.; QIAN, L.; ALSAMARAH, A.; KENT, D.; LACROIX, J. J.; BAUDRY, M.; LUO, Y. Can relative binding free energy predict selectivity of reversible covalent inhibitors? **Journal of the American Chemical Society**, Washington, v. 139, n. 49, p. 17945–17952, 2017.

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

Reprinted with permission from BONATTO, V.; SHAMIM, A.; ROCHO, F. R.; LEITÃO, A.; LUQUE, F. J.; LAMEIRA, J.; MONTANARI, C. A. Predicting the Relative Binding Affinity for Reversible Covalent Inhibitors by Free Energy Perturbation Calculations. **Journal of Chemical Information and Modeling**, Washington, v. 61, p. 4733-4744, 2021. Copyright 2023 American Chemical Society.

Cite this: <https://pubs.acs.org/doi/abs/10.1021/acs.jcim.1c00515>

The supporting information for this work can be found in Appendix C section

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

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

Cite This: *J. Chem. Inf. Model.* 2021, 61, 4733–4744

Read Online

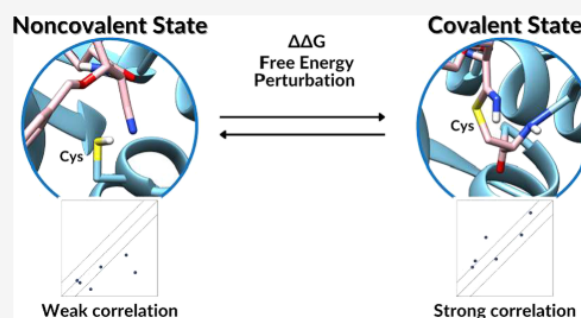
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

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



INTRODUCTION

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

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

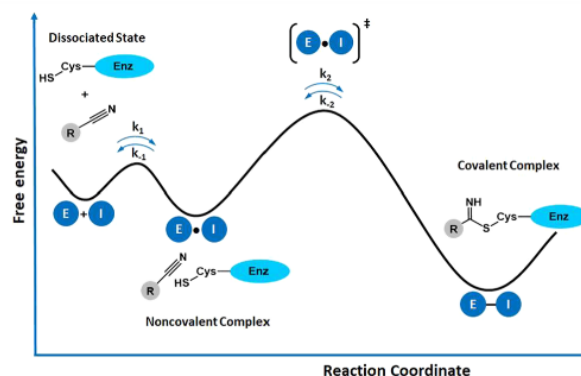


Figure 1. Schematic reaction involving a reversible covalent inhibitor (I) that incorporates a nitrile group as warhead and a reactive cysteine in the target enzyme (E) responsible for the nucleophilic attack. The rate constant k_{-2} determines the reversibility of the covalently bound adduct [E–I].

Received: May 7, 2021

Published: August 30, 2021



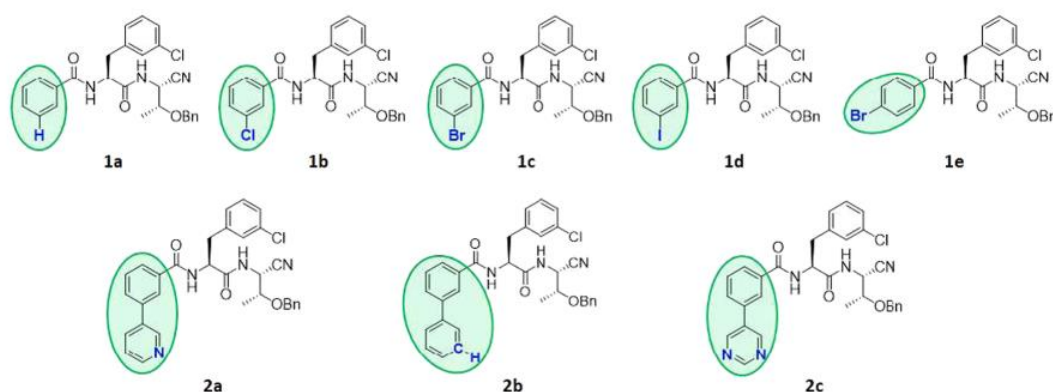


Figure 2. Two-dimensional (2D) structural representation of hCatL reversible covalent inhibitors. The P3 group is highlighted in green and the distinct R groups within each set of compounds are highlighted in blue.

complex between the inhibitor and the enzyme is subsequently formed due to the nucleophilic attack of the reactive cysteine. The step of adduct formation will guide the reversibility of the reaction through the rate constant k_{-2} . Therefore, the choice of the warhead is vital for the development of covalent inhibitors.¹³

The analysis of Figure 1 indicates that the energetic features involved in forming both noncovalent and covalent complexes are essential to gain insights into the mechanism that underlies the binding of reversible covalent inhibitors.¹³ Since the two states contribute to the overall free energy of the binding process,^{13,14} the tailored design of the chemical features required in the chemical structure of the inhibitor is a challenging task. In this context, free energy perturbation (FEP) is a powerful computational technique to explore the effects of chemical changes in the molecular structure on the binding affinity.¹⁵ This goal can be accomplished through appropriate thermodynamic cycles that enable the usage of FEP calculations to estimate absolute (ΔG_{Bind}) or relative ($\Delta\Delta G_{\text{Bind}}$) binding affinities.^{17,18} Several studies have used FEP calculations to estimate the binding free energy of inhibitors that do not form a chemical bond with the enzyme of interest.^{18–22} In the context of the current pandemic scenario, FEP has made an outstanding contribution to test potential drug candidates or finding new noncovalent compounds to inhibit Mpro of the SARS-CoV-2 virus.^{23,24}

For the specific case of reversible covalent inhibitors, Chatterjee and co-workers have demonstrated that FEP calculations can be used to determine the relative binding free energy (RBE) and predict selectivity.¹⁴ In another work, Zhang et al. showed that FEP computations could predict the differences in binding free energy of structurally related compounds that share a typical warhead.¹⁵ Moreover, we also showed that it is possible to obtain excellent agreement between theoretical and experimental binding free energy values for reversible covalent compounds through FEP calculations.²⁵ These studies support the usefulness of FEP techniques to predict the binding affinity of reversible covalent inhibitors in the lead optimization phase, providing critical guidelines for the design of specific modifications in the lead structure.²⁶ In particular, these techniques can be valuable to minimize the impact of activity cliffs (ACs) when structurally similar pairs of active compounds present a large difference in potency.^{27,28}

This study addresses the design, synthesis, and biochemical characterization of the inhibitory activity of a set of structurally related reversible covalent inhibitors of human cathepsin L (hCatL). Cathepsin L belongs to the family of lysosomal cysteine proteases (CPs), which are ubiquitous papain-like enzymes that play a vital role in protein degradation.^{12,29,30} With other secreted proteins, they can stimulate the cells in favor of tumor growth, resulting in cancer progression.^{31,32} Besides, hCatL is critical for the entry of SARS-CoV-2 and other coronaviruses, such as SARS-CoV and MERS-CoV, in human cells.^{33–36} As other CPs, hCatL presents a catalytic Cys/His dyad responsible for the enzyme's mechanism. When the proton is formally transferred from Cys to His, the thiolate ion performs a nucleophilic attack on a peptide bond's carbonyl group.^{37–40}

In this work, eight nitrile-based inhibitors (1a–e and 2a–c; see Figure 2) were subjected to FEP analyses. Derivatives 1c, 1e, 2a, 2b, and 2c were previously reported,⁴¹ and compound 1c was used as a prototype for the synthesis of *meta*-halogenated molecular pairs coded 1a, 1b, and 1d, once halogen bonding in P3 was previously identified with Gly61 of hCatL.²⁵ All ligands have the same scaffold but present differences in the terminal aromatic ring. They were split into two sets of compounds (1 and 2) to obtain a structural difference limited to a single heavy atom. Set 1 includes a series of *meta*- and *para*-halogenated compounds and set 2 corresponds to modified biphenyl derivatives. In all cases, the *in vitro* inhibitory activity was determined using kinetics assays. Finally, FEP calculations were used to rationalize the differences in the binding affinities of these inhibitors. The results point out that the changes in the covalent state might, to a large extent, be used to estimate a correct ranking of the compounds. Overall, this work applied a practical protocol for predicting the relative binding free energy of covalent inhibitors.

METHODS

Chemistry. The synthesis, characterization, and enzymatic assays of compounds 1c, 1e, 2a, 2b, and 2c were previously reported.⁴¹ The synthesis of compounds 1a, 1b, and 1d was performed similarly, where the synthetic route was developed to optimize the placement of P1, P2, and P3 groups (see the Experimental Section in the Supporting Information for details). Chemical reagents were purchased from Sigma-Aldrich (Merck) or Combi-Blocks, with at least 95% purity.

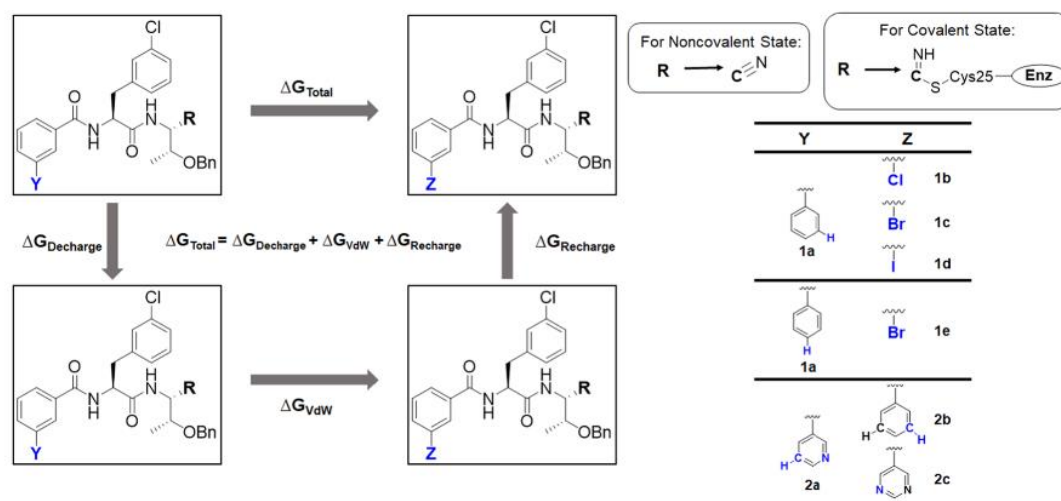


Figure 3. Schematic representation of the thermodynamic cycle for the alchemical binding free energy calculations. The cycle illustrates the path to obtain the relative free energy. The compounds with the Y group were mutated into the Z group (blue atoms), in the respective sets, in both covalent and noncovalent states. Thus, for each state, ΔG_{Total} is obtained in water and protein environments for the stepwise discharge–vdW–recharge strategy. The reference system used for the covalent state in the water environment corresponds to the ligand bonded to the side chain of Cys25, which in turn is connected to its adjacent Ser24 and Trp26.

They were used as received without further purification unless otherwise specified. All solvents were dried and distilled before use by standard procedures. ^1H and ^{13}C NMR spectra were recorded on Agilent spectrometers, models 400/54 and 500/54, operating at 400 MHz for ^1H (101 MHz for ^{13}C) or 500 MHz for ^1H (126 MHz for ^{13}C). Chemical shifts (δ) are reported relative to the internal standard TMS ($\delta = 0.0$ for ^1H and ^{13}C) or to the solvent residual signal (DMSO- d_6 : 2.50 ppm for ^1H and 39.52 ppm for ^{13}C). High-resolution mass spectra (HRMS) were obtained on a Thermo Scientific LTQ Velos Orbitrap spectrometer, operating in electrospray ionization conditions (ESI); values reported are the exact mass calculated using ChemDraw Professional 17.0 (calcd) and the mass found. Melting points were determined in a Microquímica Equipamentos, model MQAPF-302 apparatus and are not corrected.

Purification was performed by flash column chromatography on silica gel, with hexane/ethyl acetate mixtures as eluents. Purity analysis and chromatographic purifications were performed on a Shimadzu Prominence HPLC system (Kyoto, Japan) coupled to an Amazon Ion Trap mass spectrometer (Bruker GmbH, Bremen, Germany) operating in electrospray ionization conditions (ESI). The HPLC system was equipped with an LC-20AT/AD ternary pumping system, a SIL-20A autosampler, and a CTO-20A column oven (Shimadzu Corp). The column used for purity analysis was a Chiralpak IC 5.0 μm (4.6 mm \times 250.0 mm, Diacel Corporation, West Chester, PA). Gradient elution conditions were used, with a mobile phase composed of water (A) and acetonitrile (B), at a 0.5 mL/min flow rate. The gradient was set as follows: from 0.00 to 30.00 min, B was increased from 5 to 100%; from 30.01 to 40.0 min, B was kept at 100%; at 40.01 min, B was decreased to 5%, and from 40.01 to 50.00 min, B was held at 5%.

Enzyme Inhibition Study. Newly synthesized compounds were evaluated against hCatL using biochemical assays. Human cathepsin L (Enzo Life Sciences, Inc.) isolated from the human liver was assayed in a Biotek SynergyTM HT plate

reader at room temperature using an excitation wavelength of 360 nm and emission of 460 nm on Corning 96-well black flat-bottom microplates. The enzyme (1.9 nM) was activated with assay buffer (100 mM sodium acetate pH 5.5, 300 mM NaCl, 5 mM EDTA) containing 5 mM DTT, 0.014% of Triton X-100 for 20 min. A 3 mM stock solution of the fluorogenic substrate Z-Phe-Arg-AMC was prepared in DMSO. The final substrate concentration in the assay was equal to K_M ($K_M = 4.7 \mu\text{M}$). All inhibitors were solubilized in DMSO; then, it was diluted to reach the final concentrations ranging from 0.1 to 50 μM in the DMSO concentration that was 5% (10 μL). The experiments were performed in triplicate lasting for 5 min. Visual inspection and a pre-reading of the plate wells were carried out to check for possible precipitations and background fluorescence.

The initial rates of substrate hydrolysis under the first-order reaction were calculated using Gen5 Biotek software. The apparent inhibition constant K_i^{app} was determined by the nonlinear regression fit using Origin Pro 8 software, through the equation $V_s = \frac{V_0}{1 + \frac{[I]}{K_i^{\text{app}}}}$, where V_s is the steady-state rate, V_0

the rate in the absence of the inhibitor, and $[I]$ is the inhibitor concentration. The actual inhibition constant (K_i) was calculated using the Cheng–Prusoff equation for fast-binding inhibitors (see the Experimental Section in the [Supporting Information](#) for details).⁴²

System Setup for Molecular Dynamics Simulations.

Initial coordinates of the enzyme were taken from the hCatL crystal structure (PDB code: 2XU1).⁴³ The reference ligands (1a and 2a) were placed in the active site by aligning the coordinates of a dipeptidyl nitrile inhibitor in complex with cruzain⁴⁴ with the hCatL structure. The protonation states of the protein residues were evaluated through the PROPKA program.^{45,46} For each compound, the electrostatic potential was calculated using Gaussian09⁴⁷ software, with the 6-31(G) basis set⁴⁸ at the Hartree–Fock level of theory. Then, partial charges were obtained using a restrained electrostatic potential (RESP) approach⁴⁹ as implemented in the Antechamber package.⁵⁰ All of the ligands were parameterized with the

general amber force field (GAFF).⁵¹ In a further step, ligands **1a** and **2a** were solvated in a truncated octahedron TIP3P⁵² water box and Na⁺ cations were added to neutralize the system. This was done for both covalent and noncovalent systems. The AMBER18⁵³ suite of programs and the Amber ff14SB force field⁵⁴ were used to carry out the MD simulations, where the bonds containing hydrogen atoms were restrained applying the SHAKE algorithm.⁵⁵ The steepest descent and conjugate gradient algorithms were employed to minimize the systems in 10 000 cycles. Next, they were gradually heated in several steps until they reached 300 K with a Langevin thermostat.⁵⁶ Subsequently, 100 ns of MD simulations using the NPT ensemble were performed for the covalent and noncovalent states for the reference compounds **1a** and **2a**. Finally, RMSD trajectories were analyzed to select the starting point for the alchemical free energy calculations, choosing a stable structure in the trajectory.

Thermodynamic Cycle Used for Computing the Relative Binding Free Energy. According to Chatterjee and co-workers,¹⁴ covalent and noncovalent binding states are needed to estimate the overall binding free energy for reversible covalent compounds, as noted in eq 1.

$$\frac{1}{K_d} = A e^{-\Delta\Delta G_{dc}/RT} + B e^{-\Delta\Delta G_{dm}/RT} \quad (1)$$

where $\Delta\Delta G_{dc}$ and $\Delta\Delta G_{dm}$ denote the relative binding free energies of the covalent and noncovalent complexes, respectively, and K_d stands for the overall dissociation constant of the reversible covalent inhibitor.

Since the compounds in the set of alchemical transformations have the same core and warhead, A and B can be regarded as constants (eqs 2 and 3),¹⁴ and FEP calculations can be used to compute the free energy difference relative to compounds **1a** and **2a** (see below).

$$A = e^{-\Delta G_{dc}^{core}/RT} \quad (2)$$

$$B = e^{-\Delta G_{dm}^{core}/RT} \quad (3)$$

Herein, assume that the determination of constants A and B (eqs 2 and 3) is not fundamental for the relative binding free energy calculations since the structural modifications of ligands are located away from the warhead.

For the compounds bearing a halogen atom at P3, alchemical transformations were made relative to the reference inhibitor, which contains hydrogen in the benzamide ring (**1a**), to generate compounds with Cl (**1b**), Br (**1c**), and I (**1d**) at the *meta* position and Br (**1e**) at the *para* position (Figure 3). For the biphenyl derivatives, the phenylpyridine moiety (**2a**) was transformed into biphenyl (**2b**) and also in phenylpyrimidine (**2c**).

Note that we chose to carry out only the alchemical transformations within the same series of compounds since the perturbations involve relatively small changes; however, one might a priori consider the possibility to link the two series of compounds by performing the alchemical transformation of **1a** into **2a** or even preferably **2b** since benzene has a quadrupole term instead of the net dipole as in the case of pyridine, where electrostatic terms would be more important. Furthermore, transforming a hydrogen atom into a benzene ring can be a tricky task due to the large difference in the size of the perturbed groups, which can cause convergence problems. Accordingly, this transformation would require a smoother transformation protocol,¹⁸ which would require a significant

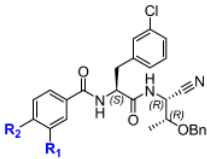
increase in computational resources. Another option could be the employment of benzene annihilation, where the transformation goes from benzene to hydrogen and seems to present fewer convergence issues, or alternatively the usage of a 'middle' scheme, where the initial λ is a mixture of both coupled and decoupled states ($\lambda = 0.5$) and its propagates in each direction ($0 \leftarrow \lambda = 0.5 \rightarrow 1$).^{57–59} For our purposes here, since the major aim is to examine the suitability of the FEP-based computational protocol proposed by Chatterjee and co-workers,¹⁴ for predicting the binding affinity of reversible covalent inhibitors of hCatL, this latter transformation was not further explored.




Alchemical Free Energy Calculations. Alchemical free energy calculations estimate the binding energy of compound B relative to reference compound A (Figure 3). Since free energy is a state function, the free energy difference between the end states represented in the thermodynamic cycle is independent on the path. Still, the choice of the pathway is significant for computational calculations.^{57–59} A coupling parameter (λ) is required to connect the thermodynamic states, ensuring a good overlap between the sampled structures at each step of the alchemical transformation along the path.

Here, we have used FEP⁶⁰ to compute the free energy change using the cycle shown in Figure 3, where the Y group was mutated into the Z group (blue atoms). Ligands were parameterized using the GAFF force field in conjunction with partial charges obtained using the RESP approach for covalent and noncovalent states. For the iodine atom in compound **1d**, the aug-cc-pVDZ-PP basis set⁶¹ was employed. To connect the thermodynamic states, 11 λ -windows were utilized from a fully coupled state ($\lambda = 1$) to a fully decoupled state ($\lambda = 0$), with an increment of 0.1 for each window. The strategy used here was a stepwise discharge–vdW–recharge. In the discharge step, where the electrostatic interactions are turned off, the following atoms were discharged separately: the hydrogen atom at the *meta* position of compound **1a**; the hydrogen atom at the *para* position of compound **1a**; the nitrogen atom of compound **2a** to be mutated into **2b**; and the CH (highlighted in blue in the Y group in Figure 3) also from compound **2a** to be transformed into **2c**. For the vdW step, the default softcore potential was used to guarantee a smooth change between the initial and final states. Finally, for the recharge step, the mutated atoms in compounds **1b**, **1c**, **1d**, **1e**, **2b**, and **2c** were gradually recharged until they reached the fully coupled state.

It is important to note that for halogenated compounds **1b–d**, an extra-point (EP) of the positive charge was applied in the recharge step. The purpose of the EP is to represent the so-called sigma-hole present in the heavier halogens and improve the description of the halogen bond between the X...OGly61 of hCatL.²⁴ Additional details about the definition of the EP can be found elsewhere.^{25,62,63} Considering that the Br atom in the P3 group in compound **1e** does not participate in a halogen bond interaction since it is at the *para* position instead of *meta*, we decided to not include the EP.

Each system was minimized in 20 000 cycles of the steepest descent. We have used 1 ns for the equilibrium phase and 5 ns for the production for each transformation in the discharge, vdW, and recharge steps, resulting in a total of 66 ns for each step. Considering that each change was accomplished for both covalent and noncovalent states in water and complexed with the protein, a total of 792 ns were performed for each alchemical transformation. All analyses of free energy out-

Table 1. Experimental K_i and Binding Free Energy ($\Delta G_{\text{Bind,exp}}$) Values for the Reversible Covalent Inhibitors against hCatL^a


Compound	R ₁ -P3 Moiety	R ₂ -P3 Moiety	K_i (nM)	$\Delta G_{\text{Bind,exp}}$ (kcal mol ⁻¹)
1a	H	H	9.80 ± 0.085	-10.92
1b	Cl	H	2.21 ± 0.085	-11.81
1c	Br	H	0.57 ± 0.092	-12.58
1d	I	H	0.42 ± 0.055	-12.79
1e	H	Br	7.65 ± 0.061	-11.07
2a		H	0.87 ± 0.028	-12.36
2b		H	47.0 ± 3.405	-10.00
2c		H	15.4 ± 0.390	-10.66

^aThe R groups attached to the phenyl ring at the *meta* and *para* positions in P3 are also represented.

comes were performed using a Python script, alchemical-analysis.py, developed by Mobley's Lab.⁶⁴

RESULTS AND DISCUSSION

Compound Selection. Our research group reported several peptidomimetic inhibitors of hCatL and other cysteine proteases in recent publications.^{25,41,65,66} These studies demonstrated that compounds bearing a halogen atom at the *meta* position of the P3 benzyl ring (Figure 2) presented selectivity for hCatL over other CPs and showed a good affinity for the enzyme. Compounds with the biphenyl ring at P3 and related derivatives also exhibit a high affinity for hCatL. Therefore, we selected three biphenyl derivatives (2a–c) and also a compound bearing a Br atom (1c) at the *meta* position and another compound with Br at the *para* position (1e) in P3,⁴⁰ which should not properly be oriented to facilitate the formation of the halogen bond interaction. In addition, we synthesized analogues of 1c by replacing bromine with chlorine (1b) or iodine (1d) to examine the influence exerted by the halogen bond formed by these atoms at the *meta* position. In this way, compounds were divided into two groups (Figure 2): group 1 includes molecules with halobenzenes at P3 and group 2 encompasses biphenyl derivatives. The inhibitory constant (K_i) for these compounds and their experimental $\Delta G_{\text{Bind,exp}}$ are shown in Table 1.

MD Simulations. In this work, 100 ns of MD simulations were performed for hCatL complexes with 1a and 2a, used as reference ligands for alchemical transformations. The two compounds adopt a similar arrangement upon binding to hCatL (see Figure S14 in the Supporting Information), and the largest structural diversity is limited to the benzyloxy moiety that fills subpocket S1, reflecting the larger size of this subpocket and the lack of specific interactions between ligands

and residues in this region. The stability of the ligand–protein complex was confirmed by monitoring selected interactions between the ligand and residues in the binding pocket (see Figures S15–S17 in the Supporting Information). The average distance from the nitrile (warhead) carbon to the Cys25 thiol group (Figure 4) for the noncovalent complexes corresponds to 3.56 (0.23) Å for the complex with 1a and 3.57 (0.25) Å for 2a. Thus, one can assume that thermal fluctuations may provide a pharmacophoric conformation for the ligand to facilitate the nucleophilic attack of the Cys25 sulfur atom, leading to the covalent adduct formation.^{38,40,67} Also, amide hydrogen bonds of the dipeptidyl compounds formed within the hCatL active site⁶⁸ (through the backbone C=O and NH of Gly68 and C=O of Asp162) were maintained in both covalent and noncovalent states for each inhibitor (see the Supporting Information for more details). Consequently, stable and suitable poses can be extracted and used as starting points for the alchemical calculations.^{69,70}

Free Energy Perturbation. Alchemical free energy methods can be used to rank ligands from the predicted RBFE values,^{16,70,71} which is less demanding in terms of computational cost than absolute binding affinities. In covalent ligands, it is also necessary to treat both the bound and unbound states. Since the compounds possess the same core and warhead, the RBFE was employed to compute the relative free energies for the alchemical transformation between a given pair of inhibitors in the covalent ($\Delta\Delta G_{\text{cov}}$) and noncovalent ($\Delta\Delta G_{\text{noncov}}$) states relative to the same process in water (eqs 4a and 4b).

$$\Delta\Delta G_{\text{noncov}} = \Delta G_{\text{total-protein,noncov}} - \Delta G_{\text{total-water}} \quad (4a)$$

$$\Delta\Delta G_{\text{cov}} = \Delta G_{\text{total-protein,cov}} - \Delta G_{\text{total-water}} \quad (4b)$$

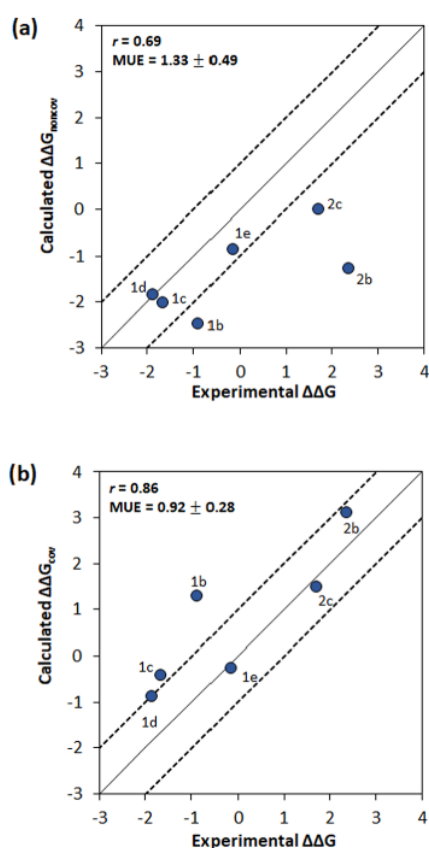


Figure 4. Comparison of the free energy changes obtained for alchemical transformations in the (a) noncovalent and (b) covalent states with the relative binding affinities derived from experimental inhibitory constants. All values are in kcal/mol. The dashed lines represent the ± 1 kcal/mol error from the experiments, and the black line represents $x = y$. Pearson's r and mean unsigned error (MUE) values are also indicated in the respective plots.

where the subscript total-protein, x stands for the addition of the free energy components related to the discharge, van der Waals, and recharge steps in the noncovalent ($x = \text{noncov}$) and covalent ($x = \text{cov}$) states as well as in aqueous solution ($x = \text{water}$).

The $\Delta\Delta G_{\text{noncov}}$ and $\Delta\Delta G_{\text{cov}}$ values are shown in Table 2, whereas the separate contributions arising from the discharge, van der Waals, and recharge steps in the complex with the protein and in aqueous solution are available in the Supporting Information (see Tables S1 and S2). The analysis of the results obtained with the overlapping distribution method supported an adequate sampling between adjacent distributions for all

alchemical transformations (see Figures S19–S34 in the Supporting Information), thus lending confidence to the convergence of estimated changes in the RBEF values.⁶⁴

The results in Table 2 reveal that there is a modest correlation between the free energy changes determined for the noncovalent state and the experimental ones (Pearson's correlation of 0.69; Figure 4a). In contrast, comparison with the free energy changes determined for the covalent state reveals a larger correspondence between the calculated and experimental data, as noted in a Pearson's correlation of 0.86 (Figure 4b). Furthermore, the mean unsigned error (MUE) is reduced from 1.33 kcal/mol for the $\Delta\Delta G_{\text{noncov}}$ values to 0.92 kcal/mol for the $\Delta\Delta G_{\text{cov}}$ ones. Besides, the difference between the $\Delta\Delta G_{\text{cov}}$ and experimental values is lower than or equal to 1 kcal/mol in four cases (changes from 1a to 1d and 1e and from 2a to 2b and 2c). Moreover, the ligands are correctly ranked when the covalent state is analyzed. Only compound 1b can be considered to be an exception, as it was predicted to be less potent than reference compound 1a. It is noteworthy that the values predicted for 1e and 2c practically matched the experimental ones.

The difference between the $\Delta\Delta G_{\text{cov}}$ and $\Delta\Delta G_{\text{noncov}}$ terms reflects the effect of the alchemical transformation on the free energy change associated with the transition from the noncovalent complex between the protein and the inhibitor to the covalently bound adduct (see Figure 5). If one assumes

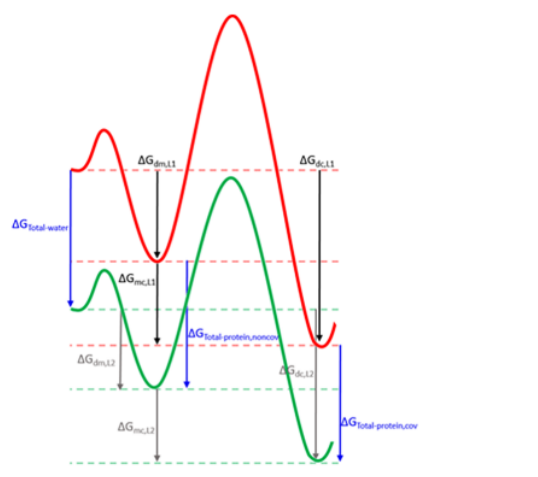


Figure 5. Profiles and components of the free energy for two hypothetical ligands, L1 (red) and L2 (green).

Table 2. RBEF Values Calculated for the Noncovalent ($\Delta\Delta G_{\text{noncov}}$) and Covalent ($\Delta\Delta G_{\text{cov}}$) States in the Alchemical Transformations between Ligands and the Free Energy Changes Determined from the Experimental Inhibition Constants^a

transformation	mutation	$\Delta\Delta G_{\text{exp}}$	$\Delta\Delta G_{\text{noncov}}$	$ \Delta x^b _{\text{noncov}}$	$\Delta\Delta G_{\text{cov}}$	$ \Delta x^b _{\text{cov}}$	$\Delta\Delta G_{\text{cov}} - \Delta\Delta G_{\text{noncov}}$
1a → 1b	H → Cl	-0.9	-2.5	1.6	+1.3	2.2	+3.8
1a → 1c	H → Br	-1.7	-2.0	0.3	-0.4	1.2	+1.6
1a → 1d	H → I	-1.9	-1.9	0.02	-0.9	1.0	+1.0
1a → 1e	H → <i>p</i> -Br	-0.1	-0.9	0.7	-0.3	0.1	+0.6
2a → 2b	N → CH	+2.4	-1.3	3.6	+3.1	0.7	+4.4
2a → 2c	CH → N	+1.7	0.0	1.7	+1.5	0.2	+1.5

^aAll values in kcal/mol. ^bAbsolute error between the calculated and experimental values.

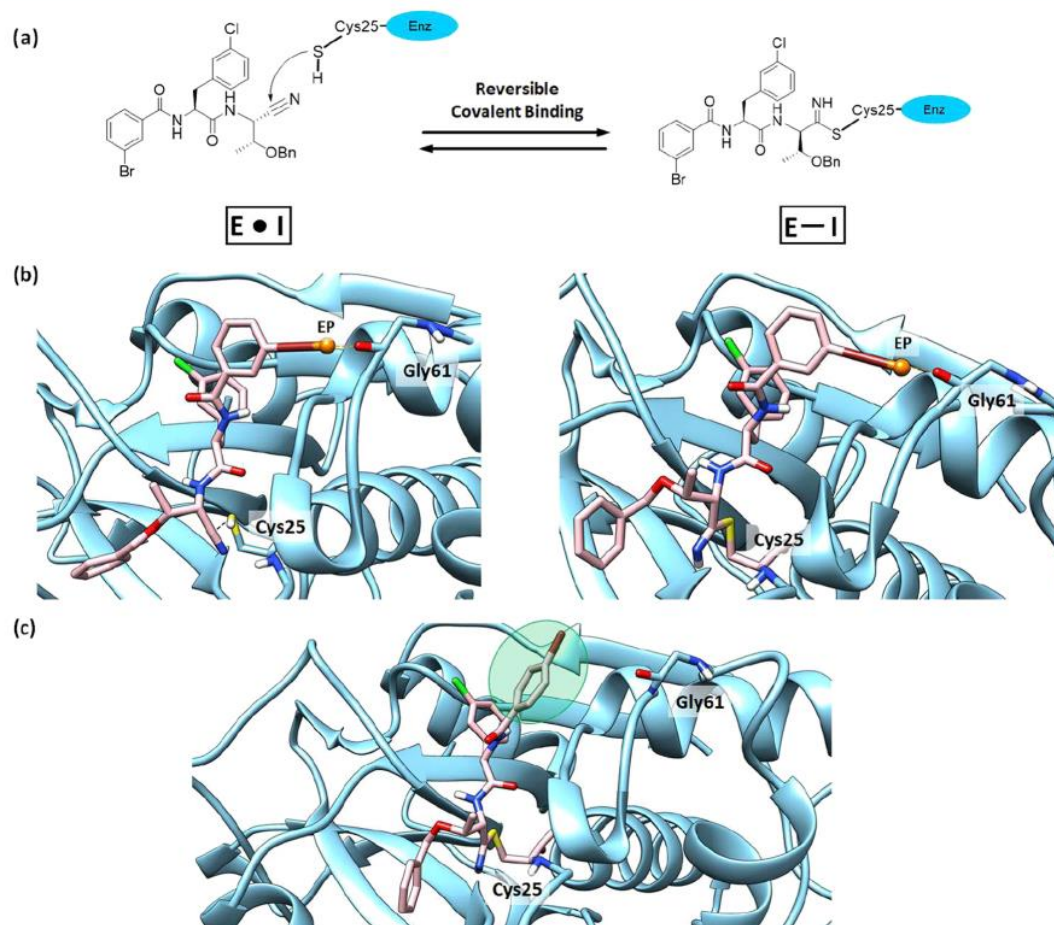


Figure 6. (a) Schematic representation of reversible covalent binding enabled by the proper positioning of the nitrile warhead close to the thiol group of the Cys25 residue ($E \bullet I$). After the reaction, the covalent adduct is formed ($E-I$). (b) The structure obtained from the last frame of 5 ns of MD for a fully coupled state ($\lambda = 1$) with the EP (orange) in the recharge step for compound **1c** in the noncovalent state (left) and the covalent state (right). (c) The putative binding mode of compound **1e** in the covalent state from the last frame of the recharge step (the P3 moiety is highlighted in green). The backbone of the inhibitor and Cys25 in the covalent state is pink, oxygen is red, nitrogen is blue, sulfur is yellow, chlorine is green, bromine is brown, and the added extra-point (EP) is orange.

that the stabilization afforded by the covalent bond between the warhead and the thiol group of Cys25 is essentially identical for two ligands (L1 and L2), the $\Delta\Delta G_{\text{cov}} - \Delta\Delta G_{\text{noncov}}$ term would account for the effect imposed by the formation of the covalent bond on the interactions formed by the chemical groups introduced in the moiety of the inhibitor that fills the P3 site (see Figure 2).

The results in Table 2 point out that this contribution is positive in all cases, suggesting that the stabilization afforded by the formation of the covalent bond imposes a destabilization in the interactions formed by the benzamide unit of the inhibitor, which is particularly important for the alchemical transformations **1a** \rightarrow **1b** (H \rightarrow Cl) and **2a** \rightarrow **2b** (pyridine \rightarrow benzene), as reflected in values of 3.8 and 4.4 kcal/mol, respectively.

The results obtained for compounds **1a–e** confirm the relevant role played by the halogen bond formed by **1b–d** with the carbonyl oxygen of Gly61 (see Figure 6b for compound **1c**). As noted above, this interaction was treated employing an EP of the positive charge in the recharge step (see Table S3 in the Supporting Information for details about the parametrization of the EP). With the EP's inclusion, the distance

between the halogen atom and Gly61 indicated the formation of the halogen bond^{25,43} in the recharge step, as noted in the average distances computed for the halogen bond in a fully coupled state (Table 3). Interestingly, the distance of the

Table 3. Average Distance (Å) between the Halogen Atom with EP and the Main Chain Oxygen of Gly61 in Noncovalent and Covalent States^a

compound	halogen atom	noncovalent state	covalent state
1b	Cl	3.49 (0.29)	4.17 (1.04)
1c	Br	3.46 (0.26)	3.46 (0.29)
1d	I	3.41 (0.21)	3.48 (0.38)

^aValues in parentheses are the corresponding standard deviations.

halogen bond in compound **1b** in the covalent state was larger than the value obtained for other compounds, which would suggest a lower stabilization in the interaction with the chlorine atom. This is in agreement with the results presented in Table 2, which show that the transformation **1a** \rightarrow **1b** is favorable in the noncovalent state ($\Delta\Delta G_{\text{noncov}} = -2.5$ kcal/mol), but penalized in the covalent state ($\Delta\Delta G_{\text{cov}} = +1.3$ kcal/mol),

which indicates that the covalent adduct should weaken the halogen bond with the chlorine atom. In turn, this reflects a better ability of the derivatives with bromine (**1c**) and iodine (**1d**) to retain the halogen bond interaction upon the structural rearrangement triggered by the formation of the covalent bond with Cys25 in the binding pocket.

On the other hand, when analyzing the binding mode obtained for compound **1e** (Figure 6c), it can be noticed that the P3 moiety does not fit the S3 subpocket of the enzyme properly. Consequently, the *p*-Br atom of the P3 benzyl ring is exposed to solvent and cannot perform the halogen bond in contrast to the other compounds in the series that present a halogen atom at the *meta* position. In this sense, due to the absence of this interaction, the K_i of **1e** is lower than the other halogenated compounds in the series (Table 1), it being very similar to the inhibition constant experimentally determined for the compound with hydrogen at the *meta* position (compound **1a**).

While halogen bonding explains the potency ranking for **1a–e**, the structural interpretation is less evident for the second set of compounds (**2a–c**). Figure 7 shows the putative binding

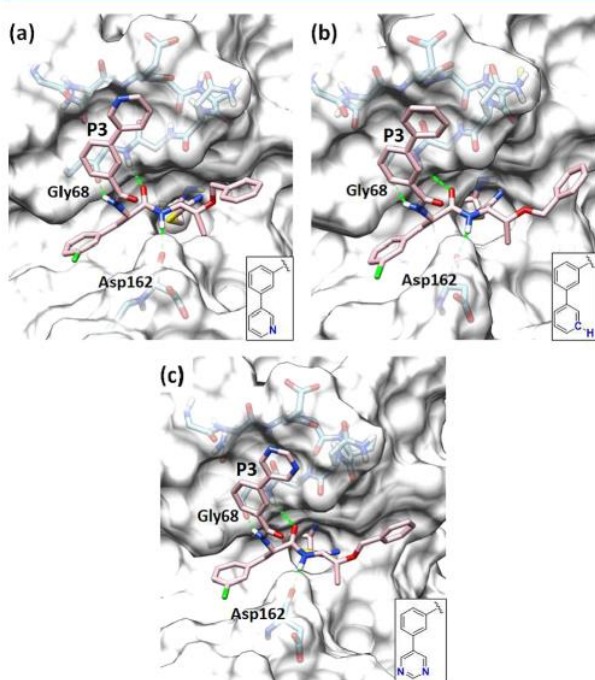


Figure 7. Putative binding mode of compounds (a) **2a**, (b) **2b**, and (c) **2c**. Hydrogen bonds are shown as dashed green lines. The 2D structural representation of the P3 group of each compound is represented in the right-bottom of each binding mode.

mode of **2a–c**, where essential hydrogen bonds are observed between the ligands and the main chain from Gly68 and Asp162 in the hCatL active site. These fundamental interactions are commonly observed for peptidomimetics interacting with hCatL.⁶⁶ It is also worth mentioning that the single-atom modification observed in **2a–c** is located at the P3 group (Figure 2), where the aromatic ring is to a large extent exposed to the solvent environment. Therefore, one can assume that the P3 group is not involved in critical interactions in the active site of hCatL. At this point, Hu and Bajorath have introduced a new category of activity cliffs that capture single-

atom modifications,⁷² where, for some cases, they could not identify the origin of ACs considering the protein–ligand complex structure. Indeed, the contribution of single-atom modification on ligand binding affinity is not always rationalized through the observation of vital ligand–protein interactions in the active site.⁷²

In this context, to gain insights into the energetic components that contribute to the destabilization in the binding of **2b** and **2c** ligands relative to **2a**, we have performed hydration free energy ($\Delta\Delta G_{\text{Hyd}}$) calculations using the same protocol detailed above for the compounds of set 2 in the gas phase to be combined with the alchemical transformation in aqueous solution and estimate the corresponding changes in the hydration free energy.

The transformation from **2a** to **2c**, which involves the alchemical change of pyridine into pyrimidine, led to a stabilization of -10.04 kcal/mol (see Table S4 in the Supporting Information). Surprisingly, the transformation of **2a** into **2b**, which involves the change of pyridine into benzene, led to an $\Delta\Delta G_{\text{Hyd}}$ of -22.74 kcal/mol, which is counterintuitive, since it implies a more favorable hydration free energy for the benzene-containing derivative. This latter finding, however, reflects the adoption of conformations markedly different from the bioactive one found in the ligand-bound complex during the in vacuum simulations. Accordingly, we performed the hydration free energy calculations for the alchemical transformations pyridine (**2a**) \rightarrow pyrimidine (**2c**) and pyridine (**2a**) \rightarrow benzene (**2b**) (see Table S5 in the Supporting Information). In the former case, the alchemical transformation leads to a $\Delta\Delta G_{\text{Hyd}}$ of -10.12 kcal/mol, which agrees with the value obtained for the whole compound (-10.04 kcal/mol; see above). Although no experimental data is available, at least to the best of our knowledge, for the hydration free energy of pyrimidine, this indicates that pyrimidine has more favorable solvation in water, which is in qualitative agreement with previous findings.⁷⁴ In turn, this suggests that binding of **2c** to the protein may be accompanied at least at some extent by a desolvation penalty relative to **2a**, in agreement with the experimental inhibitory potencies. On the other hand, the transformation pyridine (**2a**) \rightarrow benzene (**2b**) leads to an $\Delta\Delta G_{\text{Hyd}}$ of 4.03 kcal/mol, which is in agreement with the experimentally known relative hydration free energy of 3.8 kcal/mol, as deduced from the hydration free energies of pyridine (-4.7 kcal/mol) and benzene (-0.9 kcal/mol) taken from the FreeSolv database.⁷⁵ In this case, nevertheless, desolvation should favor the binding of the benzene-containing derivative **2b**, in agreement with the stabilization observed in the noncovalent state ($\Delta\Delta G_{\text{noncov}} = -1.3$ kcal/mol). However, this effect is counterbalanced by the destabilization observed in the covalent state ($\Delta\Delta G_{\text{cov}} = +3.1$ kcal/mol), which suggests that the formation of the covalent adduct alters the proper positioning of the biphenyl moiety in **2b** relative to the benzylpyridine unit in **2a**.

Importance of the Covalent State between Experimental and Theoretical $\Delta\Delta G$. Zhang and co-workers¹⁵ have shown that noncovalent state binding free energy presents no correlation with the experimental data in analyzing the reversible covalent modification of calpain-1 by ketoamide inhibitors. However, the covalent state offered a strong correlation to experimental values and the correlation to experimental data slightly improved upon the combination of data determined for covalent and noncovalent states. Interestingly, Chatterjee and colleagues¹⁴ have also shown

the importance of the bound state relative to the unbound one. They proposed that when the covalent state is at least -5.5 kcal/mol greater than the noncovalent one, only the covalent state can be used to compute the binding free energy. However, it is delicate to decide whether only the covalent complex can be used without calculations, making it necessary to consider both states. In another work, the authors predicted more potent inhibitors against hCatL only considering the covalent state.⁷³

Herein, our results comply with these works as long as the covalent state is found to have a more significant contribution to the binding free energy than the noncovalent state. Nevertheless, although the ligands can be suitably ranked considering the changes in the covalent state, our results also indicate that at least for certain compounds, the analysis of the changes determined for the noncovalent state is valuable to interpret the changes in inhibitory potency.

CONCLUSIONS

We calculated the RBFE for two sets of nitrile-based reversible covalent inhibitors and compared them with the experimental binding affinities. We have demonstrated that it is possible to access the overall binding considering only the covalent state. The results present a good correlation with the experimental data, and the ligands can be correctly ranked. Hence, it is assumed that the covalent state is more relevant for the binding free energy than the noncovalent state. However, the unbound state may be taken into account in free energy calculations to improve the analyses and interpretation of the results. In addition, our findings suggest that destabilization occurs for the compounds upon transition from the noncovalent state to the covalent one to form the covalently bound adduct.

Furthermore, for the halogenated compounds participating in a halogen bond, an extra-point of the positive charge should be used in the recharge step to represent this interaction aiming for the correct estimation of the binding free energy. Overall, the results show that RBFE calculations for covalent inhibitors using free energy perturbation are more dependent on the covalent state. The protocol employed in this work can be a powerful tool in developing new covalent drugs.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.1c00515>.

Results for synthetic and enzymatic assays; RMSD for the hCatL complex with **1a** and **2a** and the distance fluctuations of main interactions over molecular dynamics simulations; parameters used to include the extra-point in halogens; and outcomes of all FEP transformations, including the overlapping distribution graphics (ZIP) (PDF).

AUTHOR INFORMATION

Corresponding Authors

Jerônimo Lameira – Medicinal & Biological Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, 23566-590 São Carlos, SP, Brazil; Institute of Biological Science, Federal University of Pará, Rua Augusto Correa S/N, 66075-110 Belém, Pará, Brazil; orcid.org/0000-0001-7270-1517; Email: lameira@ufpa.br

Carlos A. Montanari – Medicinal & Biological Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, 23566-590 São Carlos, SP, Brazil; orcid.org/0000-0002-4963-0316; Email: carlos.montanari@usp.br

Authors

Vinicius Bonatto – Medicinal & Biological Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, 23566-590 São Carlos, SP, Brazil

Anwar Shamim – Medicinal & Biological Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, 23566-590 São Carlos, SP, Brazil

Fernanda dos R. Rocho – Medicinal & Biological Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, 23566-590 São Carlos, SP, Brazil

Andrei Leitão – Medicinal & Biological Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, 23566-590 São Carlos, SP, Brazil; orcid.org/0000-0002-6601-6609

F. Javier Luque – Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Science, Institute of Biomedicine (IBUB) and Institute of Theoretical and Computational Chemistry (IQTUB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain; orcid.org/0000-0002-8049-3567

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jcim.1c00515>

Notes

The authors declare no competing financial interest. The AMBER18 suite of programs (<https://ambermd.org/>) and the Amber ff14SB force field were used to carry out the MD simulations. The electrostatic potentials were calculated for each compound using Gaussian09 (<https://gaussian.com/glossary/g09/>). All analyses of free energy outcomes were performed using a Python script, `alchemical-analysis.py` (<https://github.com/MobleyLab/alchemical-analysis>). The files to conduct this work as the reference ligand files (pdb and parm7) can be found attached as the Supporting Information. In addition, the mol2 files for all ligands that were mutated from the reference compounds are also attached.

ACKNOWLEDGMENTS

This project was financed by Fundação de Amparo à Pesquisa do Estado de São Paulo—FAPESP grant #2018/15904-6, #2020/04653-2, and #2020/06543-0 and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant no. 304030/2018-0. The authors also thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the scholarships, the Generalitat de Catalunya (2017SGR1746), and the Ministerio de Economía y Competitividad (PID2020-117646RB-I00; MDM2017-0767; AEI/FEDER UE). For the computational resources, the authors thank the high-performance computing (HPC) provided by Superintendência de Tecnologia da Informação from University of São Paulo (STI-USP) and the Consorci de Serveis Universitaris de Catalunya (CSUC; Molecular Recognition project).

REFERENCES

(1) Sutanto, F.; Konstantinidou, M.; Dömling, A. Covalent Inhibitors: A Rational Approach to Drug Discovery. *RSC Med. Chem.* **2020**, *11*, 876–884.

- (2) Qiao, J.; Li, Y.-S.; Zeng, R.; Liu, F.-L.; Luo, R.-H.; Huang, C.; Wang, Y.-F.; Zhang, J.; Quan, B.; Shen, C.; Mao, X.; Liu, X.; Sun, W.; Yang, W.; Ni, X.; Wang, K.; Xu, L.; Duan, Z.-L.; Zou, Q.-C.; Zhang, H.-L.; Qu, W.; Long, Y.-H.-P.; Li, M.-H.; Yang, R.-C.; Liu, X.; You, J.; Zhou, Y.; Yao, R.; Li, W.-P.; Liu, J.-M.; Chen, P.; Liu, Y.; Lin, G.-F.; Yang, X.; Zou, J.; Li, L.; Hu, Y.; Lu, G.-W.; Li, W.-M.; Wei, Y.-Q.; Zheng, Y.-T.; Lei, J.; Yang, S. SARS-CoV-2 Mpro Inhibitors with Antiviral Activity in a Transgenic Mouse Model. *Science* **2021**, *371*, 1374–1378.
- (3) Ghosh, A. K.; Brindisi, M.; Shahabi, D.; Chapman, M. E.; Mesecar, A. D. Drug Development and Medicinal Chemistry Efforts toward SARS-Coronavirus and Covid-19 Therapeutics. *ChemMedChem* **2020**, *15*, 907–932.
- (4) Dai, W.; Zhang, B.; Jiang, X.-M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L.-K.; Xu, Y.; Yang, H.; Liu, H. Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease. *Science* **2020**, *368*, 1331–1335.
- (5) Breidenbach, J.; Lemke, C.; Pillaiyar, T.; Schäkel, L.; Al Hamwi, G.; Dieltz, M.; Gedschold, R.; Geiger, N.; Lopez, V.; Mirza, S.; Namasivayam, V.; Schiedel, A. C.; Sylvester, K.; Thimm, D.; Vielmuth, C.; Phuong Vu, L.; Zylina, M.; Bodem, J.; Gütschow, M.; Müller, C. E. Targeting the Main Protease of SARS-CoV-2: From the Establishment of High Throughput Screening to the Design of Tailored Inhibitors. *Angew. Chem., Int. Ed.* **2021**, *60*, 10423.
- (6) Hoffman, R. L.; Kania, R. S.; Brothers, M. A.; Davies, J. F.; Ferre, R. A.; Gajiwala, K. S.; He, M.; Hogan, R. J.; Kozminski, K.; Li, L. Y.; Lockner, J. W.; Lou, J.; Marra, M. T.; Mitchell, L. J.; Murray, B. W.; Nieman, J. A.; Noell, S.; Planken, S. P.; Rowe, T.; Ryan, K.; Smith, G. J.; Solowiej, J. E.; Stepan, C. M.; Taggart, B. Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. *J. Med. Chem.* **2020**, *63*, 12725–12747.
- (7) Pfizer unveils its oral SARS-CoV-2 inhibitor <https://cen.acs.org/acs-news/acs-meeting-news/Pfizer-unveils-oral-SARS-CoV/99/i13>.
- (8) Gehringer, M.; Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62*, 5673–5724.
- (9) Lonsdale, R.; Burgess, J.; Colclough, N.; Davies, N. L.; Lenz, E. M.; Orton, A. L.; Ward, R. A. Expanding the Armory: Predicting and Tuning Covalent Warhead Reactivity. *J. Chem. Inf. Model.* **2017**, *57*, 3124–3137.
- (10) Baillie, T. A. Targeted Covalent Inhibitors for Drug Design. *Angew. Chem., Int. Ed.* **2016**, *55*, 13408–13421.
- (11) Bonatto, V.; Batista, P. H. J.; Cianni, L.; De Vita, D.; Silva, D. G.; Cedron, R.; Tezuka, D. Y.; de Albuquerque, S.; Moraes, C. B.; Franco, C. H.; Lameira, J.; Leitão, A.; Montanari, C. A. On the Intrinsic Reactivity of Highly Potent Trypanocidal Cruzain Inhibitors. *RSC Med. Chem.* **2020**, *11*, 1275–1284.
- (12) Cianni, L.; Feldmann, C.; Gilberg, E.; Gütschow, M.; Juliano, L.; Leitao, A.; Bajorath, J.; Montanari, C. A. Can Cysteine Protease Cross-Class Inhibitors Achieve Selectivity? *J. Med. Chem.* **2019**, *62*, 10497–10525.
- (13) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The Resurgence of Covalent Drugs. *Nat. Rev. Drug Discovery* **2011**, *10*, 307–317.
- (14) Chatterjee, P.; Botello-Smith, W. M.; Zhang, H.; Qian, L.; Alsamrah, A.; Kent, D.; Lacroix, J. J.; Baudry, M.; Luo, Y. Can Relative Binding Free Energy Predict Selectivity of Reversible Covalent Inhibitors? *J. Am. Chem. Soc.* **2017**, *139*, 17945–17952.
- (15) Zhang, H.; Jiang, W.; Chatterjee, P.; Luo, Y. Ranking Reversible Covalent Drugs: From Free Energy Perturbation to Fragment Docking. *J. Chem. Inf. Model.* **2019**, *59*, 2093–2102.
- (16) Courmia, Z.; Allen, B. K.; Beuming, T.; Pearlman, D. A.; Radak, B. K.; Sherman, W. Rigorous Free Energy Simulations in Virtual Screening. *J. Chem. Inf. Model.* **2020**, *60*, 4153–4169.
- (17) Hansen, N.; van Gunsteren, W. F. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014**, *10*, 2632–2647.
- (18) Jespers, W.; Isaksen, G. V.; Andberg, T. A. H.; Vasile, S.; van Veen, A.; Åqvist, J.; Brandsdal, B. O.; Gutiérrez-de-Terán, H. QresFEP: An Automated Protocol for Free Energy Calculations of Protein Mutations in Q. *J. Chem. Theory Comput.* **2019**, *15*, 5461–5473.
- (19) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703.
- (20) Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate Calculation of the Absolute Free Energy of Binding for Drug Molecules. *Chem. Sci.* **2016**, *7*, 207–218.
- (21) Pérez-Benito, L.; Casajuaana-Martin, N.; Jiménez-Rosés, M.; Van Vlijmen, H.; Tresadern, G. Predicting Activity Cliffs with Free-Energy Perturbation. *J. Chem. Theory Comput.* **2019**, *15*, 1884–1895.
- (22) Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Predictions of Ligand Selectivity from Absolute Binding Free Energy Calculations. *J. Am. Chem. Soc.* **2017**, *139*, 946–957.
- (23) Ngo, S. T.; Nguyen, H. M.; Thuy Huong, L. T.; Quan, P. M.; Truong, V. K.; Tung, N. T.; Vu, V. V. Assessing Potential Inhibitors of SARS-CoV-2 Main Protease from Available Drugs Using Free Energy Perturbation Simulations. *RSC Adv.* **2020**, *10*, 40284–40290.
- (24) Zhang, C.-H.; Stone, E. A.; Deshmukh, M.; Ippolito, J. A.; Ghahremanpour, M. M.; Tirado-Rives, J.; Spasov, K. A.; Zhang, S.; Takeo, Y.; Kudalkar, S. N.; Liang, Z.; Isaacs, F.; Lindenbach, B.; Miller, S. J.; Anderson, K. S.; Jorgensen, W. L. Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations. *ACS Cent. Sci.* **2021**, *7*, 467–475.
- (25) Lameira, J.; Bonatto, V.; Cianni, L.; Dos Reis Rocho, F.; Leitão, A.; Montanari, C. A. Predicting the Affinity of Halogenated Reversible Covalent Inhibitors through Relative Binding Free Energy. *Phys. Chem. Chem. Phys.* **2019**, *21*, 24723–24730.
- (26) Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L. Principles of Early Drug Discovery. *Br. J. Pharmacol.* **2011**, *162*, 1239–1249.
- (27) Stumpfe, D.; Bajorath, J. Exploring Activity Cliffs in Medicinal Chemistry. *J. Med. Chem.* **2012**, *55*, 2932–2942.
- (28) Stumpfe, D.; Hu, H.; Bajorath, J. Evolving Concept of Activity Cliffs. *ACS Omega* **2019**, *4*, 14360–14368.
- (29) Turk, B.; Turk, D.; Turk, V. Lysosomal Cysteine Proteases: More than Scavengers. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **2000**, *1477*, 98–111.
- (30) Nishimura, Y.; Furuno, K.; Kato, K. Biosynthesis and Processing of Lysosomal Cathepsin L in Primary Cultures of Rat Hepatocytes. *Arch. Biochem. Biophys.* **1988**, *263*, 107–116.
- (31) De Palma, M.; Biziato, D.; Petrova, T. V. Microenvironmental Regulation of Tumour Angiogenesis. *Nat. Rev. Cancer* **2017**, *17*, 457–474.
- (32) Tabish, T. A.; Pranjol, M. Z. I.; Whatmore, J. L.; Zhang, S. Status and Future Directions of Anti-Metastatic Cancer Nanomedicines for the Inhibition of Cathepsin L. *Front. Nanotechnol.* **2020**, *2*, 1–10.
- (33) Liu, T.; Luo, S.; Libby, P.; Shi, G.-P. Cathepsin L-Selective Inhibitors: A Potentially Promising Treatment for COVID-19 Patients. *Pharmacol. Ther.* **2020**, *213*, No. 107587.
- (34) Ou, X.; Liu, Y.; Lei, X.; Li, P.; Mi, D.; Ren, L.; Guo, L.; Guo, R.; Chen, T.; Hu, J.; Xiang, Z.; Mu, Z.; Chen, X.; Chen, J.; Hu, K.; Jin, Q.; Wang, J.; Qian, Z. Characterization of Spike Glycoprotein of SARS-CoV-2 on Virus Entry and Its Immune Cross-Reactivity with SARS-CoV. *Nat. Commun.* **2020**, *11*, No. 1620.

- (35) Simmons, G.; Gosalia, D. N.; Rennekamp, A. J.; Reeves, J. D.; Diamond, S. L.; Bates, P. Inhibitors of Cathepsin L Prevent Severe Acute Respiratory Syndrome Coronavirus Entry. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11876–11881.
- (36) Kleine-Weber, H.; Elzayat, M. T.; Hoffmann, M.; Pöhlmann, S. Functional Analysis of Potential Cleavage Sites in the MERS-Coronavirus Spike Protein. *Sci. Rep.* **2018**, *8*, No. 16597.
- (37) Erez, E.; Fass, D.; Bibi, E. How Intramembrane Proteases Bury Hydrolytic Reactions in the Membrane. *Nature* **2009**, *459*, 371–378.
- (38) 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* **2018**, *57*, 3176–3190.
- (39) Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 Mpro by QM/MM Computational Methods. *Chem. Sci.* **2020**, *11*, 10626–10630.
- (40) Oanca, G.; Asadi, M.; Saha, A.; Ramachandran, B.; Warshel, A. Exploring the Catalytic Reaction of Cysteine Proteases. *J. Phys. Chem. B* **2020**, *124*, 11349–11356.
- (41) Cianni, L.; Dos Reis Rocho, F.; Bonatto, V.; Cardoso Prado Martins, F.; Lameira, J.; Leitão, A.; Montanari, C. A.; Shamim, A. Design, Synthesis and Stepwise Optimization of Nitrile-Based Inhibitors of Cathepsins B and L. *Bioorg. Med. Chem.* **2021**, No. 115827.
- (42) 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. *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
- (43) Hardegger, L. A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; Plancher, J. M.; Hartmann, G.; Banner, D. W.; Haap, W.; Diederich, F. Systematic Investigation of Halogen Bonding in Protein-Ligand Interactions. *Angew. Chem., Int. Ed.* **2011**, *50*, 314–318.
- (44) Avelar, L. A. A.; Camilo, C. D.; De Albuquerque, S.; Fernandes, W. B.; Gonçalves, 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 Trop. Dis.* **2015**, *9*, No. e0003916.
- (45) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.
- (46) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of P K a Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284–2295.
- (47) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, 2009.
- (48) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. 6-31G* Basis Set for Third-Row Atoms. *J. Comput. Chem.* **2001**, *22*, 976–984.
- (49) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem. A* **1993**, *97*, 10269–10280.
- (50) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.
- (51) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (52) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (53) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGra, S.; York, D. M.; Kollman, P. A. *Amber18*; University of California: San Francisco, 2018.
- (54) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (55) Andersen, H. C. Rattle: A “velocity” Version of the Shake Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52*, 24–34.
- (56) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin Dynamics of Peptides: The Frictional Dependence of Isomerization Rates of N-Acetylalanyl-N'-Methylamide. *Biopolymers* **1992**, *32*, 523–535.
- (57) Jespers, W.; Esguerra, M.; Åqvist, J.; Gutiérrez-de-Terán, H. QligFEP: An Automated Workflow for Small Molecule Free Energy Calculations in Q. *J. Cheminf.* **2019**, *11*, No. 26.
- (58) Michel, J.; Foloppe, N.; Essex, J. W. Rigorous Free Energy Calculations in Structure-Based Drug Design. *Mol. Inf.* **2010**, *29*, 570–578.
- (59) Kuhn, M.; Firth-Clark, S.; Tosco, P.; Mey, A. S. J. S.; MacKey, M.; Michel, J. Assessment of Binding Affinity via Alchemical Free-Energy Calculations. *J. Chem. Inf. Model.* **2020**, *60*, 3120–3130.
- (60) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
- (61) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically Convergent Basis Sets with Relativistic Pseudopotentials. II. Small- Core Pseudopotentials and Correlation Consistent Basis Sets for the Post- D Group 16 – 18 Elements. *Phys. J Chem.* **2003**, *119*, 11113–11123.
- (62) Ibrahim, M. A. A. Molecular Mechanical Study of Halogen Bonding in Drug Discovery. *J. Comput. Chem.* **2011**, *32*, 2564–2574.
- (63) Rendine, S.; Pieraccini, S.; Forni, A.; Sironi, M. Halogen Bonding in Ligand–receptor Systems in the Framework of Classical Force Fields. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19508–19516.
- (64) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the Analysis of Free Energy Calculations. *J. Comput. Aided. Mol. Des.* **2015**, *29*, 397–411.
- (65) Cianni, L.; Rocho, F.; dos, R.; Rosini, F.; Bonatto, V.; Ribeiro, J. F. R.; Lameira, J.; Leitão, A.; Shamim, A.; Montanari, C. A. Optimization Strategy of Single-Digit Nanomolar Cross-Class Inhibitors of Mammalian and Protozoa Cysteine Proteases. *Bioorg. Chem.* **2020**, *101*, No. 104039.
- (66) Cianni, L.; Lemke, C.; Gilberg, E.; Feldmann, C.; Rosini, F.; Rocho, F.; dos, R.; Ribeiro, J. F. R.; Tezuka, D. Y.; Lopes, C. D.; de Albuquerque, S.; Bajorath, J.; Laufer, S.; Leitão, A.; Gütschow, M.; Montanari, C. A. Mapping the S1 and S1' Subsites of Cysteine Proteases with New Dipeptidyl Nitrile Inhibitors as Trypanocidal Agents. *PLoS Neglected Trop. Dis.* **2020**, *14*, No. e0007755.
- (67) Dos Santos, A. M.; Cianni, L.; De Vita, D.; Rosini, F.; Leitão, A.; Laughton, C. A.; Lameira, J.; Montanari, C. A. Experimental Study and Computational Modelling of Cruzain Cysteine Protease Inhibition by Dipeptidyl Nitriles. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24317–24328.

(68) Asaad, N.; Bethel, P. A.; Coulson, M. D.; Dawson, J. E.; Ford, S. J.; Gerhardt, 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. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4280–4283.

(69) Fratev, F.; Sirimulla, S. An Improved Free Energy Perturbation FEP+ Sampling Protocol for Flexible Ligand-Binding Domains. *Sci. Rep.* **2019**, *9*, No. 16829.

(70) Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57*, 2911–2937.

(71) Song, L. F.; Merz, K. M. Evolution of Alchemical Free Energy Methods in Drug Discovery. *J. Chem. Inf. Model.* **2020**, *60*, 5308–5318.

(72) Hu, H.; Bajorath, J. Activity Cliffs Produced by Single-Atom Modification of Active Compounds: Systematic Identification and Rationalization Based on X-Ray Structures. *Eur. J. Med. Chem.* **2020**, *207*, No. 112846.

(73) Kuhn, B.; Tichý, M.; Wang, L.; Robinson, S.; Martin, R. E.; Kuglstatter, A.; Benz, J.; Giroud, M.; Schirmeister, T.; Abel, R.; Diederich, F.; Hert, J. Prospective Evaluation of Free Energy Calculations for the Prioritization of Cathepsin L Inhibitors. *J. Med. Chem.* **2017**, *60*, 2485–2497.

(74) Loeffler, J. R.; Schauperl, M.; Liedl, K. R. Hydration of Aromatic Heterocycles as an Adversary of π -Stacking. *J. Chem. Inf. Model.* **2019**, *59*, 4209–4219.

(75) Mobley, D. L.; Guthrie, J. P. FreeSolv: A Database of Experimental and Calculated Hydration Free Energies, with Input Files. *J. Comput. Aided Mol. Des.* **2014**, *28*, 711–720.

7. CONCLUSIONS

In this thesis, we conducted a comprehensive literature review, emphasizing the significance of nitrile-based groups in reactive moieties of covalent inhibitors. With the increasing interest in enzymes featuring catalytic Cys or Ser residues, such as cysteine proteases, the development of covalent reversible inhibitors becomes particularly valuable, and nitrile groups and their derivatives play a pivotal role in this context. Additionally, for non-catalytic Cys residues, nitrile-containing groups, such as cyanoacrylamides, may also be of interest for their ability to react with these amino acids.

Modeling reversible covalent inhibitors presents a considerable challenge as it requires considering both the warhead group and non-covalent interactions with the macromolecular target. In this regard, we employed QM methodology to obtain quantum descriptors that correlate the reactivity of the warhead with the experimentally measured potency against cruzain. These descriptors can be applied in future studies employing QSAR or machine learning models.

Furthermore, considering the protein environment in warhead modeling is crucial, particularly concerning selectivity and how the rest of the binding site influences the formation of the covalent adduct. For the first time, we employed the QM-cluster approach for covalent inhibitors, and the results are consistent with expectations and in agreement with other methods such as QM/MM. Using the QM-cluster methodology, we successfully predicted that an aldehyde-based warhead is more reactive than a nitrile-based warhead, with an error of only 0.25 kcal/mol compared to experimental data and matched molecular pair analysis.

Applying QM-cluster in future covalent inhibitor discovery projects holds great potential and demonstrates its value. Additionally, the QM-cluster methodology can likely be extended to irreversible inhibitors, wherein modeling the transition state is necessary.

Furthermore, using relative binding free energy calculations with FEP, we demonstrated the feasibility of modeling reversible covalent inhibitors using only the covalent state complex. While the unbound complex is helpful for analysis and interpretation, the covalent complex appears to be more critical for the binding free energy.

Overall, the literature review presented in this thesis provides valuable guidance for medicinal chemists in selecting appropriate compounds containing

nitrile groups as reactive moieties for the development of new covalent inhibitors. Moreover, the *in silico* results obtained have significant implications for the understanding and future development of reversible covalent inhibitors targeting cysteine proteases and potentially other targets. Whether for warhead modeling, non-covalent interactions, or the overall ligand modeling to obtain absolute energy against a specific target, computational chemistry proves to be a valuable tool in the development of reversible covalent inhibitors. Various methods can be applied depending on the specific research problem at hand.

In conclusion, computational chemistry is a powerful ally in the development of reversible covalent inhibitors, offering a range of methods that can be tailored to address specific research objectives. The findings presented in this thesis contribute to advancing our knowledge and pave the way for further research and innovation in covalent inhibitor design.

APPENDIX A – Supporting Information for On the Intrinsic Reactivity of Highly Potent Trypanocidal Cruzain Inhibitors

Reproduced from BONATTO, V.; BATISTA, P. H. J.; CIANNI, L.; DE VITA, D. ; SILVA, D. G. ; CEDRON, R.; TEZUKA, D. Y.; DE ALBUQUERQUE, S.; MORAES, C. B.; FRANCO, C. H.; LAMEIRA, J.; LEITÃO, A.; MONTANARI, C. A. . On the intrinsic reactivity of highly potent trypanocidal cruzain inhibitors. **RSC Medicinal Chemistry**, Cambridge, v. 11, p. 1275-1284, 2020 with permission from the Royal Society of Chemistry.

Cite this: <https://pubs.rsc.org/en/content/articlelanding/2020/md/d0md00097c/unauth>

Supporting Information for:

**On the Intrinsic Reactivity of Highly Potent Trypanocidal
Cruzain Inhibitors**

Vinicius Bonatto^{1§}, Pedro Henrique Jatai Batista^{1§}, Lorenzo Cianni¹, Daniela De Vita¹, Daniel G. Silva¹, Rodrigo Cedron¹, Daiane Y. Tezuka^{1,2}, Sérgio de Albuquerque², Carolina Borsoi Moraes³, Caio Haddad Franco³, Jerônimo Lameira^{1,4}, Andrei Leitão^{1*}, Carlos A. Montanari^{1*}

¹Medicinal Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, Avenue Trabalhador Sancarlene, 400, 23566-590, São Carlos/SP, Brazil

²Ribeirão Preto School of Pharmaceutical Sciences, University of São Paulo, Ribeirão Preto, São Paulo, Brazil

³Laboratório Nacional de Biotecnologia (LNBio), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, São Paulo, Brazil

⁴Laboratório de Planejamento e Desenvolvimento de Fármacos. Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará. Rua Augusto Corrêa 01, CP 66075-110. Belém-PA

[§]These authors contributed equally to the work.

*corresponding authors: carlos.montanari@usp.br (CAM); andleitao@iqsc.usp.br (AL)

INDEX

1. Synthesis and characterization	3
2. Electronic parameters	10
3. Half-life measurement and decay constant by HPLC	12
4. Reference	16
Figure S1. ¹ H-NMR (400 MHz, CDCl ₃) of compound 4	7
Figure S2. ¹ H-NMR (400 MHz, DMSO- <i>d</i> ₆) of compound Neq0673	8
Figure S3. ¹³ C-NMR (100 MHz, DMSO- <i>d</i> ₆) of compound Neq0673	8
Figure S4. ¹ H-NMR (400 MHz, DMSO- <i>d</i> ₆) of compound Neq0646	9
Figure S5. ¹³ C-NMR (100 MHz, DMSO- <i>d</i> ₆) of compound Neq0646	9
Figure S6. The putative linear equation and the coefficient of determination obtained through the linear correlation between the calculated parameter and pK _i against Cz. a) Ionization potential b) electron affinity c) chemical potential d) hardness e) electrophilic Fukui function and f) global electrophilicity, all parameters values are in eV but electrophilic Fukui function	11
Figure S7. Decay curve of Neq0490 with cysteine.	12
Figure S8. Decay curve of Neq0570 with cysteine	12
Figure S9. Decay curve of Neq0656 with cysteine	13
Figure S10. Decay curve of Neq0656 with glutathione.	13
Figure S11. Decay curve of Nilvadipine with cysteine.	14
Figure S12. Blank of Neq0570.	14
Figure S13. Blank of Neq0409.	15
Figure S14. Blank of Neq0656.	15
Figure S15. Blank of Neq0690.	15

1. Synthesis and characterization

General Consideration

Melting points were determined on a Büchi 510 oil bath apparatus and are uncorrected. Infrared spectra were obtained from FT-IR Thermo Scientific Nicolet 380. Reagents, starting materials and solvents were of commercial quality and were used without further purification unless otherwise stated. All syntheses started with enantiopure amino acids. TLC analysis was carried out on Merck 60 F₂₅₄ silica gel plates and visualized under UV light at 254 nm and 365 nm or by using a ninhydrin staining solution.

Purity was determined with an LC-MS instrument (AmaZon SL ESI-MS, Shimadzu LC) with a cellulose-2 Phenomenex column (250 x 4.6 mm, 5 µm) or a Diacel column (IC-chiralpak, 250 x 4.6 mm, 5 µm). Isocratic elution with MeCN and water was applied as specified, stop time 60 min, flow 0.5 mL/min. NMR spectra were recorded on Bruker Avance 400 MHz and Bruker Avance DRX 500 MHz NMR spectrometers. Chemical shifts are reported in ppm relative to TMS or the residual proton peak of the re-protonated deuterated solvent, and the spectra were calibrated against the residual proton peak of the used deuterated solvent. The following symbols indicate spin multiplicities: s (singlet), s br (broad singlet), d (doublet), dd (doublet of doublet), t (triplet), tt (triplet of triplet), q (quartet), sept (septet), and m (multiplet). HRMS spectra were recorded on a Thermo Scientific LTQ Velos Orbitrap, in electrospray ionization (ESI) mode by direct injection.

Synthesis of 4-nitroisoxazole (**1**): isoxazole (15 mmol, 960 μ L) was dissolved in TFAA (7.3 mL); then, NH_4NO_3 (22.5 mmol, 1.81 g) was added in 0.3-g portions, each 15 min, keeping the reaction mixture at 25-30°C. After complete addition, the mixture was kept at room temperature for 2 h after that poured in ice water (30 mL) and this aqueous washing was extracted with CHCl_3 (3 x 15 mL); the combined organic extracts were dried over Na_2SO_4 and evaporated (bath at room temperature) to give an oil that was triturated with *n*-hexane to give a yellow solid (50 % yield). $^1\text{H-NMR}$ (CDCl_3) δ = 9.29 (s, 1H), 8.83 (s, 1H) ppm¹.

Synthesis of 4-aminoisoxazole (**2**): to a yellow solution of 4-nitroisoxazole (**1**, 160 mg, 1.4 mmol) in 6 M HCl (7 mL) SnCl_2 (1.327 g, 7 mmol) was added in one portion. After 1.5 h at room temperature, the resulting orange solution was treated with a saturated solution of Na_2CO_3 until pH was 9. The formed solid was removed by filtration, and the aqueous solution was extracted with ethylacetate (5 x 50 mL); the organic phase, dried over MgSO_4 , was evaporated to give a brown oil (R_f = 0.64 ethylacetate 100% / silica) stored at 4°C and inert atmosphere (65% yield). $^1\text{H-NMR}$ ($\text{DMSO-}d_6$) δ = 8.16 (s, 1H), 8.13 (s, 1H), 4.26 (s br, 2H) ppm.

Synthesis of 2-benzamido-3-phenylpropanoic acid (**3**) or 3-(1H-indol-3-yl)-2-(phenylformamido)propanoic acid (**4**) 2.75 mmol of the corresponding amino acid was dissolved in 1M NaOH (6 mL) in an ice-bath. Benzoyl chloride (261 μ L, 2.25 mmol) was added. After 5 min, the reaction mixture was allowed to stand at room temperature. After 20 min, the solution was cooled in ice and 1M KHSO_4 (16 mL) was added slowly. The obtained white solid was washed with 1 M KHSO_4 (3 x 5 mL), H_2O (10 x 3 mL), 9:1 EtOH: H_2O (3 x 3 mL) and dried under vacuum on P_2O_5

(yield 88%). For compound **3**: ^1H NMR (500 MHz, $\text{DMSO-}d_6$): 12.62 (s br, 1H, D_2O exchange), 10.75 (s, 1H, D_2O exchange), 8.56 (m, 1H, D_2O exchange), 7.28 (m, 8H), 7.13 (m, 2H), 4.54 (m, 1H), 3.29 (m, 1H), 3.19 (m, 1H). For compound **4**: ^1H NMR (500 MHz, $\text{DMSO-}d_6$): 12.69 (s br, 1H, D_2O exchange), 10.80 (s, 1H, D_2O exchange), 8.61 (d, $J = 8.0$ Hz, 1H, D_2O exchange), 7.81 (d, $J = 8.5$ Hz, 2H), 7.59 (d, $J = 8$ Hz, 1H), 7.51 (t, $J = 7.0$ Hz, 1H), 7.44 (t, $J = 7.5$ Hz, 2H), 7.31 (d, $J = 8.0$ Hz, 1H), 7.20 (s, 1H), 7.05 (t, $J = 7.5$ Hz, 1H), 6.98 (t, $J = 7.5$ Hz, 1H), 4.65 (m, 1H), 3.30 (m, 1H, H_2O overlapping), 3.19 (m, 1H).

Synthesis of N-(1-(isoxazol-4-ylamino)-1-oxo-3-phenylpropan-2-yl)benzamide (**Neq0646**): to a suspension of (\pm)-2-benzamido-3-phenylpropanoic acid (**3**, 216 mg, 0.70 mmol), HOBt (123 mg, 0.91 mmol) and EDC (175 mg, 0.91 mmol) in CH_2Cl_2 (8 mL) were added under argon at 0°C . After stirring of 1 hour at room temperature, the mixture was kept on ice-bath, and a solution of 4-aminoisoxazole (235 mg, 2.80 mmol) in dry CH_2Cl_2 (2 mL) was added. The resulting mixture was kept overnight at room temperature, then the solvent was evaporated, and the residue treated with AcOEt (30 mL) and washed H_2O (2 x 20 mL) and brine (2 x 20 mL). The organic phase was dried over MgSO_4 and evaporated to give a crude residue that was purified by chromatographic column on silica using as mobile phase $\text{CHCl}_3/\text{AcOEt}$ (1:1), to give with solid ($R_f = 0.4$) crystallized from AcOEt (36% yield).

Secondary purification was carried out on cellulose-2 Phenomenex column, in isocratic elution with a flow rate of 2.36 mL min^{-1} , at 32°C ; the mobile phase composition was *n*-hexane/ethanol (70:30) (v/v) to give Neq0646. ^1H -NMR (500

MHz, DMSO-*d*₆) δ = 10.80 (s br, 1H), 10.51 (s, 1H), 9.13 (s, 1H), 8.72 (d, *J* = 8.0 Hz, 1H), 8.63 (s, 1H), 7.84 (m, 2H), 7.69 (d, *J* = 8.0 Hz, 1H), 7.52 (tt, *J* = 7.5 Hz, *J* = 1.5 Hz, 1H), 7.44 (m, 2H), 7.31 (dt, *J* = 8.0 Hz, *J* = 1.0 Hz, 1H), 7.23 (d, *J* = 2.0 Hz, 1H), 7.05 (m, 1H), 6.98 (m, 1H), 4.83 (qd, *J* = 9.5 Hz, *J* = 8.0 Hz, *J* = 5.0 Hz, 1H), 3.30 (dd, *J* = 14.5 Hz, *J* = 5.0 Hz, 1H), 3.22 (dd, *J* = 14.5 Hz, *J* = 9.5 Hz, 1H) ppm. ¹³C NMR (DMSO-*d*₆): 170.72, 166.90, 147.56, 144.85, 136.52, 134.23, 131.86, 128.63, 127.97, 127.59, 124.23, 121.41, 120.17, 118.89, 118.73, 111.83, 110.55, 54.89, 27.74 ppm. HRMS (+) Calc. for [C₁₉H₁₇N₃O₃]⁺ 335.12699, found: 336.12663 [M+H]⁺.

Synthesis of N-(3-(1H-indol-3-yl)-1-(isoxazol-4-ylamino)-1-oxopropan-2-yl)benzamide (**Neq0673**): to a suspension of (±)-2-benzamido-3-phenylpropanoic acid (216 mg, 0.70 mmol), HOBT (124 mg, 0.91 mmol) and EDC (175 mg, 0.91 mmol) in CH₂Cl₂ (8 mL) were added under argon at 0°C. After stirring of 1 hour at room temperature, the mixture was kept on ice-bath and a solution of 4-aminoisoxazole (235 mg, 2.80 mmol) in dry CH₂Cl₂ (2 mL) was added. The resulting mixture was kept overnight at RT, then the solvent was evaporated, and the residue treated with AcOEt (30 mL) and washed H₂O (2 x 20 mL) and brine (2 x 20 mL). The organic phase was dried over MgSO₄ and evaporated to give a crude residue that was purified by chromatographic column on silica using as mobile phase CHCl₃/AcOEt (1:1), to give with solid (*R*_f = 0.4) crystallized from AcOEt (36% yield).

Secondary purification was carried out on cellulose-2 Phenomenex column, in isocratic elution with a flow rate of 2.36 mL min⁻¹, at 32°C; the mobile phase

composition was *n*-hexane/ethanol (70:30) (v/v) to give Neq0673. ¹H-NMR (500 MHz, DMSO-*d*₆) δ = 10.80 (s br, 1H), 10.51 (s, 1H), 9.13 (s, 1H), 8.72 (d, *J* = 8.0 Hz, 1H), 8.63 (s, 1H), 7.84 (m, 2H), 7.69 (d, *J* = 8.0 Hz, 1H), 7.52 (tt, *J* = 7.5 Hz, *J* = 1.5 Hz, 1H), 7.44 (m, 2H), 7.31 (dt, *J* = 8.0 Hz, *J* = 1.0 Hz, 1H), 7.23 (d, *J* = 2.0 Hz, 1H), 7.05 (m, 1H), 6.98 (m, 1H), 4.83 (qd, *J* = 9.5 Hz, *J* = 8.0 Hz, *J* = 5.0 Hz, 1H), 3.30 (dd, *J* = 14.5 Hz, *J* = 5.0 Hz, 1H), 3.22 (dd, *J* = 14.5 Hz, *J* = 9.5 Hz, 1H) ppm. ¹³C NMR (DMSO-*d*₆): 170.72, 166.90, 147.56, 144.85, 136.52, 134.23, 131.86, 128.63, 127.97, 127.59, 124.23, 121.41, 120.17, 118.89, 118.73, 111.83, 110.55, 54.89, 27.74 ppm. HRMS (+) Calc. for [C₂₁H₁₈N₄O₃]⁺ 374.13789, found: 375.13895 [M+H]⁺.

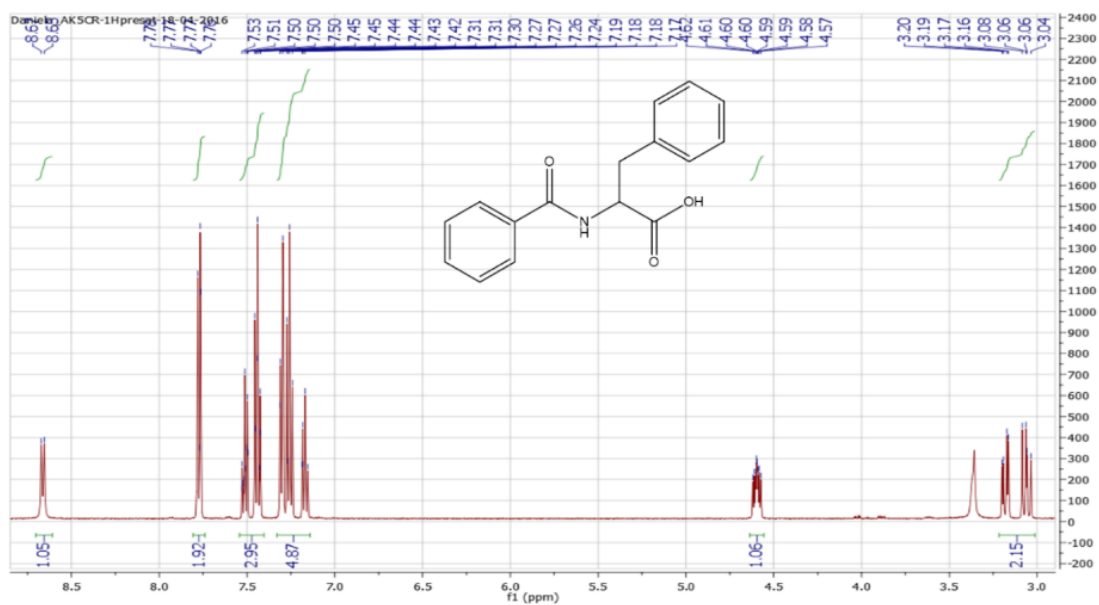


Figure S1. ¹H-NMR (400 MHz, CDCl₃) of compound 4

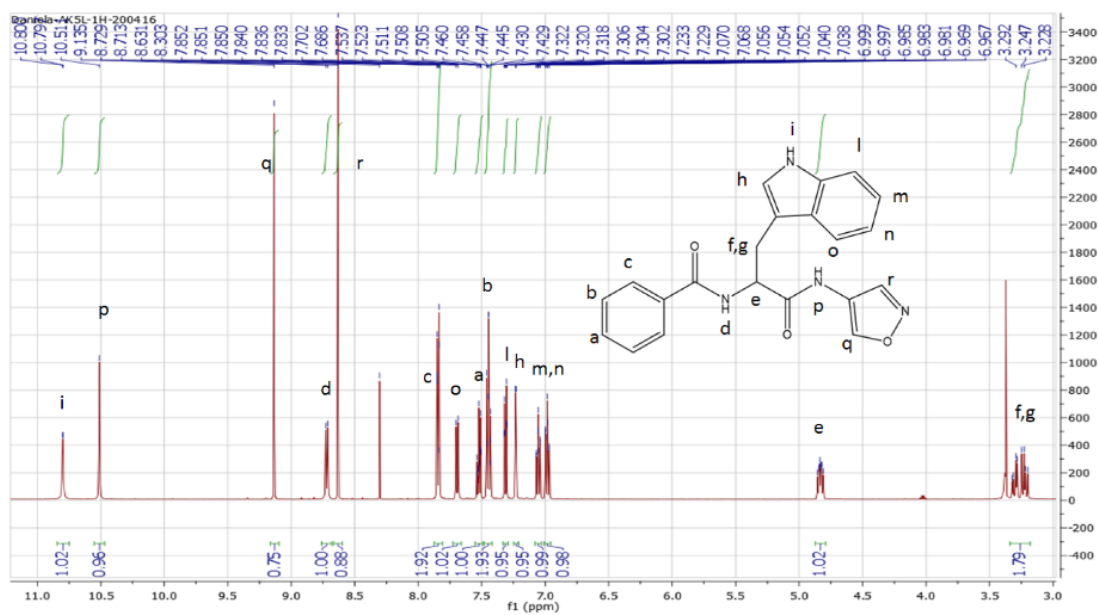


Figure S2. $^1\text{H-NMR}$ (400 MHz, $\text{DMSO-}d_6$) of compound **Neq0673**

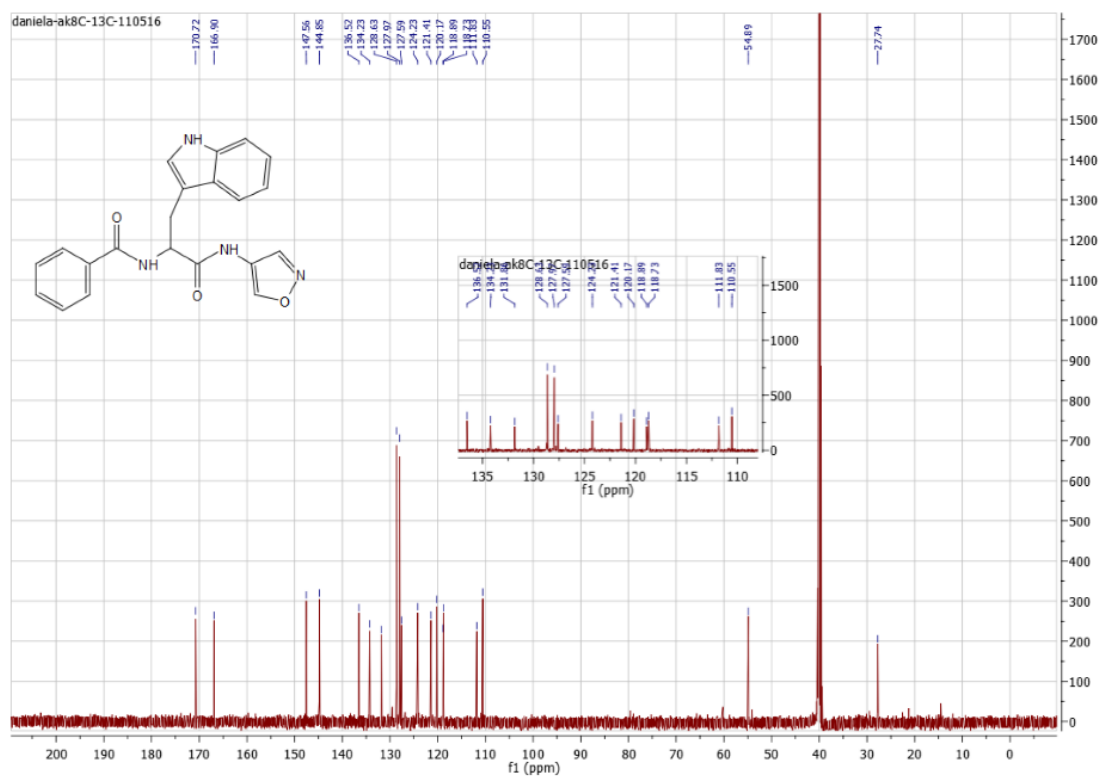


Figure S3. $^{13}\text{C-NMR}$ (100 MHz, $\text{DMSO-}d_6$) of compound **Neq0673**

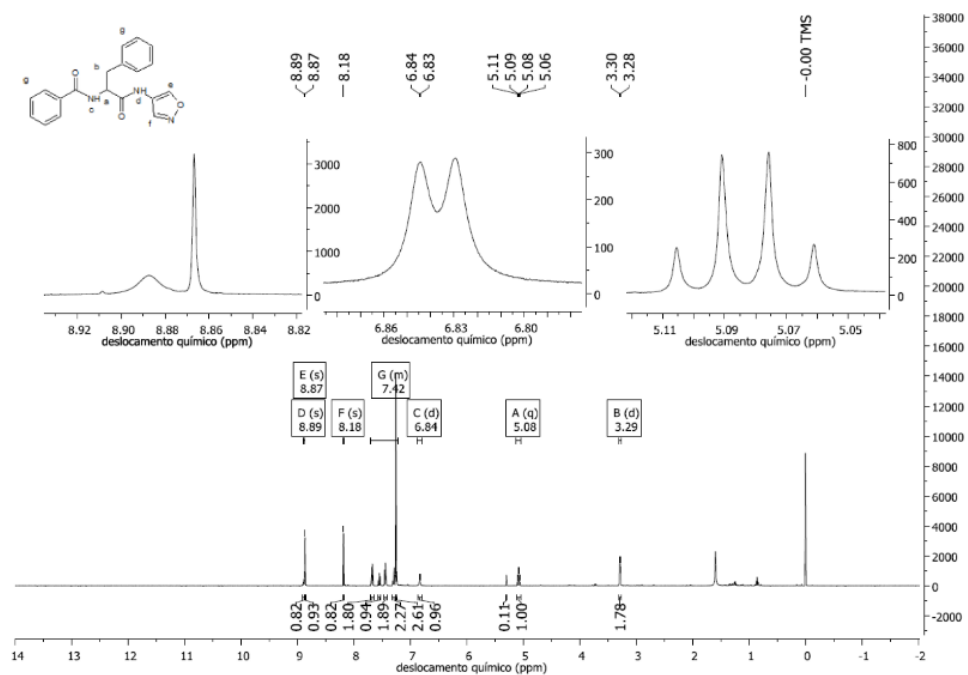


Figure S4. $^1\text{H-NMR}$ (400 MHz, $\text{DMSO-}d_6$) of compound Neq0646

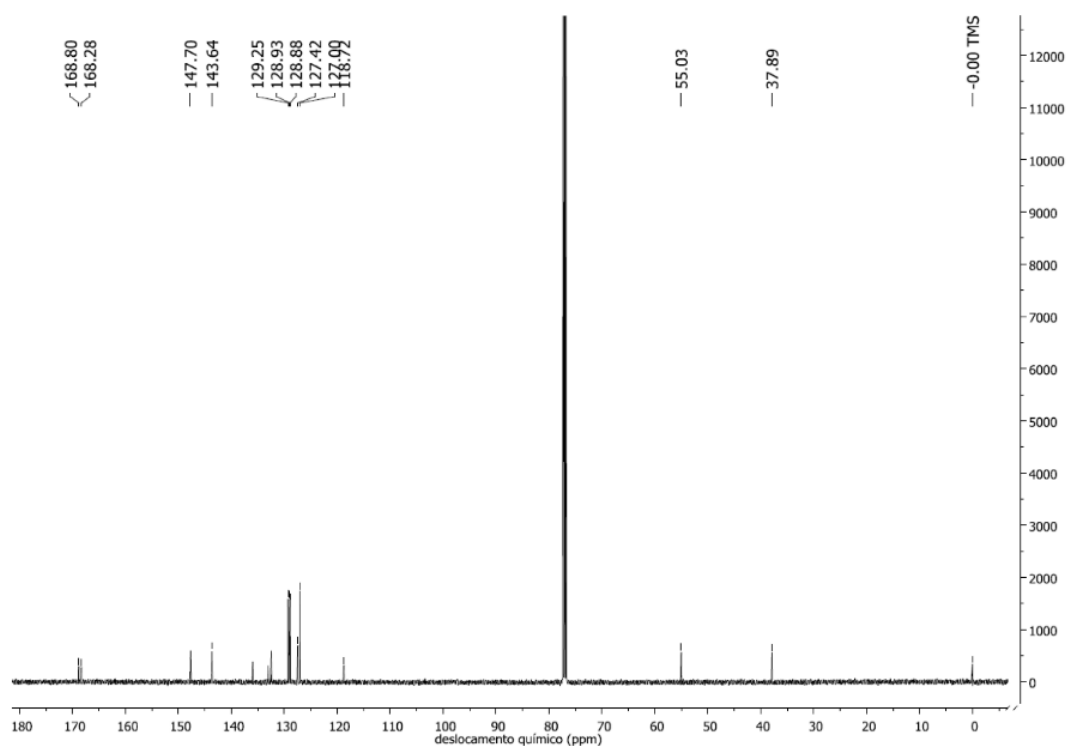


Figure S5. $^{13}\text{C-NMR}$ (100 MHz, $\text{DMSO-}d_6$) of compound Neq0646

2. Electronic parameters

Table S1. The energy (E) of neutral, anionic and cationic form, ionization potential (IP), electron affinity (EA), chemical potential (μ) and hardness (η) for compounds.

Compound (Neq)	$E_{\text{Neutral}} (E_h)$	$E_{\text{Anion}} (E_h)$	$E_{\text{Cation}} (E_h)$	IP (eV)	EA (eV)	μ (eV)	η (eV)
500	-1122.56	-1122.54	-1122.20	9.799	-0.473	-5.136	10.272
539	-1235.59	-1235.58	-1235.27	8.844	-0.462	-4.653	9.306
615	-1352.25	-1352.22	-1351.89	9.722	-0.836	-5.279	10.558
646	-1121.32	-1121.30	-1120.99	8.809	-0.353	-4.581	9.161
652	-1509.98	-1509.96	-1509.66	8.738	-0.459	-4.599	9.198
653	-1274.83	-1274.81	-1274.52	8.653	-0.636	-4.645	9.289
654	-1451.45	-1451.43	-1451.12	9.030	-0.534	-4.782	9.564
655	-1199.53	-1199.52	-1199.24	7.731	-0.302	-4.016	8.032
656	-1220.78	-1220.76	-1220.43	9.669	-0.494	-5.082	10.163
657	-1392.40	-1392.38	-1392.06	9.228	-0.684	-4.956	9.912
673	-1252.56	-1252.55	-1252.24	8.712	-0.397	-4.555	9.109
675	-1198.78	-1198.74	-1198.39	10.383	-0.830	-5.606	11.213
677	-1216.92	-1216.90	-1216.57	9.683	-0.593	-5.138	10.276
690	-989.82	-989.80	-989.46	9.701	-0.689	-5.195	10.389

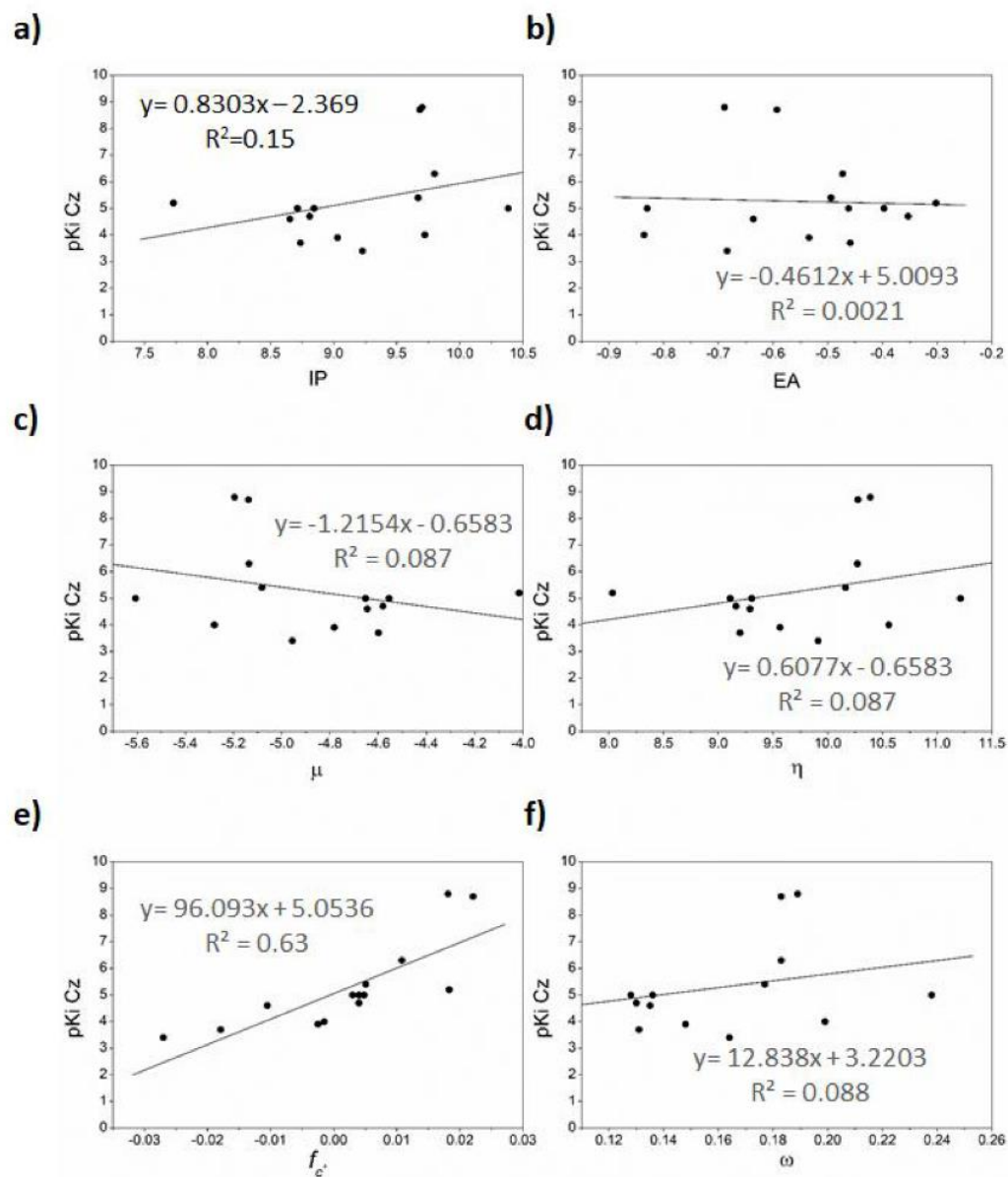


Figure S6. The putative linear equation and the coefficient of determination obtained through the linear correlation between the calculated parameter and pKi against Cz. **a)** Ionization potential **b)** electron affinity **c)** chemical potential **d)** hardness **e)** electrophilic Fukui function and **f)** global electrophilicity, all parameters values are in eV but electrophilic Fukui function.

3. Half-life measurement and decay constant by HPLC

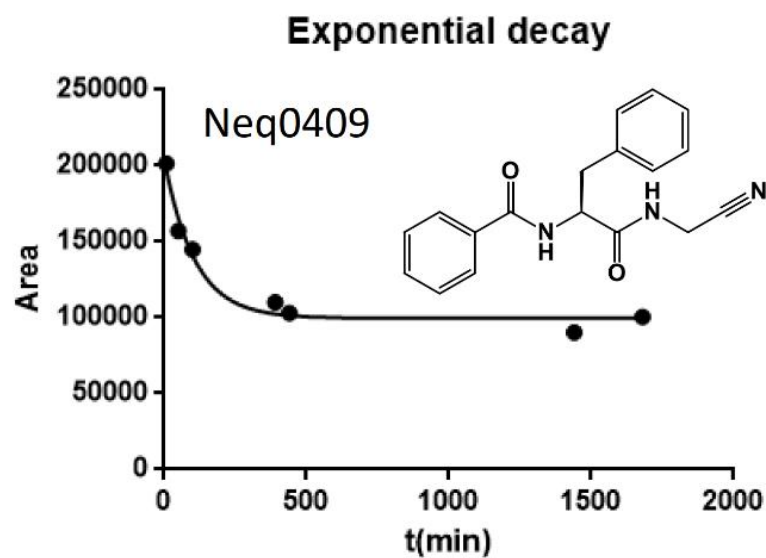


Figure S7. Decay curve of Neq0490 with cysteine.

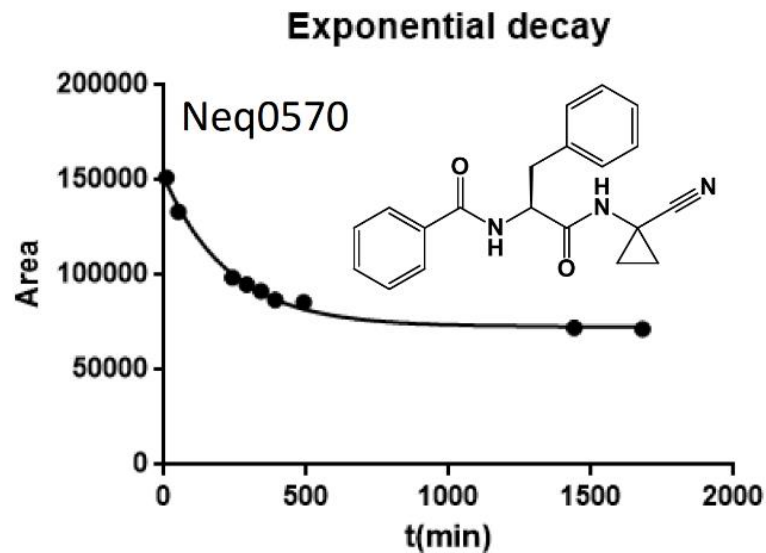


Figure S8. Decay curve of Neq0570 with cysteine.

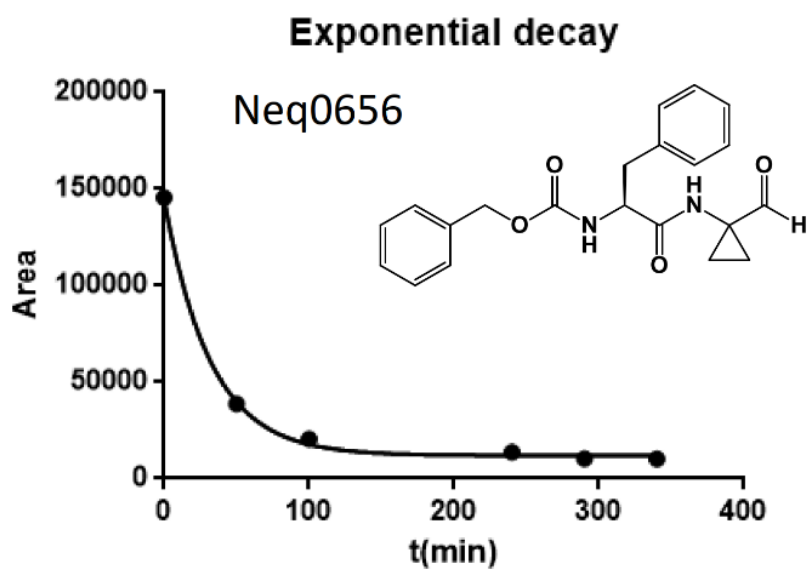


Figure S9. Decay curve of **Neq0656** with cysteine.

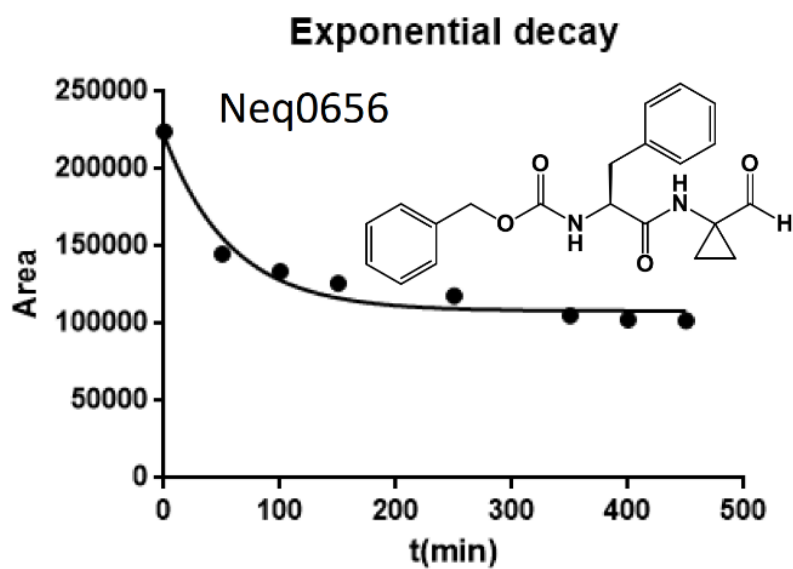


Figure S10. Decay curve of **Neq0656** with glutathione.

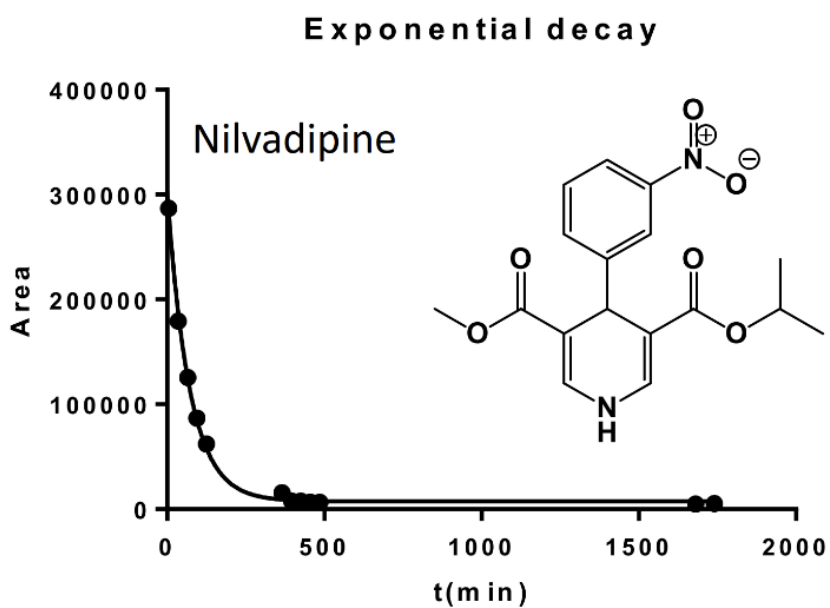


Figure S11. Decay curve of Nilvadipine with cysteine.

We also performed a blank HPLC run in order to show the stability of compounds. The decays are caused by reaction with the thiol.

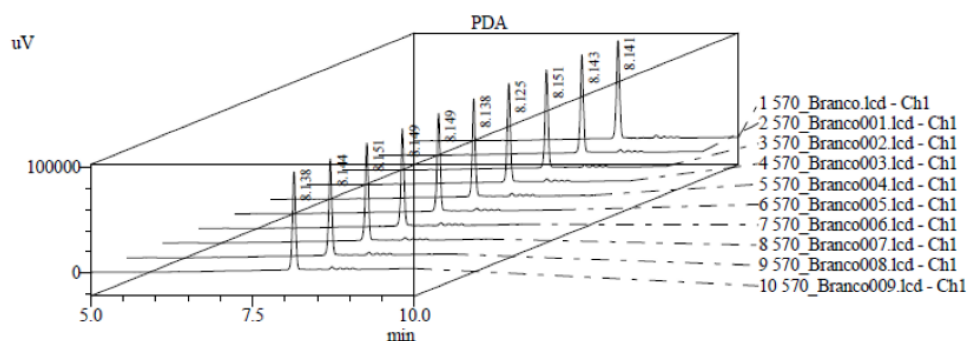


Figure S12. Blank of Neq0570.

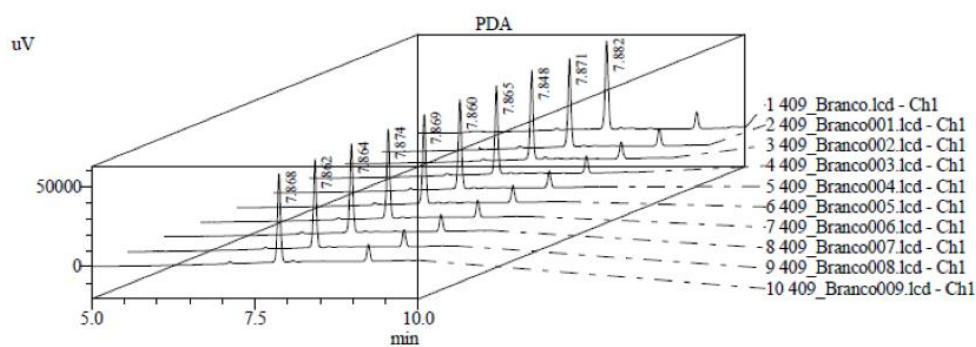


Figure S13. Blank of Neq0409.

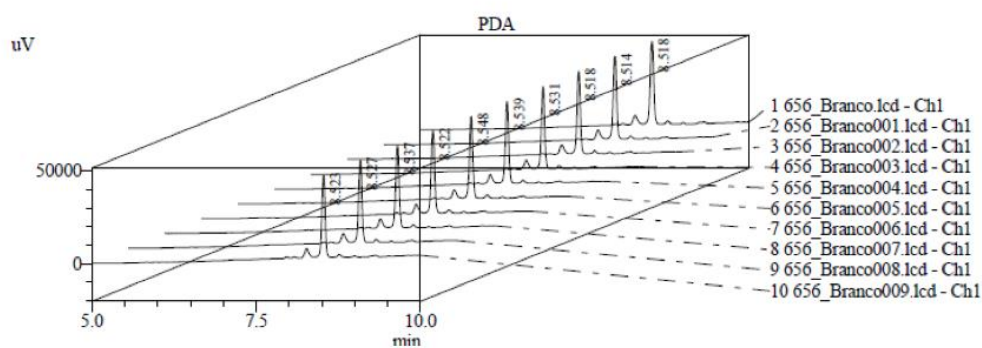


Figure S14. Blank of Neq0656.

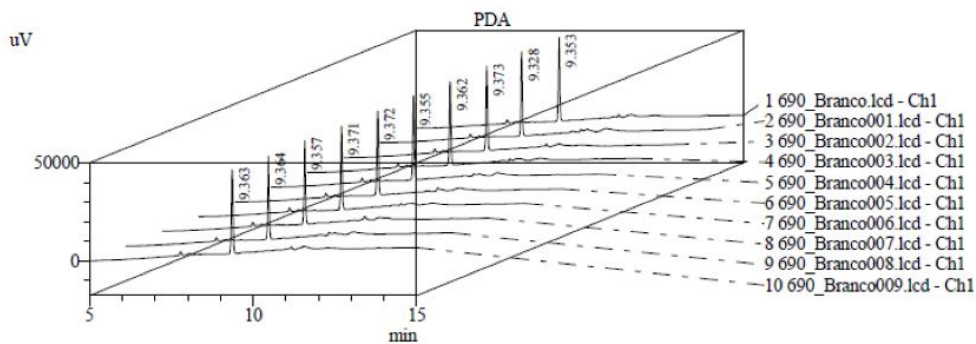


Figure S15. Blank of Neq0690.

4. Reference

- (1) Wriede, U.; Fernandez, M.; West, K. F.; Harcour, D.; Moore, H. W. Synthesis of Halodimethoxy-1,2-Benzoquinones. *J. Org. Chem.* **1987**, 52 (20), 4485–4489. <https://doi.org/10.1021/jo00229a011>.

APPENDIX B – Cartesian Coordinates for the QM-cluster Systems

Nitrile-based inhibitor coordinates for the non-covalent state:

Atoms followed by the number 1 are fixed and those followed by the number 0 are free.

C 1	0.87465100	-7.35305500	-2.23699300
C 0	1.23723700	-6.25064500	-3.23133500
C 0	0.57372200	-4.94285700	-2.82211000
N 0	1.17784600	-3.78917600	-3.18889400
O 0	-0.48182400	-4.96106100	-2.18088700
H 0	-0.20982800	-7.42651900	-2.14054500
H 0	1.29192500	-7.12188400	-1.25196800
H 0	2.32248000	-6.11906900	-3.31095800
H 0	0.87675100	-6.50644800	-4.23738600
H 0	0.69763500	-2.89157200	-2.97095500
H 0	1.94065000	-3.80848000	-3.84866900
N 1	-1.91243600	-5.42615200	-6.50788500
C 1	-1.66676900	-4.19324600	-5.74807000
C 1	-2.78814600	-3.83412300	-4.83570500
O 0	-3.57959300	-2.92412500	-5.12161700
H 0	-2.12699500	-6.18090700	-5.85849500
H 0	-0.72911200	-4.30806000	-5.19098100
H 0	-1.53781500	-3.36347300	-6.44458900
N 1	-2.94997200	-4.60499200	-3.77199900
C 1	-4.00743300	-4.63866100	-2.73679800
C 1	-3.63936800	-4.03369500	-1.39291300
O 0	-4.18818800	-4.49665500	-0.37925800
C 1	-4.44020400	-6.05548000	-2.44233800
O 1	-3.39794100	-6.92925600	-2.12365300
H 0	-2.16761400	-5.19550200	-3.50280400
H 0	-4.86630100	-4.09380300	-3.13924300
H 0	-5.21386200	-6.02068700	-1.67232000
H 0	-4.89135900	-6.46330800	-3.35447800
H 0	-3.05151400	-6.66123700	-1.25984200
C 0	-1.73538000	-3.18423700	-0.12485800
N 0	-2.61193200	-3.17013300	-1.29962300
O 0	-1.88368800	-2.96898900	2.26078400
S 0	-0.39552700	-1.31048000	-1.85540200
C 0	-2.38466200	-2.69632200	1.16805100
N 0	-2.58250400	2.06320200	-3.49990900
O 0	-4.57895500	1.00905000	-3.23979900
C 0	-0.39481100	-2.43136500	-0.39490800
N 0	0.66334300	1.10319500	-3.94687000
O 0	-5.70442400	3.77649600	-4.69479100
N 0	-5.77849700	3.42131200	-2.45308300
C 0	-3.93004300	3.19706200	-0.87904600
C 0	-4.34431000	3.27044300	-2.35301100
C 0	-3.86287100	2.00886100	-3.09385400
C 0	-1.95107600	0.92617600	-4.14378400
C 0	-0.49133100	1.03073000	-4.03355300
C 0	-6.35933300	3.64780800	-3.66399300
C 0	-7.87441300	3.76996300	-3.64341000
H 0	-4.38922300	2.33250500	-0.38871300
H 0	-2.84443900	3.12404400	-0.79648100
H 0	-4.25673700	4.09915300	-0.35517800
H 0	-8.30548800	3.55140300	-2.66475800
H 0	-8.14358900	4.78769600	-3.94131500
H 0	-8.29631600	3.08495800	-4.38181800
H 0	-3.86187900	4.13207500	-2.82409500
H 0	-6.35552200	3.13045300	-1.66920300
H 0	-2.02588600	2.90344900	-3.32250100
H 0	-2.27091000	-0.00486600	-3.66468800
H 0	-2.22263400	0.86110200	-5.20576600

H 0	-0.14634400	-1.86730300	0.50784700
H 0	0.38245600	-3.18764300	-0.53201100
H 0	-1.49149800	-4.23028100	0.08682500
H 0	-2.15218900	-2.83055300	-2.14286700
N 0	-3.50683900	-1.97750800	1.00134700
C 1	-4.42836200	-1.64004400	2.04772100
C 1	-4.93024100	-2.87051500	2.87568100
O 1	-4.96455800	-2.86779000	4.07271100
C 0	-5.61459400	-0.81287500	1.44172900
C 0	-5.95219200	-1.21516300	0.03674300
C 0	-5.26496700	-0.77295000	-1.07520400
C 0	-6.83472200	-2.25859600	-0.42585200
C 0	-6.63168900	-2.38174100	-1.83468200
C 0	-7.77569400	-3.09055200	0.20291200
N 0	-5.64913000	-1.48232000	-2.18857100
C 0	-7.35410700	-3.29244000	-2.61179600
C 0	-8.48119200	-4.00742500	-0.56722800
C 0	-8.27348200	-4.10510900	-1.95903500
H 0	-3.85329400	-1.90378700	0.05366900
H 0	-3.93523500	-1.02756200	2.80727100
H 0	-6.47907300	-0.90259400	2.10587200
H 0	-5.31769100	0.24010200	1.45415000
H 0	-4.51277200	-0.00806700	-1.16669500
H 0	-5.30896500	-1.27092100	-3.11680300
H 0	-7.93859600	-3.02107700	1.27373700
H 0	-7.21150800	-3.34911800	-3.68625100
H 0	-9.21287900	-4.65497300	-0.09354800
H 0	-8.85461600	-4.81799100	-2.53542300
N 1	-5.25108200	-4.02885100	2.22085200
C 1	-5.54185800	-5.33971500	2.78682100
C 1	-4.41404400	-5.82631800	3.74081400
O 1	-4.70158500	-6.24501500	4.89227600
C 1	-5.80134700	-6.43390200	1.70516800
H 0	-5.03533400	-4.04423000	1.22515500
H 0	-6.42109700	-5.26797900	3.43579500
H 0	-6.07886200	-7.36315500	2.20651800
H 0	-6.61334100	-6.11550100	1.04556900
H 0	-4.91194200	-6.59679100	1.09106800
N 0	-2.06515500	-3.73892600	5.16215800
C 1	-2.01948900	-2.68726700	6.19146000
C 1	-3.11908600	-3.03844200	7.19307800
O 1	-2.92506500	-2.87093800	8.42454200
C 1	-2.15928400	-1.29789300	5.64346000
O 1	-1.95001400	-0.37781600	6.75002400
H 0	-1.38879600	-3.52742300	4.43116400
H 0	-1.07543300	-2.74905900	6.74389700
H 0	-1.38233600	-1.10193700	4.89916500
H 0	-3.13329700	-1.14626000	5.16994400
H 0	-2.26375500	-0.81991600	7.55209000
C 1	-7.19040900	0.11604500	-4.86699200
C 0	-8.22010000	0.09741700	-3.71350900
O 0	-9.34150100	-0.38060100	-3.86932300
H 0	-7.70403900	0.47515900	-5.76071600
H 0	-6.31163400	0.72687900	-4.66049500
H 0	-6.87755000	-0.91527000	-5.05923400
N 0	-7.84384100	0.63200300	-2.50554200
C 1	-8.72445600	0.53905000	-1.34033700
C 0	-8.22388200	1.37125200	-0.17585200
O 0	-7.40153300	2.28472200	-0.28684500
H 0	-6.88737500	0.94190600	-2.37707400
H 0	-9.71998800	0.90184800	-1.62352000
H 0	-8.84953500	-0.50562400	-1.02889000
N 0	-8.82794000	1.08017700	1.00364600
C 1	-8.56766900	1.81837700	2.22446500

C 1	-8.75533500	0.84584600	3.42095800	H 0	0.91877000	7.49825000	-4.03943300
O 0	-9.08202800	-0.31304700	3.28934500	N 0	0.27930100	4.68149800	-0.95283200
C 1	-9.57512800	3.02079900	2.42633100	C 0	0.55766500	3.30473100	-0.61219500
C 1	-9.55736900	4.19499900	1.34249700	C 1	0.51428300	3.07471000	0.88667100
C 1	-10.73081700	5.21141700	1.44528200	O 0	0.32721100	3.92413600	1.72909300
C 1	-8.26283100	4.91993700	1.52401300	C 0	1.94354600	2.85315300	-1.19804000
H 0	-9.31981000	0.20110500	1.10903700	C 0	2.52020900	1.57570400	-0.66447000
H 1	-7.51736900	2.18981500	2.16402700	C 0	3.66989900	1.36836400	0.06439900
H 1	-10.59749800	2.60492100	2.44685200	N 0	1.96878000	0.31301200	-0.84905800
H 1	-9.39323700	3.46902600	3.41467200	C 0	2.73257600	-0.60674600	-0.25742800
H 1	-9.60660200	3.70986300	0.33997800	N 0	3.77732900	0.00820700	0.30466900
H 1	-11.69603000	4.71164500	1.33832600	H 0	0.81574200	5.39277100	-0.46656900
H 1	-10.72272300	5.72873000	2.40763100	H 0	-0.22430500	2.67615400	-1.05853300
H 1	-10.64579100	5.96263700	0.65508500	H 0	1.83779300	2.81004100	-2.28481900
H 1	-8.16883000	5.78941500	0.85640200	H 0	2.65626000	3.65826300	-1.01753200
H 1	-8.12402700	5.29857200	2.54348800	H 0	4.41086000	2.07445800	0.39961900
H 1	-7.39539600	4.27696700	1.29114500	H 0	2.54150800	-1.66842600	-0.24024300
C 1	-6.65436500	0.52696000	6.60819100	H 0	4.49339800	-0.50184700	0.84885400
C 1	-5.30478200	0.28867900	5.90444700	C 1	6.13646300	-3.50595300	4.17739400
S 1	-4.66469400	1.34706400	4.55560200	C 0	6.91209500	-3.29721900	2.86950100
C 1	-3.70689600	2.60961400	5.45264900	C 0	6.49427200	-2.07658000	2.05189900
H 0	-6.73451200	1.53507200	7.02829800	N 0	7.50906500	-1.36032300	1.53254300
H 0	-6.74767400	-0.19357100	7.42700000	O 0	5.30347500	-1.79143500	1.85461000
H 0	-4.49922800	0.24657400	6.64546500	H 0	6.34780300	-2.70492600	4.89192200
H 0	-5.33632000	-0.70556500	5.44371800	H 0	7.98836300	-3.25259800	3.06363500
H 0	-4.34555800	3.24164100	6.07446400	H 0	6.75212000	-4.15640100	2.20480900
H 0	-2.94286700	2.11957900	6.06176700	H 0	7.32464200	-0.67378100	0.81618700
H 0	-3.22234800	3.22955100	4.69552500	H 0	8.47435000	-1.61297300	1.68886600
N 0	1.28398300	6.31810300	3.19927300	C 0	3.81510500	-4.72693000	1.23648200
C 1	2.55757600	6.88146900	2.81805800	O 0	2.80616000	-4.87451400	0.22879100
C 1	2.78052400	7.08753600	1.30619100	H 0	3.54369500	-3.93523900	1.94600500
O 0	2.10883400	6.49531300	0.49165900	H 0	4.69707800	-4.37234200	0.69985500
C 1	3.77853100	5.99405600	3.37518300	H 0	1.95731400	-5.02228800	0.66682100
C 0	3.67116300	4.55434900	2.83967100	C 1	6.57990600	-6.59586500	-0.87947300
C 0	3.78633600	6.05949000	4.90207100	C 1	7.30135100	-5.28424000	-1.30178700
C 0	4.92635800	3.70763000	3.06993300	C 1	6.42214900	-4.10337700	-1.55908200
H 0	0.56454300	7.03310100	3.17493300	C 1	5.10954400	-4.12039400	-1.89879100
H 0	2.67128900	7.86635100	3.29293000	C 1	6.76954900	-2.62261500	-1.71665300
H 0	4.70418700	6.44288000	2.98502600	C 1	5.64733500	-1.89802200	-2.19567000
H 0	2.80170100	4.07391200	3.30228300	C 1	7.95338100	-1.86586100	-1.63116300
H 0	3.46411400	4.58410100	1.76345900	N 1	4.60121900	-2.82886300	-2.16500300
H 0	4.57278300	5.42987100	5.32958100	C 1	5.67417500	-0.56669000	-2.65805000
H 0	3.95504500	7.08402900	5.25136400	C 1	7.99045600	-0.53560000	-2.07320700
H 0	2.81865300	5.72435700	5.28471600	C 1	6.90781300	0.09420300	-2.66049100
H 0	4.77882500	2.68725000	2.69887600	H 0	5.85513300	-6.91265100	-1.63549100
H 0	5.79224800	4.12828200	2.54633600	H 0	7.31062600	-7.39940200	-0.75451600
H 0	5.18057600	3.62925600	4.13182400	H 0	8.03331400	-5.02671700	-0.52308600
C 1	-0.03223000	9.36103800	-0.01730300	H 0	7.90425500	-5.48753300	-2.19855600
C 1	-0.62382900	8.50906800	-1.15909400	H 0	4.40554000	-4.94073000	-1.88101400
O 1	-1.63572400	8.82856800	-1.77089500	H 0	3.65393800	-2.67040100	-2.47315700
C 1	-0.71204700	8.96824000	1.31291200	H 0	8.86709600	-2.31301000	-1.24643900
H 0	-0.24211400	10.40782600	-0.25269400	H 0	4.77493500	-0.10046300	-3.04782900
H 0	1.05078600	9.22311600	0.04534900	H 0	8.94133100	-0.01597300	-2.00447200
H 1	-0.40141300	7.95923200	1.62291000	H 0	6.97628800	1.10913800	-3.03687100
H 1	-0.44645300	9.65941300	2.11285200	H 0	6.41624000	-4.45735000	4.63721400
H 1	-1.80092100	8.96609500	1.20359700	H 0	1.27167200	-8.31973600	-2.56031200
N 0	0.11206500	7.39298500	-1.43123200	H 0	3.60991300	7.75203500	0.98693500
C 1	-0.25315200	6.51251100	-2.52563500	C 1	6.02480900	5.58599800	-1.55068300
C 0	-0.41362500	5.05123500	-2.05882600	C 1	5.49122300	4.18543400	-1.19773100
O 0	-1.08510700	4.25242200	-2.72387700	H 0	5.07987800	4.18120200	-0.17965900
C 0	0.87848400	6.49833200	-3.59539600	H 0	4.68168600	3.94461400	-1.89250400
C 0	2.25967600	6.15784300	-2.94256600	H 0	6.26629700	3.41637200	-1.24666400
O 0	2.71695200	5.00744700	-3.14082200	C 1	12.17169800	3.75657200	-1.62479200
O 0	2.74830100	7.08177700	-2.22368200	H 0	12.73820000	3.85472400	-0.69275800
H 0	1.13587700	7.45610200	-1.31561100	C 1	10.74165900	4.26716800	-1.46290200
H 0	-1.20376200	6.83729100	-2.94908900	H 0	10.77219700	5.33743800	-1.22486000
H 0	0.62876800	5.76470600	-4.36487500	H 0	10.21154500	4.18281300	-2.41649900

C 1	9.88427600	3.55380900	-0.38421900	H 0	2.27155900	-5.83751700	-1.56286100
C 1	9.91217600	3.88266000	0.97926100	H 0	2.60767000	-6.16886800	-3.26019900
H 0	10.62808600	4.63562500	1.30525300	H 0	1.16248200	-2.94589400	-3.74348100
C 1	9.10659300	3.26740300	1.91144900	H 0	2.56887000	-3.94390900	-3.92026800
H 0	9.14528900	3.53203000	2.96223600	N 1	-1.75477000	-5.28431500	-6.74945100
C 1	8.14884500	2.32953500	1.42990400	C 1	-1.51136200	-4.06612600	-5.96541000
H 0	7.47756100	1.84058700	2.12887400	C 1	-2.64425200	-3.71320300	-5.06491300
C 1	8.09799300	1.98270200	0.06573700	O 0	-3.36043900	-2.73414100	-5.29857600
H 0	7.34851200	1.27781400	-0.27613900	H 0	-2.53917900	-5.12135800	-7.37753000
C 1	8.90656500	2.60676000	-0.84603600	H 0	-0.61300300	-4.24055300	-5.35940600
H 0	8.83580100	2.38806700	-1.90613100	H 0	-1.33173300	-3.16483900	-6.56732500
C 1	13.23600600	-5.38902400	0.28567300	N 1	-2.82902700	-4.50010000	-4.01680200
H 0	13.80117200	-5.15709100	-0.62207200	C 1	-3.90311200	-4.54256000	-2.99896000
C 1	11.78219600	-4.80697200	0.19675200	C 1	-3.55083000	-3.96220700	-1.64020600
H 0	11.27483600	-5.31005500	-0.63665800	O 0	-4.12400700	-4.38310700	-0.63281300
H 0	11.23780300	-5.11887200	1.09604600	C 1	-4.35137200	-5.96024600	-2.73501100
C 1	11.66111700	-3.33914900	0.03424900	O 1	-3.32118700	-6.84763300	-2.41481000
C 1	11.90120700	-2.64287200	-1.10009700	H 0	-2.17013900	-5.26153900	-3.89193300
H 0	12.24523900	-2.94039900	-2.07846300	H 0	-4.74969900	-3.96758000	-3.39397200
N 1	11.74563600	-1.26826400	-0.77629700	H 0	-5.13033300	-5.92884500	-1.96882700
H 1	11.85535800	-0.54356700	-1.45638600	H 0	-4.79809500	-6.35291800	-3.65499200
C 1	11.30443700	-1.05801500	0.50251400	H 0	-3.02179400	-6.63369400	-1.51920900
C 1	10.86995000	0.10125700	1.16655900	C 0	-1.65191800	-2.87981500	-0.45488500
H 0	10.86926800	1.07459000	0.69055300	N 0	-2.51646200	-3.08947100	-1.61586600
C 1	10.47506100	-0.06035300	2.51137800	O 0	-1.94119900	-3.06607900	1.92091300
H 0	10.14993200	0.80724700	3.07197400	S 0	-1.59451000	-0.13720000	-1.15013900
C 1	10.45600500	-1.33787800	3.10484700	C 0	-2.40348100	-2.67061300	0.85147100
H 0	10.14379500	-1.43530700	4.14009600	N 0	-2.66451200	1.89396400	-3.12482200
C 1	10.80351400	-2.49418300	2.38511000	O 0	-4.80006200	1.11408200	-2.89402000
H 0	10.75770900	-3.46704600	2.86627600	C 0	-0.71231100	-1.68559900	-0.69808600
C 1	11.20434200	-2.37050400	1.04319700	N 0	0.07346200	-0.28321100	-3.36147200
H 0	13.20028900	-6.47557900	0.40482400	O 0	-5.58010100	3.74831700	-4.73941200
H 0	13.77423800	-4.96035600	1.13620500	N 0	-5.75289700	3.69595200	-2.46886700
H 0	6.04767200	-6.46086100	0.06303800	C 0	-4.01518700	3.49714500	-0.76987800
H 0	6.52375500	5.57354100	-2.52588800	C 0	-4.34595800	3.42175300	-2.26495300
H 0	5.18185100	6.28187000	-1.62313800	C 0	-3.97275300	2.03134100	-2.80724800
H 0	6.74680400	5.94836700	-0.80865700	C 0	-2.12910100	0.69195300	-3.73966900
H 0	-3.37220700	-5.76548700	3.37627400	C 0	-1.05340900	0.04953800	-2.88344600
H 0	-8.59808300	1.30310300	4.41505100	C 0	-6.26973100	3.79892700	-3.72353700
H 0	-7.49309400	0.36523900	5.92731500	C 0	-7.77402200	4.02277800	-3.77875900
H 0	5.06067400	-3.51150800	3.98885200	H 0	-4.56147100	2.72806600	-0.21428100
H 0	0.67207400	2.01711500	1.18389900	H 0	-2.94326100	3.36670500	-0.61839400
C 1	4.12529000	-6.03033600	1.97483700	H 0	-4.29782200	4.47842100	-0.37982900
H 0	4.35252000	-6.82977000	1.26456900	H 0	-7.96241000	5.03740500	-4.14250300
H 0	4.98180000	-5.90064200	2.64499400	H 0	-8.21182700	3.32365400	-4.49379900
H 0	3.27393000	-6.35482800	2.58406400	H 0	-8.25889700	3.89385000	-2.80855200
H 0	12.18398200	2.69729200	-1.90681200	H 0	-3.75277000	4.16599300	-2.80128200
H 0	12.69884800	4.31760200	-2.40257700	H 0	-6.38271100	3.45668500	-1.71047200
H 0	-2.96118500	-3.65117600	4.68393500	H 0	-2.06350700	2.71736000	-3.02238700
H 0	-4.06068400	-3.45729200	6.79674800	H 0	-2.95997000	-0.00538300	-3.88086800
H 0	1.06710400	-0.00417700	-1.32810800	H 0	-1.68410500	0.92139300	-4.71162800
H 0	-2.75579700	-5.29768400	-7.06436700	H 0	-0.11930000	-1.50697200	0.19861900
H 0	0.99670200	5.59460300	2.54408500	H 0	-0.01857400	-1.92776600	-1.50163700

Nitrile-based inhibitor coordinates for the covalent state:

Atoms followed by the number 1 are fixed
and those followed by the number 0 are
free.

C 1	0.93922900	-7.31342700	-2.44094600	C 0	-6.29833400	-1.40345500	0.17390400
C 0	1.79653800	-6.04407000	-2.53215300	C 0	-6.02064900	-1.28012300	-1.17143900
C 0	0.99387800	-4.79159900	-2.86010400	C 0	-7.35760800	-2.38098100	0.29149800
N 0	1.67349200	-3.78251400	-3.48543200	C 0	-7.66511000	-2.81803500	-1.02484800
O 0	-0.18406100	-4.65565100	-2.53153200	C 0	-8.07417400	-2.91793800	1.37469000
H 0	0.52581700	-7.58829600	-3.41568400	N 0	-6.81316500	-2.15976200	-1.89433800
H 0	0.09888800	-7.15027300	-1.76290800	C 0	-8.66536700	-3.75945800	-1.28490400

C 0	-9.06216100	-3.85996300	1.12171300	S 1	-4.63785600	1.31816800	4.40786300
C 0	-9.35473600	-4.27678500	-0.19444800	C 1	-3.68446100	2.55792600	5.34079500
H 0	-3.86536400	-1.74743700	-0.20641000	H 0	-6.74629200	1.48350000	6.84722900
H 0	-3.82761500	-1.04833500	2.58091700	H 0	-6.77722100	-0.24997100	7.22300200
H 0	-6.22375900	-0.54945400	2.14359600	H 0	-4.51484000	0.18229300	6.48213400
H 0	-5.14157900	0.20880800	0.98254000	H 0	-5.33964000	-0.74358100	5.25150700
H 0	-5.34692400	-0.60453100	-1.68559500	H 0	-4.32807200	3.18250700	5.96485100
H 0	-7.03820300	-2.03755200	-2.87207200	H 0	-2.93161100	2.05264800	5.95122200
H 0	-7.86113100	-2.60265700	2.38982300	H 0	-3.18657300	3.18840600	4.60136400
H 0	-8.89842400	-4.06140200	-2.30083100	N 0	1.35182600	6.34043900	3.25535800
H 0	-9.62458000	-4.28004300	1.94979700	C 1	2.65536200	6.82238200	2.87664300
H 0	-10.14057200	-5.00775000	-0.35947400	C 1	2.90221400	7.05074700	1.37173100
N 1	-5.22009700	-4.00446300	1.94787900	O 0	1.97496400	7.14155200	0.60781000
C 1	-5.52995200	-5.32215200	2.48731100	C 1	3.85947200	5.91512100	3.43721000
C 1	-4.42128800	-5.83365500	3.45068800	C 0	3.80851800	4.51827700	2.79096200
O 1	-4.73021300	-6.26898800	4.59033400	C 0	3.81191800	5.87117200	4.96302500
C 1	-5.78088100	-6.39616300	1.38364700	C 0	5.05026500	3.66023800	3.04821400
H 0	-5.03262200	-3.99090100	0.95024600	H 0	0.66960800	7.08553100	3.16665300
H 0	-6.42352000	-5.25406900	3.11688000	H 0	2.80994900	7.81406700	3.33349400
H 0	-6.04804700	-7.33795000	1.86692700	H 0	4.78717000	6.41125000	3.11541900
H 0	-6.60395700	-6.07298800	0.74000300	H 0	2.91076200	3.99777400	3.14198900
H 0	-4.89141600	-6.54017100	0.76454500	H 0	3.68371700	4.62789100	1.70561500
N 0	-2.09101900	-3.78949800	4.95913700	H 0	4.61238300	5.25083000	5.37836900
C 1	-2.05069900	-2.76414200	6.01746500	H 0	3.91665300	6.87606500	5.38761400
C 1	-3.16872100	-3.12273200	6.99582400	H 0	2.84823100	5.47061100	5.28846700
O 1	-2.99297100	-2.97719600	8.23278800	H 0	4.95497000	2.68249700	2.56339300
C 1	-2.17090600	-1.36470500	5.49085500	H 0	5.95266600	4.13764600	2.64984600
O 1	-1.97196200	-0.46475200	6.61556800	H 0	5.20865900	3.47725800	4.11612500
H 0	-1.38890600	-3.57439800	4.25372600	C 1	0.13018100	9.36873200	0.04257800
H 0	-1.11468900	-2.84732000	6.57992700	C 1	-0.45251600	8.54092600	-1.12236700
H 0	-1.37790700	-1.16163500	4.76523700	O 1	-1.45057100	8.87852400	-1.74429700
H 0	-3.13563000	-1.19736700	5.00330100	C 1	-0.57274100	8.96000000	1.35346600
H 0	-2.30760200	-0.91368100	7.40481600	H 0	-0.09000000	10.41595200	-0.18464900
C 1	-7.02358800	0.26390700	-5.07248900	H 0	1.20899900	9.23976600	0.12038200
C 0	-7.95001100	0.05616500	-3.85740400	H 1	-0.28000500	7.94330900	1.65535400
O 0	-8.83983000	-0.80092900	-3.89053200	H 1	-0.31631300	9.63502700	2.17079700
H 0	-6.59541900	1.26712900	-5.12125200	H 1	-1.66143900	8.96807000	1.22832100
H 0	-6.18873900	-0.44257400	-5.01206300	N 0	0.27177300	7.44475300	-1.43427200
H 0	-7.60784600	0.04133500	-5.96544900	C 1	-0.07123300	6.56437700	-2.51675500
N 0	-7.75618300	0.82456500	-2.73977800	C 0	-0.40247700	5.10926900	-2.01671900
C 1	-8.61244500	0.64043800	-1.56325200	O 0	-1.34608100	4.46261400	-2.50659100
C 0	-8.14642900	1.45627600	-0.37403000	C 0	1.10238600	6.48036900	-3.52955500
O 0	-7.35133000	2.39501000	-0.45930800	C 0	2.54688900	6.18326900	-2.98624200
H 0	-6.85702100	1.29287400	-2.62397000	O 0	3.46524600	6.35918400	-3.79874800
H 0	-9.63373600	0.96166600	-1.80606700	O 0	2.67365300	5.78798000	-1.76978500
H 0	-8.68502400	-0.41831100	-1.29403800	H 0	1.15517200	7.27174500	-0.97155700
N 0	-8.75558700	1.13284500	0.79330100	H 0	-0.97477500	6.93870400	-2.99594600
C 1	-8.49985800	1.85895500	2.02455100	H 0	0.87231900	5.72041100	-4.28649100
C 1	-8.71232000	0.86901400	3.20104500	H 0	1.16390700	7.43436000	-4.06220800
O 0	-9.06654000	-0.27923400	3.05191800	N 0	0.44871700	4.62832900	-1.07699700
C 1	-9.50116800	3.06639600	2.23042200	C 0	0.51257500	3.23918300	-0.66670800
C 1	-9.45619600	4.25888400	1.16664700	C 1	0.61169000	3.06255400	0.84984200
C 1	-10.62338900	5.28276700	1.26769100	O 0	0.72654600	3.94452500	1.67043200
C 1	-8.15932800	4.96993900	1.38025100	C 0	1.71925500	2.52758300	-1.40055800
H 0	-9.19006200	0.22213900	0.89619200	C 0	2.34748000	1.30977400	-0.77524900
H 1	-7.44607800	2.22490300	1.98569600	C 0	3.44534600	1.26582500	0.06229000
H 1	-10.52648600	2.65910100	2.22724900	N 0	1.90719900	0.01092900	-1.00066700
H 1	-9.33065900	3.49741800	3.22835600	C 0	2.71337100	-0.77901100	-0.30484800
H 1	-9.49369200	3.79060900	0.15527200	N 0	3.66122500	-0.06542900	0.35440500
H 1	-11.59069400	4.79270600	1.13729200	H 0	1.37090300	5.14116100	-1.10149200
H 1	-10.62644400	5.78400100	2.23854200	H 0	-0.41598100	2.73221800	-0.95731400
H 1	-10.51995600	6.04624100	0.49155800	H 0	1.37047600	2.27378200	-2.40520100
H 1	-8.04793300	5.84955500	0.72878600	H 0	2.48730200	3.29747200	-1.51745600
H 1	-8.03368900	5.33051500	2.40795000	H 0	4.07684600	2.05346300	0.44231600
H 1	-7.29340800	4.32387500	1.15032700	H 0	2.65614600	-1.85798800	-0.24131200
C 1	-6.66655900	0.48051100	6.41521500	H 0	4.37659200	-0.48628200	0.95248500
C 1	-5.30751700	0.24295300	5.72861200	C 1	6.12914000	-3.61574600	4.11811700

H 0	-3.99957900	-6.13261000	-3.84271900	H 0	-6.46968900	-5.58999500	3.07040600
O 1	-2.48487800	-6.79001100	-2.67285900	C 1	-5.69585100	-6.53229500	1.27039200
H 0	-2.26261200	-6.73466700	-1.73040200	H 0	-6.51679600	-6.20452900	0.62541700
C 1	-2.84482600	-4.11101600	-1.34426600	H 0	-4.77458300	-6.51935800	0.67860100
O 0	-3.11745000	-4.84425700	-0.37236700	H 0	-5.89474500	-7.56143700	1.58495700
N 0	-2.15434600	-2.96847800	-1.20710600	C 1	-4.44236400	-6.09950400	3.39657300
C 0	-1.45815200	-2.69809000	0.04626500	O 1	-4.71775200	-6.46506800	4.56870600
C 0	-0.50893700	-1.52014700	-0.17052400	N 0	-2.04424600	-4.27574700	4.89748700
S 0	-1.30814500	0.01748800	-0.77157200	H 0	-3.00732600	-4.23338000	4.56789000
C 0	-2.33663900	-2.44432200	1.28079100	C 1	-1.95108700	-3.40773500	6.07973400
O 0	-1.86706900	-2.62142700	2.40481000	H 0	-1.01705000	-3.60673400	6.61950200
C 0	-4.16966600	3.07115300	-0.78202200	C 1	-2.07262200	-1.90327100	5.77728500
C 0	-4.39525900	3.10127600	-2.29997600	H 0	-1.36938400	-1.58880900	4.99606900
C 0	-3.97247100	1.75424900	-2.91045300	H 0	-3.07530100	-1.64613500	5.44479600
O 0	-4.80646400	0.87935100	-3.18621000	O 1	-1.83247000	-1.23254600	7.03798100
N 0	-2.63978800	1.56501400	-3.04608200	H 0	-1.00798200	-1.59017800	7.38986500
C 0	-2.10955300	0.31449400	-3.55753400	C 1	-3.09423500	-3.72971800	7.01058800
C 0	-0.89539100	-0.28174200	-2.83210200	O 1	-2.86087900	-4.21777100	8.14703000
N 0	-5.78579600	3.38707400	-2.57386600	C 1	-7.16584500	0.50305500	-5.27685400
C 0	-6.20042000	3.68568900	-3.83345200	H 0	-6.60614500	-0.43806700	-5.31540100
O 0	-5.42833800	3.81557500	-4.78136500	H 0	-6.43800600	1.29712200	-5.10504700
C 0	-7.70082500	3.89572100	-3.97596200	C 0	-8.21922500	0.41529800	-4.18010900
O 0	-0.69332500	-1.54559400	-3.07759800	O 0	-9.40388100	0.18390200	-4.42024200
H 0	-1.78772900	-2.47367100	-2.04269700	H 0	-7.67750400	0.64841000	-6.22846500
H 0	-0.85186000	-3.56969500	0.31493400	N 0	-7.75871900	0.57894300	-2.90353200
H 0	-0.01771400	-1.30902300	0.78180500	H 0	-6.76275300	0.73649300	-2.75595300
H 0	0.22596800	-1.83504900	-0.91676900	C 1	-8.64482600	0.33999900	-1.78969700
H 0	-3.77676700	3.88233700	-2.74426200	H 0	-9.65526200	0.63475400	-2.09378800
H 0	-3.12569400	2.85106100	-0.55843600	H 0	-8.69211300	-0.72446800	-1.51980100
H 0	-4.80323000	2.31115000	-0.31341700	C 0	-8.24150300	1.15552500	-0.54592200
H 0	-2.91173500	-0.41951900	-3.54324300	O 0	-7.49930600	2.13817900	-0.61335700
H 0	-1.80535700	0.42997700	-4.60858700	N 0	-8.79909800	0.78917200	0.64865200
H 0	-6.47081300	3.04433800	-1.90662700	H 0	-9.22125200	-0.12628900	0.74572200
H 0	-2.03746800	2.37119800	-2.88850900	C 1	-8.53762800	1.51401000	1.89307800
H 0	-0.00827000	0.37590000	-2.90449000	H 1	-7.47868900	1.91308100	1.83999500
H 0	0.74542800	0.53810500	-0.63832600	C 1	-9.51773900	2.74964300	2.12674600
H 0	-4.42263300	4.04384700	-0.35223200	H 1	-10.54944500	2.39812500	1.93589900
H 0	-8.05432300	3.36526400	-4.86215100	H 1	-9.49200000	3.01714700	3.19866700
H 0	-7.89247100	4.96258900	-4.12842700	C 1	-9.21253400	3.97812200	1.27626000
H 0	-8.26057300	3.55225600	-3.10367300	H 1	-9.15031900	3.70119400	0.20072700
N 0	-3.57366400	-1.97333400	1.03762000	C 1	-10.28873400	5.09573900	1.45024900
H 0	-3.85082100	-1.89837000	0.06724300	H 1	-10.36031200	5.41394500	2.49259300
C 1	-4.61353100	-1.79573900	2.03899500	H 1	-10.02466800	5.97043600	0.84847100
H 0	-4.21386100	-1.23084200	2.88062700	H 1	-11.27199900	4.74618600	1.13012400
C 0	-5.82592900	-1.06087500	1.39131400	C 1	-7.84991000	4.49299700	1.77065900
H 0	-6.70918000	-1.23292800	2.01060600	H 1	-7.59599900	5.45598800	1.30712100
H 0	-5.60914900	0.01197900	1.42887500	H 1	-7.82973900	4.64362800	2.85218700
C 0	-6.06704700	-1.46783400	-0.03490100	H 1	-7.03657800	3.80663300	1.49423600
C 0	-5.33810300	-0.98034000	-1.10231300	C 1	-8.67132300	0.59737100	3.09666400
H 0	-4.62052300	-0.17831400	-1.13242500	O 0	-8.92598000	-0.58709200	3.02701300
N 0	-5.61832700	-1.69457000	-2.24180000	H 0	-8.55566700	1.10875700	4.06929400
H 0	-5.16328600	-1.53546000	-3.13446400	C 1	-6.52770500	0.47747900	6.09025700
C 0	-6.55284000	-2.65991500	-1.94980200	H 0	-6.32591000	1.20260100	6.88440700
C 0	-7.14286400	-3.62583500	-2.77316700	H 0	-7.15740200	-0.31780400	6.50018500
H 0	-6.88581100	-3.68898900	-3.82581900	C 1	-5.19766800	-0.12560000	5.53184000
C 0	-8.07169000	-4.48127800	-2.19390000	H 0	-4.61026600	-0.52995600	6.36390200
H 0	-8.55195400	-5.23890100	-2.80546200	H 0	-5.42604600	-0.97658800	4.88252800
C 0	-8.41572300	-4.37380400	-0.82875800	S 1	-4.16363500	0.95008200	4.55868000
H 0	-9.15598000	-5.05132900	-0.41390200	C 1	-3.41394200	1.99667300	5.78064100
C 0	-7.82682700	-3.41812500	-0.01006800	H 0	-4.14586700	2.60146600	6.32437500
H 0	-8.08603000	-3.35106300	1.04140300	H 0	-2.82567100	1.40292900	6.48812900
C 0	-6.87002500	-2.54677000	-0.55951700	H 0	-2.73800500	2.67192700	5.25059600
C 1	-4.98288300	-3.20256000	2.65663500	H 0	-7.07131100	0.99312400	5.29860900
O 1	-4.86872800	-3.35198800	3.83188100	N 0	1.18950300	5.82192000	3.40325600
N 1	-5.38799400	-4.24271800	1.84090300	H 0	0.31530500	6.32990000	3.31841800
H 0	-5.09430200	-4.21922300	0.87109600	C 1	2.31577300	6.67066500	3.06823900
C 1	-5.56703600	-5.60896400	2.45489700	H 0	2.21958900	7.62083100	3.61180300

C 1	3.66270300	5.99989900	3.53396600	H 0	1.94627900	3.45464100	-1.60239700
H 0	4.48808000	6.60764900	3.13276600	H 0	2.26652500	3.93162600	0.04431100
C 0	3.73447600	6.01364000	5.06340200	C 0	2.40432500	1.85287900	-0.21382400
H 0	4.64175800	5.52431500	5.42913600	N 0	1.75819900	0.64004600	-0.39043700
H 0	3.73325900	7.03959500	5.44720700	C 0	2.54489900	-0.36363900	0.02059000
H 0	2.86455800	5.49507100	5.47490400	H 0	2.28843000	-1.41009800	0.02443300
C 0	3.78299000	4.58036400	2.95010600	N 0	3.69801700	0.16310800	0.43632600
H 0	3.02647800	3.93778700	3.41247600	H 0	4.42583200	-0.38968400	0.93573600
H 0	3.54378000	4.60958900	1.88068900	C 0	3.63738600	1.53981600	0.30514000
C 0	5.16861500	3.95740900	3.11821100	H 0	4.45952300	2.17864500	0.57836200
H 0	5.93286600	4.53620800	2.59014700	C 0	-0.16043400	2.90958500	1.11316700
H 0	5.46236000	3.88754400	4.17052700	O 0	0.61363700	2.81073100	2.04567100
H 0	5.18724000	2.94229500	2.71067400	C 0	5.73598700	-2.88709200	4.44576200
C 1	2.44296700	7.03907800	1.56925600	H 0	4.89862100	-2.19107100	4.51923900
O 0	1.83569000	6.45870700	0.69736800	C 1	6.86405800	-2.27254500	3.61798300
C 1	5.33111800	5.81910200	-1.62451800	H 0	7.26918000	-1.39246600	4.13156400
H 0	6.02519500	6.26050200	-0.89843400	H 0	7.69277000	-2.97249700	3.48308200
C 1	4.89796500	4.41127900	-1.21141700	C 0	6.36594800	-1.78883100	2.25084600
H 0	4.11582500	4.06551500	-1.89193300	O 0	5.21865900	-1.32843900	2.12381100
H 0	5.73049500	3.70077100	-1.21507200	N 0	7.21939100	-1.88775600	1.21925200
H 0	4.47801900	4.42640900	-0.19944900	H 0	6.93756300	-1.59532200	0.29385900
H 0	4.44310900	6.45215200	-1.72460000	H 0	8.19001300	-2.14043400	1.35311100
H 0	5.83687100	5.80150800	-2.59666100	H 0	5.36634100	-3.80580800	3.98269500
C 1	11.45946500	4.54536400	-1.73422300	H 0	6.08550200	-3.12960100	5.45285600
H 0	12.03499000	4.57897600	-0.80197100	C 1	4.05441300	-5.82993800	1.92681100
C 1	10.00949500	4.86024900	-1.50530900	H 0	4.35741600	-5.49624200	2.92571900
H 0	9.93263700	5.89288400	-1.13985300	C 0	3.89628900	-4.65696300	0.97322200
H 0	9.49713600	4.86733400	-2.47640000	H 0	3.12591700	-3.96751900	1.34527200
C 1	9.19766900	3.98719200	-0.56876900	H 0	4.83074800	-4.08109300	0.91482100
C 1	9.17118700	4.27187700	0.87337500	O 0	3.57366800	-5.12852900	-0.32772000
H 0	9.71618900	5.12447400	1.26481500	H 0	2.65034500	-4.86603800	-0.54629300
C 1	8.46666400	3.38495400	1.71319300	H 0	4.80449100	-6.53348000	1.55572700
H 0	8.44656100	3.55219400	2.78587900	H 0	3.10845000	-6.37157500	2.02277400
C 1	7.72458500	2.35632600	1.15698100	C 1	6.94556900	-6.07589600	-0.59811800
H 0	7.14874000	1.71371600	1.81846200	H 0	6.45660800	-6.50949300	-1.47577800
C 1	7.71709600	2.09409500	-0.19615200	C 1	7.70792200	-4.79751900	-0.95339900
H 0	7.14747800	1.25864900	-0.59234000	H 0	8.21512600	-4.43431600	-0.04678700
C 1	8.48837700	2.84865700	-1.03711900	H 0	8.52927200	-5.05927800	-1.63571500
H 0	8.51468500	2.62083300	-2.09939900	C 1	6.98424500	-3.62405800	-1.56080800
H 0	11.59299700	3.54389500	-2.16066700	C 1	5.67856500	-3.37801100	-1.77460400
H 0	11.90352400	5.26489100	-2.42898700	H 0	4.81140600	-3.97226900	-1.50316500
C 1	-0.25980200	9.17847100	0.59644500	N 1	5.46971100	-2.10793900	-2.39691400
H 0	0.82016300	9.07195500	0.47461700	H 0	4.54325200	-1.75454200	-2.58291900
C 1	-0.71379700	8.88239500	2.07688700	C 1	6.61767000	-1.44348600	-2.59912800
H 1	-1.80904100	8.87853900	2.13255900	C 1	6.90337700	-0.32871100	-3.34027300
H 1	-0.36816600	7.89351900	2.39982800	H 0	6.13030300	0.32709600	-3.72658900
C 1	-1.01969400	8.21892000	-0.31254400	C 1	8.28819700	-0.03030700	-3.54125200
O 1	-2.33593300	8.14717200	-0.30967600	H 0	8.54952200	0.86766500	-4.09338600
H 0	-0.54485100	10.20099300	0.33148600	C 1	9.34232400	-0.93860300	-3.12304000
H 1	-0.32691800	9.63375800	2.76072700	H 0	10.37514000	-0.72666300	-3.36730000
N 0	-0.29438500	7.39662400	-1.08654800	C 1	8.96925200	-2.10764500	-2.47351100
H 0	0.72582000	7.49947800	-1.23591900	H 0	9.74601800	-2.81360700	-2.19393500
C 1	-0.91265700	6.48905700	-2.03999600	C 1	7.63040700	-2.43108700	-2.13246700
H 0	-1.93663200	6.80416700	-2.23874100	H 0	7.63983200	-6.81512900	-0.18894700
C 0	-0.01110200	6.46106700	-3.31066100	H 0	6.16785800	-5.88584100	0.14035200
H 0	-0.43771300	5.76174200	-4.03004100	C 1	13.01646100	-5.21812000	0.47160600
H 0	0.00832100	7.46975500	-3.73513800	H 0	13.88238500	-4.81836500	-0.06471800
C 0	1.43514800	6.02154200	-2.90187300	C 1	11.68939400	-4.70033800	-0.17000200
O 0	1.77510400	4.84460800	-3.17247800	H 0	11.64152000	-5.07586200	-1.19918900
O 0	2.09043800	6.89663500	-2.25732200	H 0	10.84944500	-5.17816800	0.35097500
C 0	-0.94579600	4.99804800	-1.59298300	C 1	11.49965600	-3.24244100	-0.19079800
O 0	-1.57414300	4.17766800	-2.26408800	C 1	12.01947500	-2.44497900	-1.17603100
N 0	-0.17602400	4.63026900	-0.52899600	H 0	12.54530000	-2.75694900	-2.06918300
H 0	0.39517900	5.35355300	-0.10056400	N 1	11.76696200	-1.11482700	-0.97404200
C 1	0.26475100	3.23970900	-0.30806600	H 1	11.97320100	-0.38430300	-1.62412200
H 0	-0.30738500	2.62245700	-1.00358000	C 1	11.05143500	-0.94453000	0.19431700
C 0	1.77788500	3.16262300	-0.55454700	C 1	10.54805100	0.20030700	0.83414300

H 0	10.58966500	1.18418800	0.38125500
C 1	10.12597900	0.03642300	2.19273800
H 0	9.80839600	0.92157500	2.73237200
C 1	9.94340300	-1.23547000	2.77041600
H 0	9.55832700	-1.33052100	3.77753700
C 1	10.32921800	-2.41657100	2.04762300
H 0	10.20070900	-3.39046300	2.51415400
C 1	10.97054200	-2.31913000	0.77589900
H 0	13.05704500	-6.31075200	0.43650900
H 0	13.09279000	-4.90007700	1.51575000
H 0	-1.05290000	-3.91837900	-7.48138200
H 0	-3.41097300	-6.11921900	3.00117000
H 0	-1.52027300	-3.84513300	4.13597500
H 0	-4.12009600	-3.54297800	6.64361200
H 0	3.16014200	7.84926400	1.31737100
H 0	1.13303600	5.02652000	2.77216200
H 0	-1.25598700	2.83198300	1.25785800

C 0	-0.84988100	0.46390900	-2.55857900
N 0	-5.98497200	3.36600600	-2.60912200
C 0	-6.41509900	3.68799000	-3.85550800
O 0	-5.65003500	3.85520700	-4.80397600
C 0	-7.92062700	3.86126000	-3.98840700
O 0	0.05408000	-0.51926400	-2.95548100
H 0	-2.15641600	-2.26081000	-2.01946400
H 0	-0.72605400	-3.38389600	0.15901500
H 0	-0.23537100	-0.98072500	0.71184000
H 0	0.13198000	-1.44292700	-0.95059400
H 0	-4.01012000	3.99919200	-2.72866000
H 0	-3.31090600	2.93612700	-0.59019900
H 0	-4.94471700	2.24597500	-0.38854300
H 0	-2.60555100	-0.13455300	-3.72909200
H 0	-1.43547000	0.91247400	-4.55865300
H 0	-6.65217500	3.03040800	-1.92075300
H 0	-2.21646600	2.72300000	-2.93852900
H 0	-0.31131300	1.39896300	-2.38175500
H 0	0.90062700	-0.23127700	-2.49265100
H 0	-4.70529300	4.00025500	-0.34645300
H 0	-8.27680500	3.27911200	-4.84076300
H 0	-8.13344700	4.91461900	-4.19443400
H 0	-8.46454300	3.55309000	-3.09304900
N 0	-3.54901100	-1.99561600	0.99916700
H 0	-3.83880700	-1.87134900	0.03761000
C 1	-4.56724400	-1.85901200	2.01180400
H 0	-4.16817200	-1.30940700	2.86464600
C 0	-5.80535600	-1.12524400	1.40445200
H 0	-6.67892700	-1.34211900	2.02346500
H 0	-5.61407500	-0.04988100	1.48281800
C 0	-6.04327900	-1.47782500	-0.03680300
C 0	-5.34395400	-0.90308800	-1.07999000
H 0	-4.67370500	-0.06115000	-1.07626700
N 0	-5.57325500	-1.58750600	-2.24810900
H 0	-5.14716700	-1.33337500	-3.13195200
C 0	-6.45730400	-2.61113200	-2.00542300
C 0	-6.99599200	-3.56608300	-2.87598200
H 0	-6.73359600	-3.56614300	-3.92920600
C 0	-7.88479900	-4.49097600	-2.34249500
H 0	-8.32646900	-5.24107900	-2.99116500
C 0	-8.24051400	-4.46127100	-0.97655100
H 0	-8.95144500	-5.18956200	-0.59823800
C 0	-7.70211500	-3.51681400	-0.11129900
H 0	-7.97132200	-3.50982900	0.93950900
C 0	-6.78460800	-2.57647200	-0.61258900
C 1	-4.92749000	-3.27694600	2.60890100
O 1	-4.80893500	-3.44390400	3.78149400
N 1	-5.33158900	-4.30601400	1.77862300
H 0	-5.07081200	-4.25799700	0.80188700
C 1	-5.50245400	-5.68236900	2.37227900
H 0	-6.40309000	-5.67696900	2.99098200
C 1	-5.63170500	-6.58780600	1.17395100
H 0	-6.45137900	-6.25173700	0.53205400
H 0	-4.70859000	-6.56857600	0.58441900
H 0	-5.83015400	-7.62151900	1.47294500
C 1	-4.37187000	-6.18241600	3.30178700
O 1	-4.64127500	-6.56730100	4.46922100
N 0	-1.99892500	-4.38006100	4.86304400
H 0	-2.95961000	-4.32517900	4.52865000
C 1	-1.90356300	-3.53306800	6.06047300
H 0	-0.96645700	-3.73812300	6.59185200
C 1	-2.03182300	-2.02461000	5.78234800
H 0	-1.33172100	-1.69562300	5.00381000
H 0	-3.03669300	-1.76577500	5.45758700
O 1	-1.79052300	-1.37252400	7.05235700
H 0	-0.94906800	-1.71034400	7.38311300
C 1	-3.04095200	-3.87436700	6.99125300

**Aldehyde-based inhibitor
coordinates for the covalent
state:**

**Atoms followed by the number 1 are fixed
and those followed by the number 0 are
free.**

C 1	1.48130000	-7.11619100	-2.26840200
H 0	0.44078000	-7.33100200	-2.53116900
H 0	1.56215900	-7.08087200	-1.18011100
C 0	1.93050500	-5.77685900	-2.86972200
H 0	2.97504500	-5.58796000	-2.59340500
H 0	1.88195400	-5.80564200	-3.96408000
C 0	1.13781900	-4.58645500	-2.33872900
O 0	0.78876100	-4.50849900	-1.15756100
N 0	0.86124500	-3.60274900	-3.23518700
H 0	0.43237400	-2.72877200	-2.94077500
H 0	1.23710600	-3.64394200	-4.16890800
H 0	2.10978600	-7.92904500	-2.64055300
N 1	-1.66767400	-4.05672900	-6.79532800
H 0	-2.56176600	-4.09517200	-7.27979900
C 1	-1.68507300	-2.93489600	-5.83924200
H 0	-0.67488500	-2.83645400	-5.42421000
H 0	-1.94007300	-1.96541400	-6.28706500
C 1	-2.64985900	-3.08762700	-4.68574900
O 0	-3.60960200	-2.30921400	-4.56614200
N 1	-2.39673300	-4.03507900	-3.82760600
H 0	-1.62386000	-4.66260900	-4.01735800
C 1	-3.28784100	-4.47869000	-2.76015400
H 0	-4.22822900	-3.93297800	-2.91646300
C 1	-3.58034400	-5.93807400	-2.92561700
H 0	-4.36722600	-6.22418000	-2.21776100
H 0	-3.95894700	-6.10173400	-3.93784500
O 1	-2.43471200	-6.77028800	-2.78475900
H 0	-2.18857500	-6.70256200	-1.84828000
C 1	-2.80111600	-4.11366100	-1.41366300
O 0	-2.95869400	-4.90005700	-0.46586600
N 0	-2.18000200	-2.92611800	-1.25684600
C 0	-1.44848200	-2.58986300	-0.03808500
C 0	-0.67370600	-1.27955600	-0.24115100
S 0	-1.70871800	0.08623000	-0.89449000
C 0	-2.31082600	-2.47650800	1.21979000
O 0	-1.84460000	-2.72824900	2.32796000
C 0	-4.36818700	3.07195700	-0.81495300
C 0	-4.58235900	3.15766500	-2.33308100
C 0	-4.05199700	1.87165400	-2.99387600
O 0	-4.78653000	0.90319700	-3.23648100
N 0	-2.71929300	1.87276200	-3.19800800
C 0	-1.93817100	0.71992200	-3.60230900

O 1	-2.80143500	-4.37920000	8.11878800	C 1	4.88377800	4.42803300	-1.13392800
C 1	-7.17822500	0.52975600	-5.21407000	H 0	4.11461400	4.08160800	-1.82981600
H 0	-6.58192500	-0.38892700	-5.24082200	H 0	5.72037300	3.72307100	-1.13531800
H 0	-6.48229400	1.35350000	-5.05636500	H 0	4.44935500	4.41629900	-0.12744400
C 0	-8.22967100	0.41559300	-4.11811100	H 0	4.41785200	6.47192100	-1.62939200
O 0	-9.41066200	0.16817200	-4.35797700	H 0	5.82602800	5.84212800	-2.49357100
H 0	-7.69549700	0.64150200	-6.16725200	C 1	11.44261200	4.59930100	-1.67922700
N 0	-7.76978500	0.57783900	-2.84070000	H 0	12.02130100	4.62078300	-0.74847200
H 0	-6.77540300	0.74292800	-2.70036400	C 1	9.99215800	4.90417400	-1.44002500
C 1	-8.64308800	0.30614500	-1.72433200	H 0	9.91248400	5.93091200	-1.05869500
H 0	-9.66078600	0.58408000	-2.02004600	H 0	9.47620100	4.92416600	-2.40909100
H 0	-8.66824500	-0.76293200	-1.47018100	C 1	9.18776100	4.01313600	-0.51406900
C 0	-8.25107800	1.11132100	-0.47193200	C 1	9.16551000	4.27538800	0.93240400
O 0	-7.53152200	2.11111700	-0.53330600	H 0	9.70714400	5.12524800	1.33466500
N 0	-8.78833800	0.71503700	0.72144600	C 1	8.46797900	3.37242800	1.76103200
H 0	-9.19665600	-0.20769200	0.80815300	H 0	8.44833400	3.52528500	2.83596400
C 1	-8.52697800	1.42359700	1.97573400	C 1	7.72838900	2.34931300	1.19172400
H 1	-7.47004700	1.82810800	1.92488400	H 0	7.15543600	1.69481000	1.84286800
C 1	-9.51162500	2.65106300	2.23221600	C 1	7.71693300	2.10795100	-0.16526000
H 1	-10.54250100	2.29797200	2.03981900	H 0	7.14139500	1.28175200	-0.57078700
H 1	-9.48296200	2.90205700	3.30805000	C 1	8.48168200	2.87884300	-0.99732600
C 1	-9.21503900	3.89389900	1.39972900	H 0	8.50242000	2.67012700	-2.06373900
H 1	-9.15573600	3.63391500	0.31979900	H 0	11.57871600	3.60481500	-2.12116900
C 1	-10.29545700	5.00390200	1.59503500	H 0	11.88178100	5.33103600	-2.36455600
H 1	-10.36443700	5.30562700	2.64244100	C 1	-0.28783300	9.14379400	0.76690200
H 1	-10.03752100	5.88896800	1.00588800	H 0	0.79180300	9.04277200	0.63958000
H 1	-11.27839800	4.65498600	1.27323000	C 1	-0.73487500	8.82283300	2.24428000
C 1	-7.85281000	4.40709900	1.89691200	H 1	-1.82984600	8.81325600	2.30400100
H 1	-7.60488900	5.37825900	1.44738900	H 1	-0.38365600	7.83062300	2.55055200
H 1	-7.82916500	4.54106800	2.98056000	C 1	-1.04699300	8.19505000	-0.15399000
H 1	-7.03754300	3.72870600	1.60683900	O 1	-2.36286700	8.11746000	-0.14729000
C 1	-8.65208300	0.48783500	3.16551400	H 0	-0.57837400	10.16914400	0.51881900
O 0	-8.89449900	-0.69801200	3.07651300	H 1	-0.34864700	9.56523200	2.93821600
H 0	-8.54077100	0.98478100	4.14590200	N 0	-0.31963700	7.40666800	-0.96101500
C 1	-6.49654600	0.33114600	6.14877500	H 0	0.70285700	7.50271000	-1.12004200
H 0	-6.29472700	1.04460500	6.95329700	C 1	-0.93894900	6.49252700	-1.90829200
H 0	-7.12115200	-0.47314600	6.54875200	H 0	-1.97250700	6.79308400	-2.07883300
C 1	-5.16602000	-0.25738400	5.57610000	C 0	-0.08109100	6.48441800	-3.20570300
H 0	-4.57389200	-0.67245300	6.39944600	H 0	-0.52572000	5.79012000	-3.92027700
H 0	-5.39314600	-1.09829400	4.91307100	H 0	-0.09165000	7.49726800	-3.62204500
S 1	-4.14045900	0.83787100	4.61585800	C 0	1.38728600	6.05607700	-2.86883500
C 1	-3.39064500	1.86871800	5.85106800	O 0	1.76866300	4.93954800	-3.27961400
H 0	-4.12413600	2.45834700	6.40890500	O 0	2.02084600	6.88626200	-2.14140500
H 0	-2.79413200	1.26718200	6.54483000	C 0	-0.95991500	5.01465700	-1.45321800
H 0	-2.72342300	2.55894500	5.32940200	O 0	-1.70696200	4.21371600	-2.02625800
H 0	-7.04536500	0.85678900	5.36732400	N 0	-0.07276300	4.62899100	-0.50245900
N 0	1.18662800	5.76390600	3.53825900	H 0	0.57941200	5.32756800	-0.15571500
H 0	0.30788000	6.22030300	3.31547200	C 1	0.25925900	3.22215200	-0.23148200
C 1	2.30811400	6.60946400	3.18985400	H 0	-0.54337600	2.62436400	-0.66948100
H 0	2.22062000	7.55942700	3.73883000	C 0	1.64143100	2.86214900	-0.82113200
C 1	3.65988500	5.93760800	3.64006700	H 0	1.59810100	3.05354700	-1.89777500
H 0	4.47490400	6.56629000	3.25037900	H 0	2.34891100	3.60008600	-0.43993900
C 0	3.74600400	5.90984600	5.16815500	C 0	2.20285200	1.49623700	-0.51556100
H 0	4.65111500	5.39598100	5.50641500	N 0	2.00435800	0.34403900	-1.27554700
H 0	3.76329000	6.92397900	5.58199100	C 0	2.71637800	-0.61725900	-0.69170200
H 0	2.87428900	5.39145500	5.57569400	H 0	2.73403900	-1.66196100	-0.97383500
C 0	3.80051700	4.53619600	3.01771000	N 0	3.37609600	-0.14893100	0.39502900
H 0	3.13877900	3.84157500	3.54560800	H 0	3.98934200	-0.69009800	1.01727700
H 0	3.44450300	4.55575100	1.98185700	C 0	3.05947500	1.18353300	0.51896200
C 0	5.23212300	4.00341100	3.02723600	H 0	3.43609900	1.79518400	1.32145400
H 0	5.89712800	4.64296100	2.43866800	C 0	0.13908600	3.10125700	1.27313900
H 0	5.63946000	3.93359700	4.04208400	O 0	1.07169600	2.97986400	2.04338500
H 0	5.27749300	3.00442700	2.58457900	C 0	5.88785100	-3.03992800	4.47597000
C 1	2.42802100	7.00157400	1.69632100	H 0	5.01248900	-2.40986700	4.64244100
O 0	1.81445800	6.44424100	0.81558900	C 1	6.89740500	-2.32087000	3.58390200
C 1	5.30921700	5.84401100	-1.52675900	H 0	7.28983100	-1.43391600	4.09597500
H 0	5.99172800	6.28486400	-0.78879500	H 0	7.75693300	-2.95636400	3.35518300

C 0	6.25908900	-1.83191300	2.27892200
O 0	5.05079800	-1.56895500	2.22180200
N 0	7.08364700	-1.70623400	1.21936900
H 0	6.71601600	-1.40218000	0.32917200
H 0	8.08078800	-1.85624900	1.29619000
H 0	5.54280200	-3.96704100	4.00965500
H 0	6.33523200	-3.29032000	5.44184800
C 1	4.09717100	-5.86417300	1.84875200
H 0	4.64641000	-5.74533500	2.78913000
C 0	3.98315500	-4.52879000	1.11569200
H 0	3.44126200	-3.79646700	1.72640200
H 0	4.97808000	-4.10785600	0.94071900
O 0	3.39076100	-4.66035800	-0.17529000
H 0	2.42104200	-4.71570000	-0.11804200
H 0	4.61771300	-6.60229500	1.23249200
H 0	3.10631400	-6.26490100	2.08810800
C 1	6.97953200	-6.05820000	-0.69109400
H 0	6.49289500	-6.48180100	-1.57569200
C 1	7.73505400	-4.77102000	-1.02885400
H 0	8.24205000	-4.41784000	-0.11828100
H 0	8.55605800	-5.02038700	-1.71607900
C 1	7.00387100	-3.59157800	-1.61555800
C 1	5.69639300	-3.34810900	-1.82069300
H 0	4.82552600	-3.92654800	-1.53867700
N 1	5.47943500	-2.06940400	-2.42229000
H 0	4.55349600	-1.67643000	-2.51230800
C 1	6.62380800	-1.39689800	-2.61863800
C 1	6.90178000	-0.26950000	-3.34344000
H 0	6.12385800	0.39175400	-3.70993800
C 1	8.28450300	0.03812400	-3.54497400
H 0	8.53969600	0.94731300	-4.08129900
C 1	9.34419100	-0.87185600	-3.14483700
H 0	10.37548100	-0.65095200	-3.38775200
C 1	8.97872200	-2.05244300	-2.51208300
H 0	9.76040100	-2.75811800	-2.24537600
C 1	7.64260800	-2.38709200	-2.17115200
H 0	7.67588900	-6.80180600	-0.29392100
H 0	6.20477800	-5.88314600	0.05637700
C 1	13.05073400	-5.19013800	0.36927200
H 0	13.91271200	-4.77821200	-0.16422500
C 1	11.71897300	-4.66837900	-0.25925400
H 0	11.66903900	-5.02846000	-1.29396800
H 0	10.88329300	-5.15794600	0.25784000
C 1	11.52276900	-3.21119500	-0.25675500
C 1	12.03533100	-2.39628800	-1.23146300
H 0	12.55825600	-2.69196000	-2.13189200
N 1	11.77777300	-1.07055000	-1.00795600
H 1	11.97833300	-0.32915100	-1.64740800
C 1	11.06596500	-0.92152700	0.16557200
C 1	10.55999700	0.21103400	0.82492900
H 0	10.59257200	1.20181800	0.38664600
C 1	10.14392600	0.02428900	2.18244500
H 0	9.82241300	0.89975400	2.73534500
C 1	9.96920800	-1.25724800	2.74104900
H 0	9.58818000	-1.36910100	3.74787200
C 1	10.35731800	-2.42522300	1.99853900
H 0	10.22776500	-3.40702700	2.44733300
C 1	10.99333900	-2.30530700	0.72610900
H 0	13.09648700	-6.28204200	0.31773400
H 0	13.12962400	-4.88724200	1.41772300
H 0	-0.96811000	-3.86407600	-7.50682900
H 0	-3.34197100	-6.19012500	2.90244000
H 0	-1.46445500	-3.94450400	4.11197900
H 0	-4.06944700	-3.68436200	6.63262600
H 0	3.14184800	7.82132000	1.46208000
H 0	1.21071300	4.89475000	3.01210400
H 0	-0.90014900	3.21238900	1.64856600

APPENDIX C – Supporting Information for Predicting the Relative Binding Affinity for Reversible Covalent Inhibitors by Free Energy Perturbation Calculations

Reprinted with permission from BONATTO, V.; SHAMIM, A.; ROCHO, F. R.; LEITÃO, A.; LUQUE, F. J.; LAMEIRA, J.; MONTANARI, C. A. Predicting the Relative Binding Affinity for Reversible Covalent Inhibitors by Free Energy Perturbation Calculations. **Journal of Chemical Information and Modeling**, Washington, v. 61, p. 4733-4744, 2021. Copyright 2023 American Chemical Society.

Cite this: <https://pubs.acs.org/doi/abs/10.1021/acs.jcim.1c00515>

SUPPORTING INFORMATION FOR:

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

Vinícius Bonatto,^a Anwar Shamim,^a Fernanda dos R. Rocho,^a Andrei Leitão,^a F. Javier Luque,^b Jerônimo Lameira,^{a,c*} Carlos A. Montanari^{a*}

^a Medicinal & Biological Chemistry Group, Institute of Chemistry of São Carlos, University of São Paulo, Avenue Trabalhador Sancarlense, 400, 23566-590, São Carlos/SP, Brazil

^b Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Science, Institute of Biomedicine (IBUB) and Institute of Theoretical and Computational Chemistry (IQTCUB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

^c Institute of Biological Science. Federal University of Pará. Rua Augusto Correa S/N. Belém/PA, Brazil.

*Corresponding Authors: J. L., lameira@ufpa.br; C.A.M., carlos.montanari@usp.br

Contents

1. Experimental Section	S3
1.1. Chemistry	S3
1.2. Spectral data for compound 1a . (Fig. S1-S4)	S5
1.3. Spectral data for compound 1b . (Fig. S5-S8)	S7
1.4. Spectral data for compound 1d . (Fig. S9-S12)	S9
1.5. Determination of the inhibition constant (K_i) against hCatL. (Fig. S13)	S11
2. Computational Results	S13
Figure S14: Representation of the superposed arrangements of 1a (carbon atoms as orange sticks) and 2a (carbon atoms as pink sticks) in the covalent complex formed in the binding pocket with hCatL (protein surface shown in blue).	S13
Figure S15: RMSD fluctuations (Å) along 100ns MD simulations in the noncovalent complexes a) 1a -hCatL b) 2a -hCatL, and covalent complexes c) 1a -CatL d) 2a -CatL.	S14
Figure S16: Investigated interactions between the dipeptidyl nitrile-based inhibitor and hCatL in a) noncovalent and b) covalent complexes.	S14
Figure S17: Distance fluctuations of selected noncovalent interactions (Å) in the state analyzed over 100 ns MD simulations between a) SCys25, c) HGly68, e) OGly68 and g) OAsp162 with compound 1a , and b) SCys25, d) HGly68, f) OGly68 and h) OAsp162 with 2a . The interactions are depicted in Figure S16.	S15

Figure S18: Distance fluctuations of selected noncovalent interactions (Å) in the covalent state analyzed over a 100ns MD simulations between a) HGly68, c) OGly68 and e) OAsp162 with compound **1a**, and b) HGly68, d) OGly68, and f) OAsp162 with **2a**. The interactions are depicted in Figure S16. S16

Table S1: Outcome of all FEP transformation steps of the ligand-hCatL noncovalent complex. All values in kcal/mol. The standard deviation is represented in parenthesis. S17

Table S2: Outcome of all FEP transformation steps of the ligand-hCatL covalent complex. All values in kcal/mol. The standard deviation is represented in parenthesis. S17

Table S3: Parameters used to include the Extra-Point (EP) in MD simulations and alchemical free energy calculations. S18

Table S4: Outcome of all FEP transformation steps performed for the relative hydration free energies of compounds **2a**, **2b** and **2c**. All values in kcal/mol. The standard deviation is represented in parenthesis. S18

Table S5: Outcome of all FEP transformation steps performed for the relative hydration free energies of pyridine, benzene and pyrimidine. All values in kcal/mol. The standard deviation is represented in parenthesis. S18

3. Analysis of the decharge, vdW and recharge steps in alchemical transformations.

(Figures S19-S34)

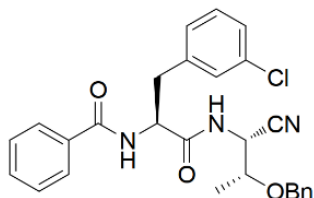
S219-S34

1. Experimental Section

1.1. Chemistry

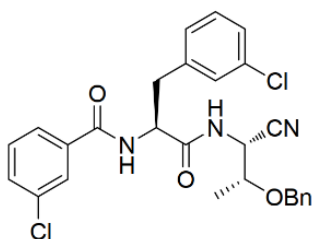
Compound **1a**;

N-((S)-1-(((1R,2R)-2-(benzyloxy)-1-cyanopropyl)amino)-3-(3-chlorophenyl)-1-oxopropan-2-yl)benzamide



Global yield = 27.95%; White solid; m.p. = 137.9-138.8 °C; ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.00 (d, *J* = 8.3 Hz, 1H), 8.68 (d, *J* = 8.3 Hz, 1H), 7.82 – 7.72 (m, 2H), 7.52 – 7.47 (m, 2H), 7.43 (t, *J* = 7.4 Hz, 2H), 7.39 – 7.18 (m, 7H), 5.07 (dd, *J* = 8.3, 4.1 Hz, 1H), 4.82 – 4.72 (m, 1H), 4.60 (q, *J* = 11.9 Hz, 2H), 3.85 (dd, *J* = 6.2, 4.1 Hz, 1H), 3.12 – 2.94 (m, 2H), 1.18 (d, *J* = 6.2 Hz, 3H); ¹³C NMR (126 MHz, DMSO-*d*₆) δ 172.12, 166.89, 141.11, 138.39, 134.19, 133.12, 131.86, 130.31, 129.52, 128.69 (2C), 128.65 (2C), 128.45, 128.17 (2C), 128.05, 127.84 (2C), 126.80, 118.42, 73.91, 70.91, 54.90, 45.23, 36.77, 16.03; Purity (HPLC) = > 99%; HRMS: *m/z* calcd for C₂₇H₂₇ClN₃O₃ [M+H]⁺: 476.1741; found: 476.1698.

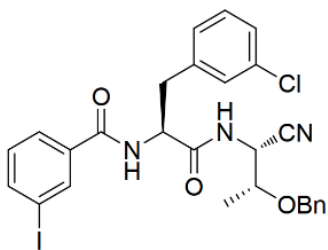
Compound **1b**;



N-((S)-1-(((1R,2R)-2-(benzyloxy)-1-cyanopropyl)amino)-3-(3-chlorophenyl)-1-oxopropan-2-yl)-3-chlorobenzamide

Global yield = 27.76%; White solid; m.p. = 117.8-118.6 °C; ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.03 (d, *J* = 8.3 Hz, 1H), 8.84 (d, *J* = 8.2 Hz, 1H), 7.82 (t, *J* = 1.8 Hz, 1H), 7.72 (d, *J* = 7.8 Hz, 1H), 7.58 (ddd, *J* = 8.0, 2.1, 1.0 Hz, 1H), 7.47 (t, *J* = 7.9 Hz, 2H), 7.41 – 7.19 (m, 7H), 5.07 (dd, *J* = 8.3, 4.2 Hz, 1H), 4.76 (ddd, *J* = 10.7, 8.2, 4.4 Hz, 1H), 4.60 (q, *J* = 11.9 Hz, 2H), 3.85 (dd, *J* = 6.2, 4.2 Hz, 1H), 3.14 – 2.91 (m, 2H), 1.18 (d, *J* = 6.2 Hz, 3H); ¹³C NMR (126 MHz, DMSO-*d*₆) δ 172.12, 166.89, 141.11, 138.39, 134.19, 133.12, 131.86, 130.31, 129.52, 128.69 (2C), 128.65 (2C), 128.45, 128.17 (2C), 128.05, 127.84 (2C), 126.80, 118.42, 73.91, 70.91, 54.90, 45.23, 36.77, 16.03; Purity (HPLC) = > 99%; HRMS: *m/z* calcd for C₂₇H₂₆Cl₂N₃O₃ [M+H]⁺: 510.1351; found: 510.1303.

Compound **1d**;



N-((S)-1-(((1R,2R)-2-(benzyloxy)-1-cyanopropyl)amino)-3-(3-chlorophenyl)-1-oxopropan-2-yl)-3-iodobenzamide

Global yield = 27.10%; White solid; m.p. = 135.9-136.5 °C; ^1H NMR (500 MHz, DMSO-*d*6) δ 9.01 (d, J = 8.3 Hz, 1H), 8.81 (d, J = 8.3 Hz, 1H), 8.12 (t, J = 1.6 Hz, 1H), 7.87 (ddd, J = 7.8, 1.7, 1.0 Hz, 1H), 7.78 – 7.74 (m, 1H), 7.46 (s, 1H), 7.37 – 7.21 (m, 7H), 5.06 (dd, J = 8.3, 4.2 Hz, 1H), 4.74 (ddd, J = 10.7, 8.2, 4.4 Hz, 1H), 4.60 (dd, J = 27.7, 11.9 Hz, 2H), 3.88 – 3.81 (m, 1H), 3.02 (ddd, J = 24.4, 13.6, 7.6 Hz, 2H), 1.18 (d, J = 6.2 Hz, 3H); ^{13}C NMR (101 MHz, DMSO-*d*6) δ 171.91, 165.41, 140.95, 140.39, 138.37, 136.26, 136.17, 133.14, 130.90, 130.33, 129.52, 128.69 (2C), 128.39, 128.17 (2C), 128.05, 127.31, 126.85, 118.37, 95.01, 73.90, 70.91, 54.91, 45.23, 36.76, 16.05; Purity (HPLC) = > 99%; HRMS: m/z calcd for $\text{C}_{27}\text{H}_{26}\text{ClIN}_3\text{O}_3$ $[\text{M}+\text{H}]^+$: 602.0707; found: 602.0700.

1.2. Spectral data for compound 1a.

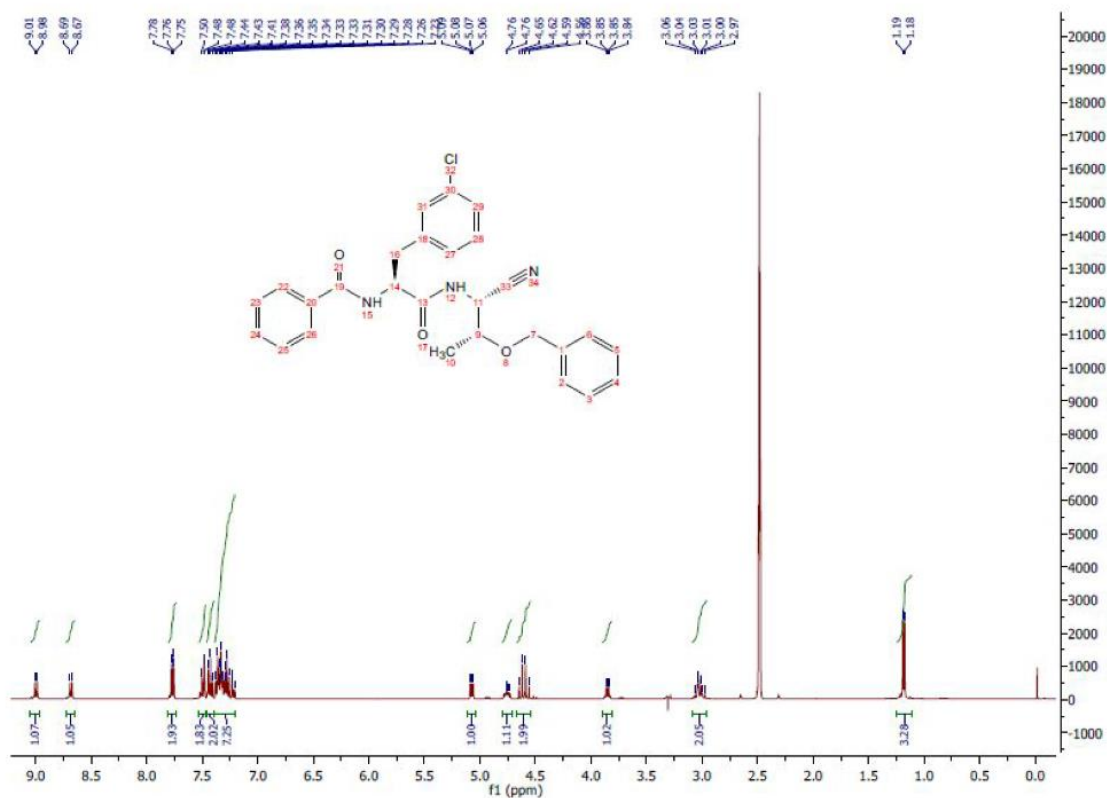


Figure S2: ^1H NMR (400 MHz, DMSO- d_6) spectrum of 1a.

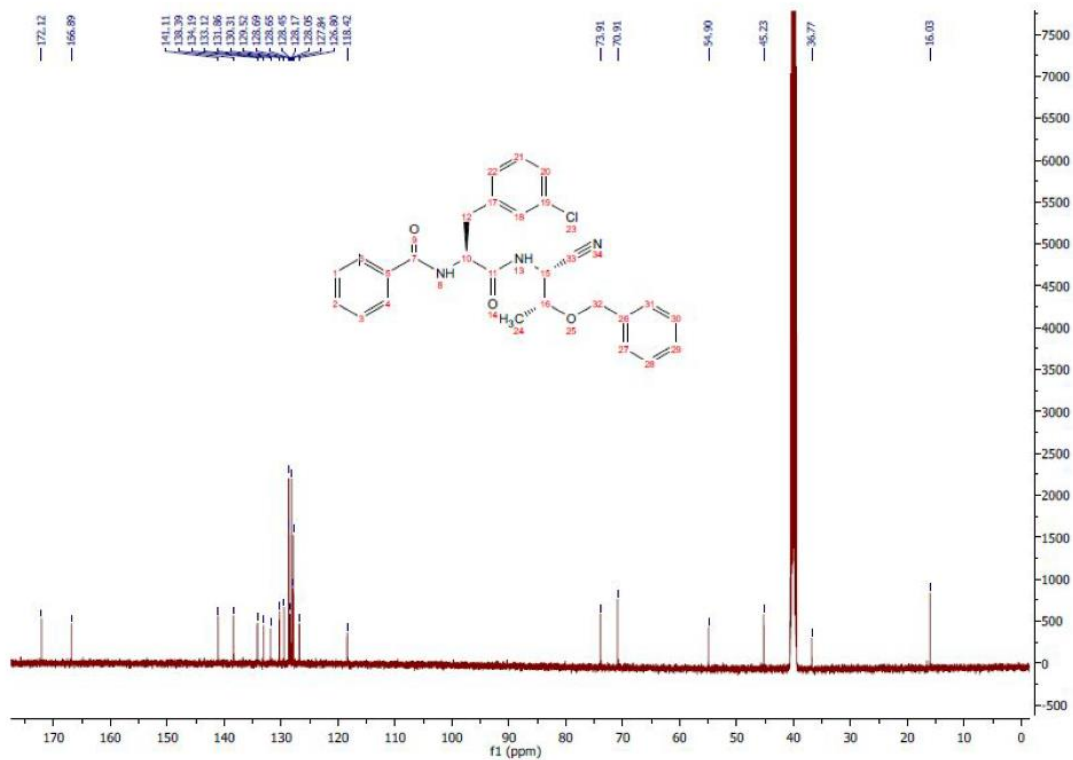
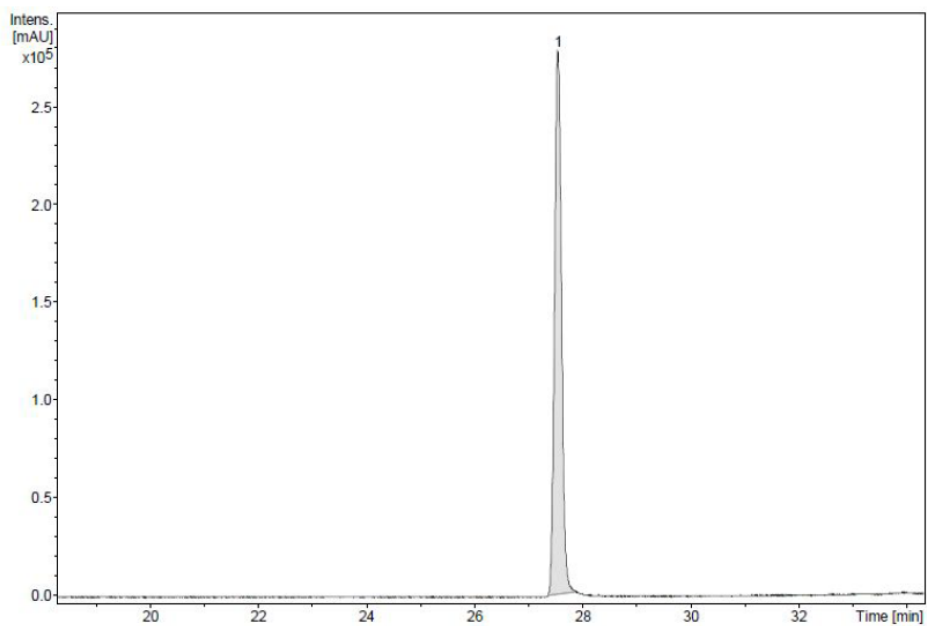


Figure S3: ^{13}C NMR (126 MHz, DMSO- d_6) spectrum of 1a.



#	RT [min]	Chromatogram	Area	Area %	S/N
1	27.6	UV Chromatogram, 190-800 nm	2546207	100.00	936.6

Figure S4: HPLC chromatogram of **1a**.

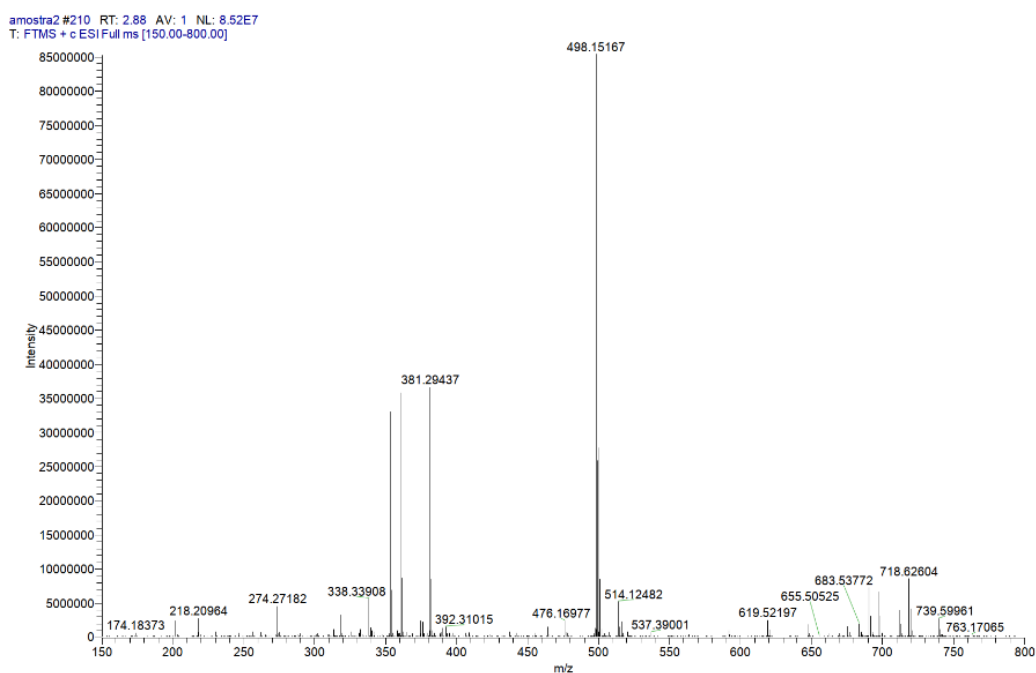


Figure S5: HRMSESI spectrum of **1a**.

1.3. Spectral data for compound 1b.

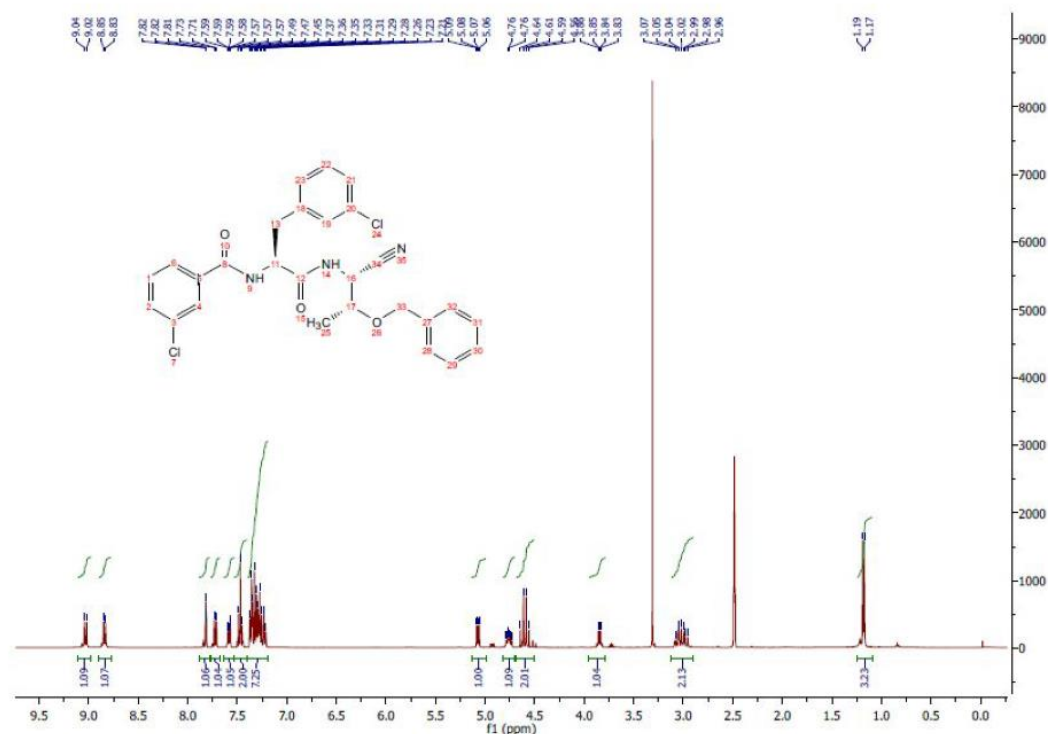


Figure S6: ¹H NMR (400 MHz, DMSO-d₆) spectrum of 1b.

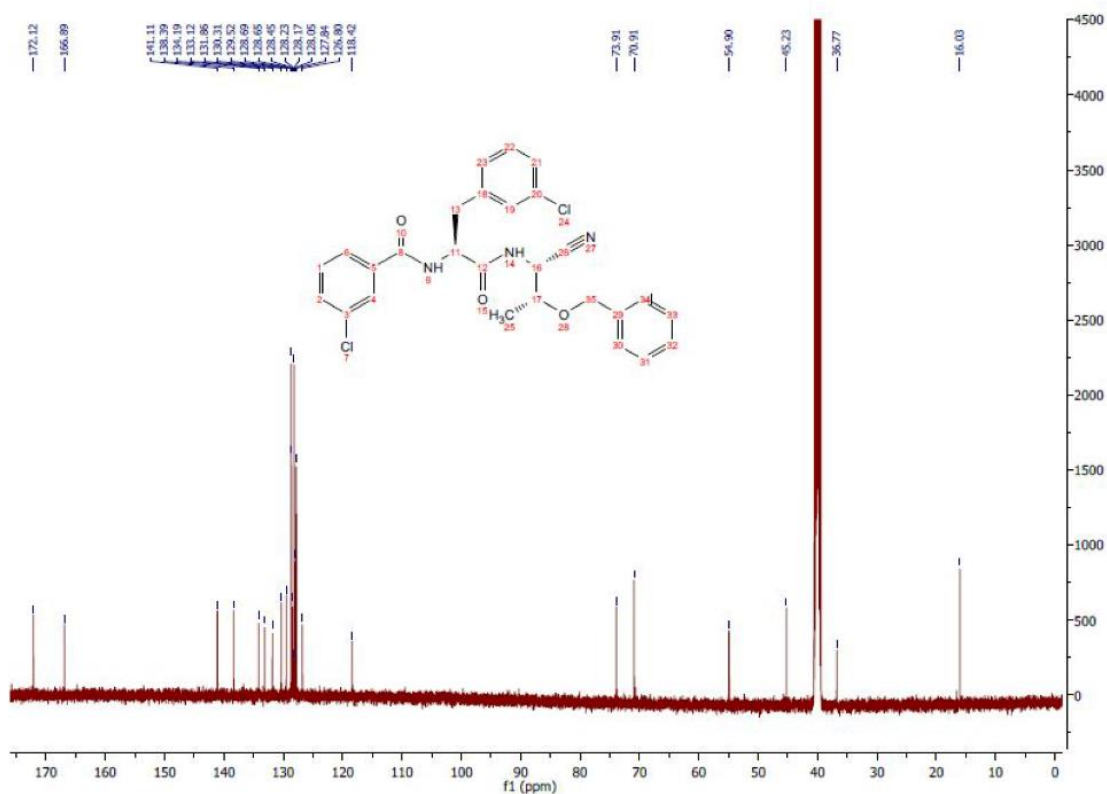
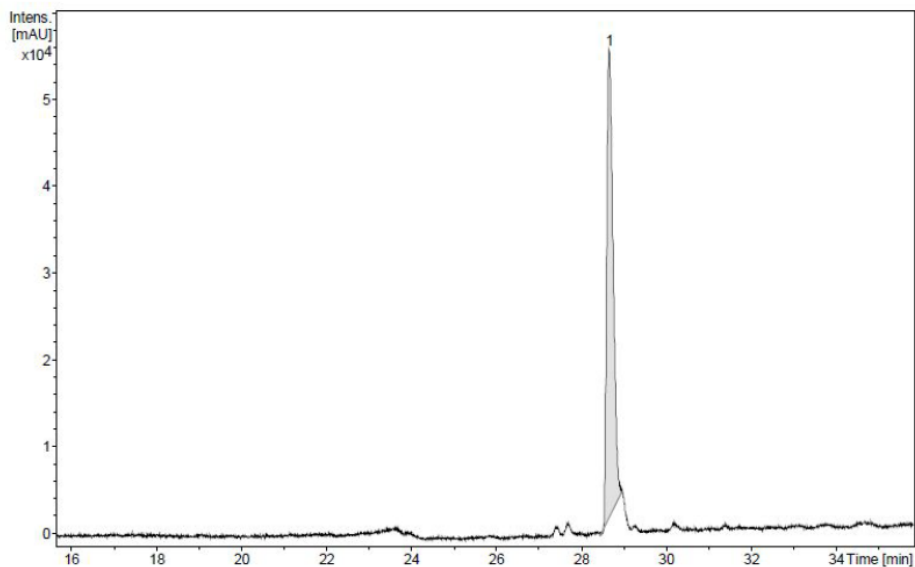


Figure S7: ¹³C NMR (126 MHz, DMSO-d₆) spectrum of 1b.



#	RT [min]	Chromatogram	Area	Area %	S/N
1	28.7	UV Chromatogram, 190-800 nm	562862	100.00	180.9

Figure S8: HPLC chromatogram of **1b**.

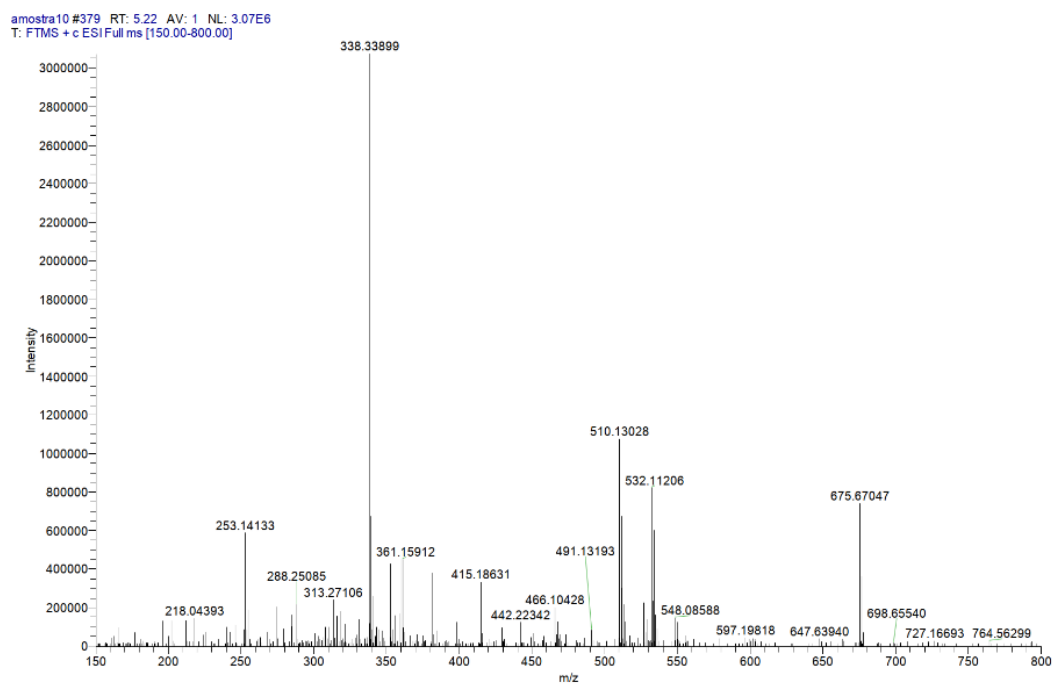
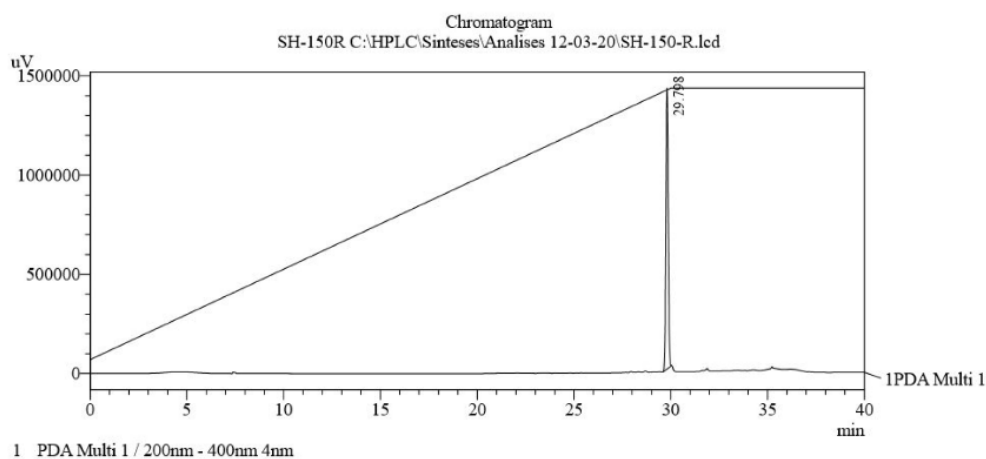


Figure S9: HRMSESI spectrum of **1b**.



PeakTable

PDA Ch1 200nm - 400nm 4nm

Peak#	Ret. Time	Area	Height	Area %	Height %
1	29.798	11768311	1415167	100.000	100.000
Total		11768311	1415167	100.000	100.000

Figure S12: HPLC chromatogram of 1d.

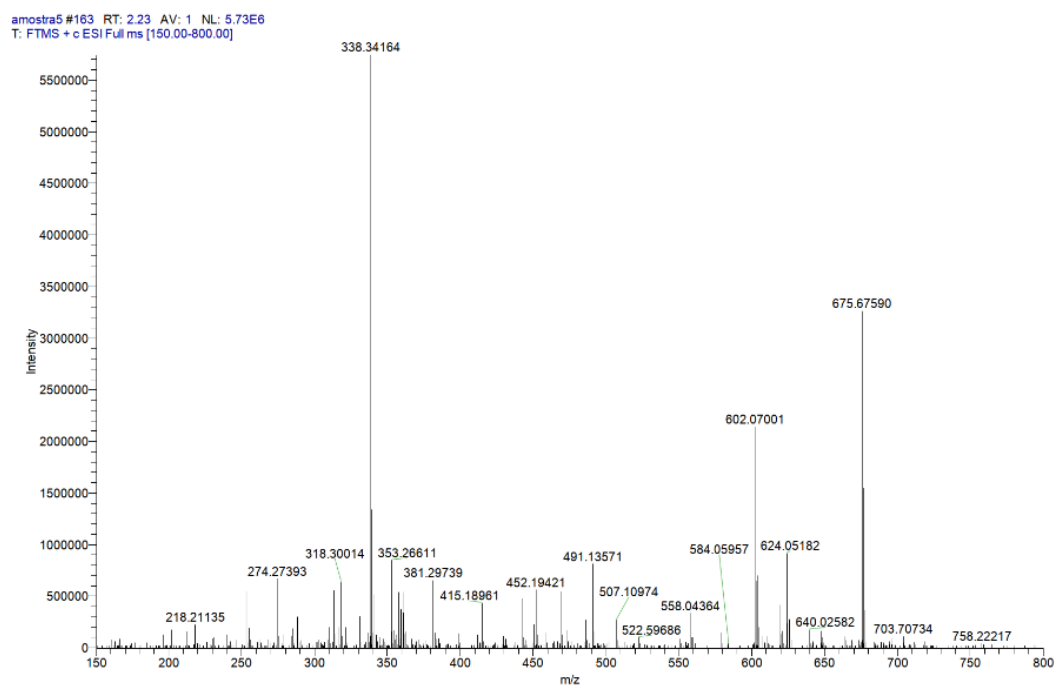
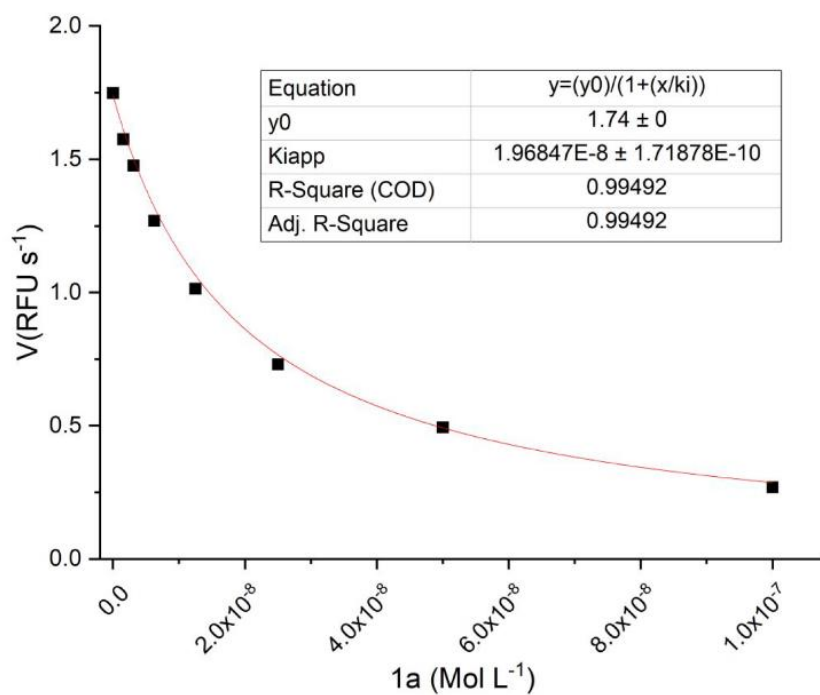
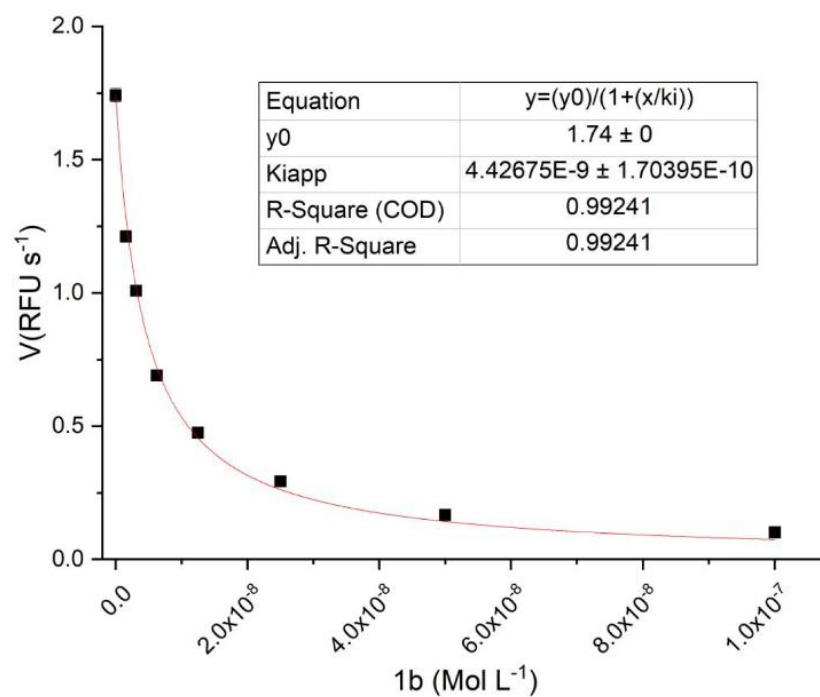


Figure S13: HRMSESI spectrum of 1d.

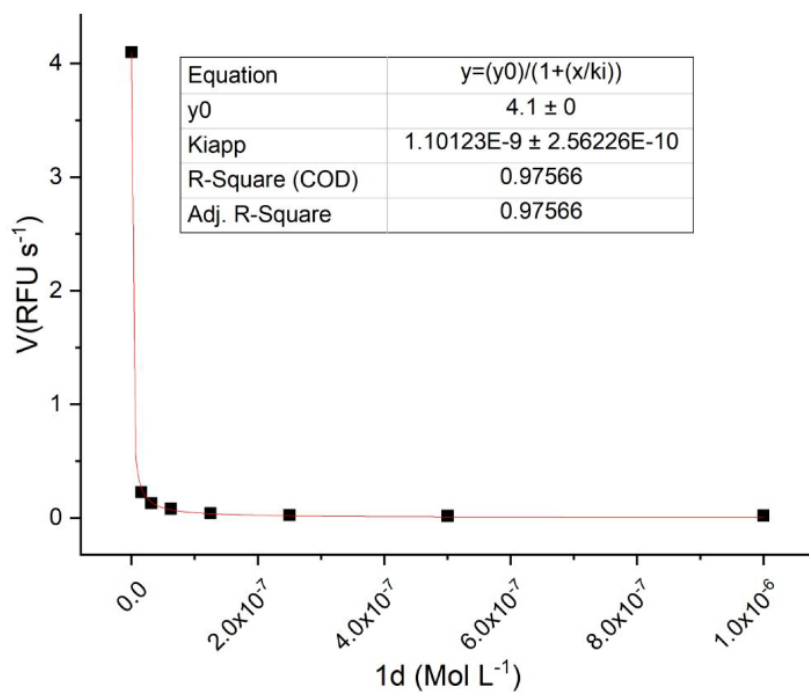
1.5. Determination of the inhibition constant (K_i) against hCatL.



$$pK_i = 8.0 \pm 0.01$$



$$pK_i = 8.6 \pm 0.02$$



$$pK_i = 9.4 \pm 0.02$$

Figure S13: Determination of the inhibition constant (K_i) against human Cathepsin L for (top) compound **1a**, (middle) **1b** and (bottom) **1d**.

2. Computational Results

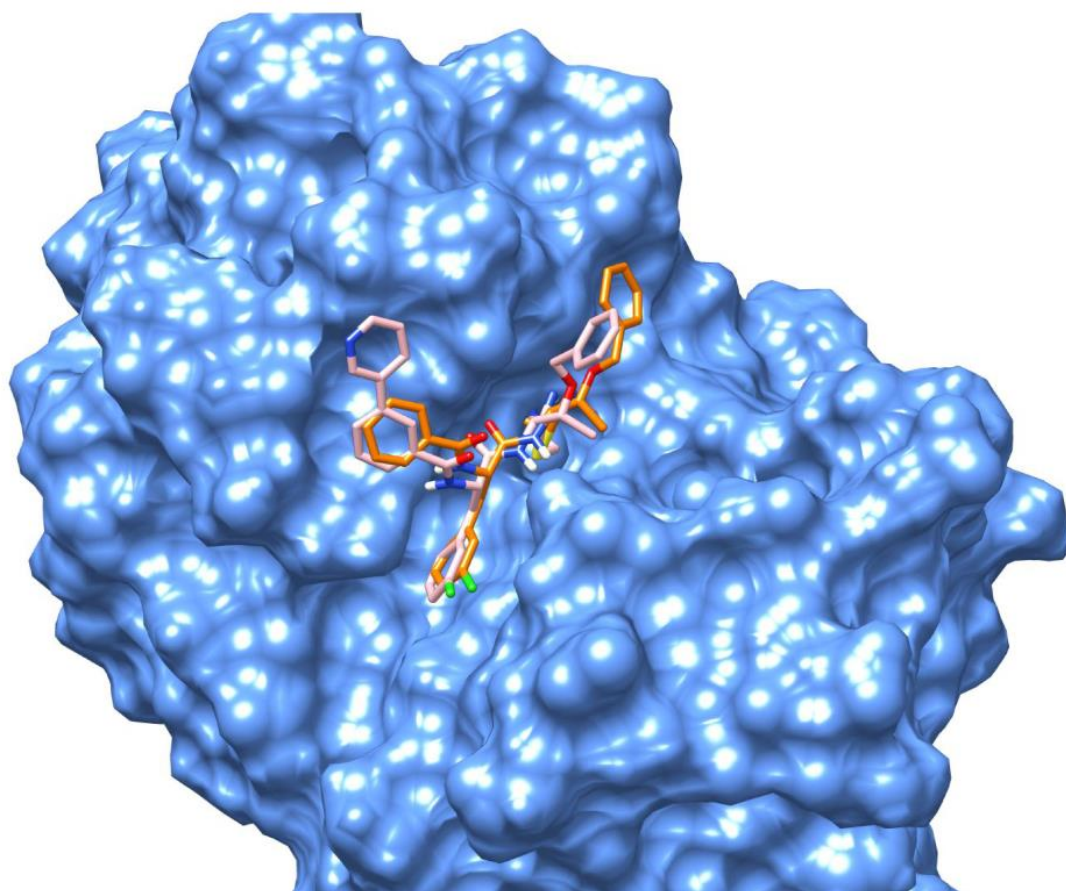


Figure S14: Representation of the superposed arrangements of **1a** (carbon atoms as orange sticks) and **2a** (carbon atoms as pink sticks) in the covalent complex formed in the binding pocket with hCatL (protein surface shown in blue).

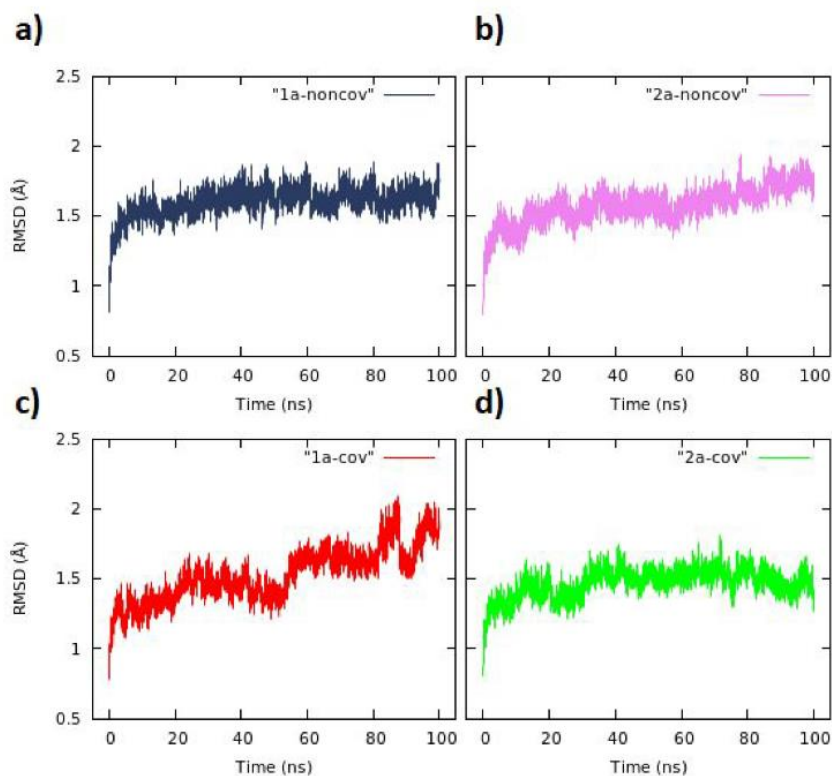


Figure S15: RMSD fluctuations (Å) along 100ns MD simulations in the noncovalent complexes a) **1a**-hCatL b) **2a**-hCatL, and the covalent complexes c) **1a**-CatL d) **2a**-CatL.

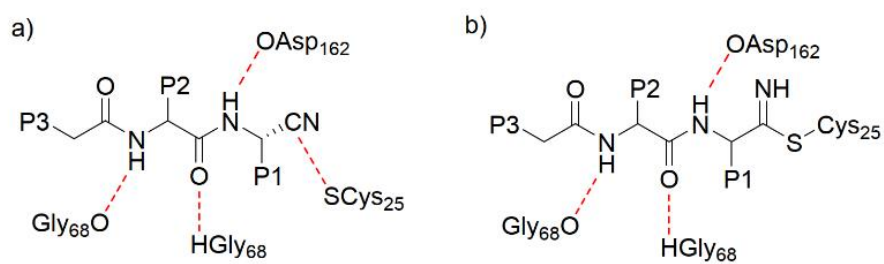


Figure S15: Investigated interactions between the dipeptidyl nitrile-based inhibitor and hCatL in a) noncovalent and b) covalent complexes.

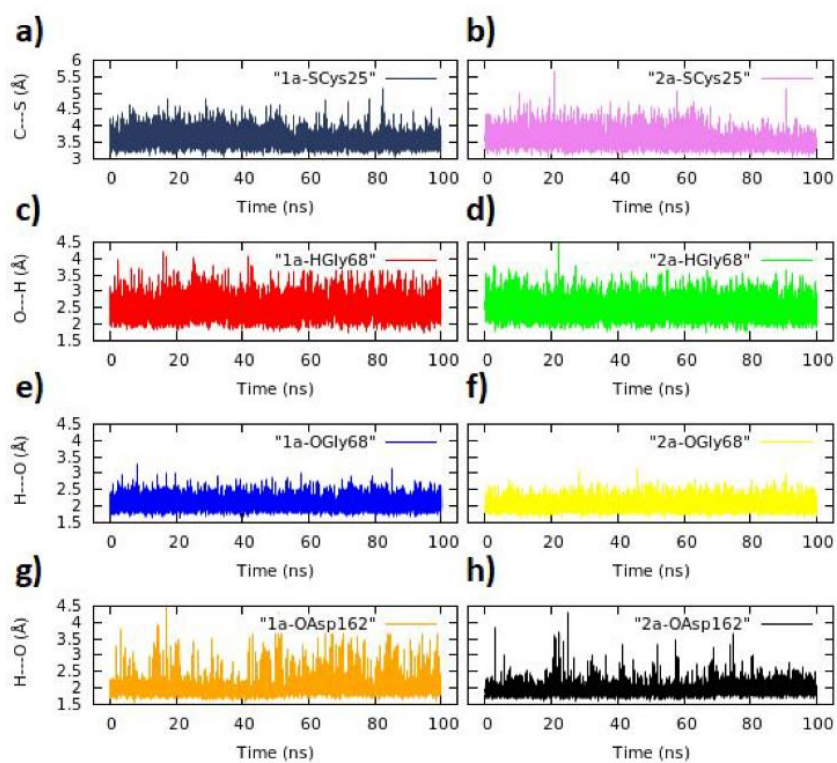


Figure S16: Distance fluctuations of selected noncovalent interactions (Å) in the noncovalent state analyzed over 100 ns MD simulations between a) SCys25, c) HGly68, e) OGly68 ,and g) OAsp162 with compound **1a**, and b) SCys25, d) HGly68, f) OGly68 and h) OAsp162 with **2a**. The interactions are depicted in Figure S16.

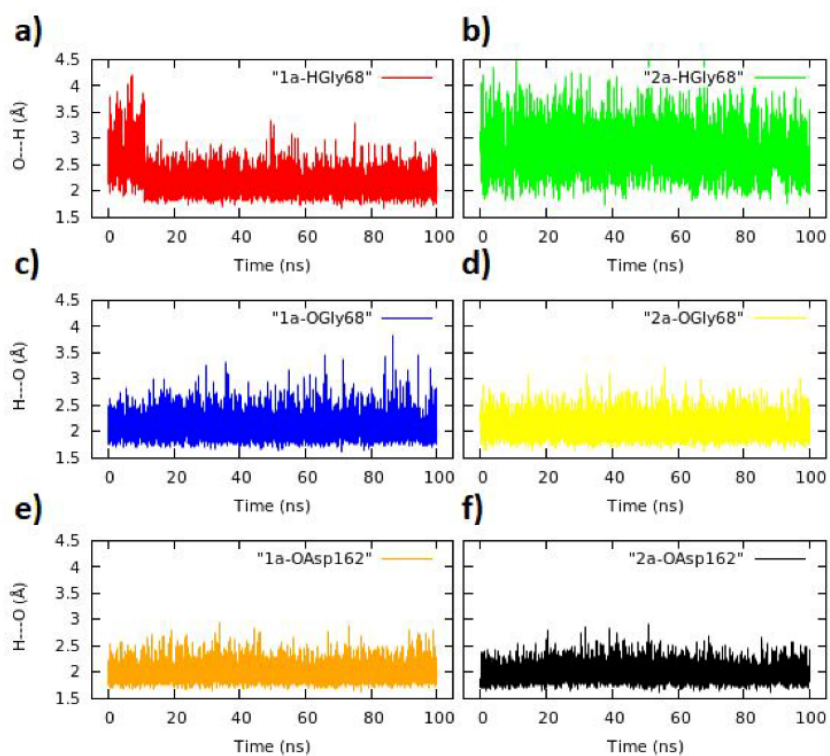


Figure S18: Distance fluctuations of selected noncovalent interactions (Å) in the covalent state analyzed over a 100ns MD simulations between a) HGly68, c) OGly68, and e) OAsp162, with compound **1a**, and b) HGly68, d) OGly68, and f) OAsp162 with **2a**. The interactions are depicted in Figure S16.

Table S1: Outcome of all FEP transformation steps of the ligand-hCatL noncovalent complex. All values in kcal/mol. The standard deviation is represented in parenthesis.

Transformation	Ligand in Water				Complex				$\Delta\Delta G_{\text{Total}}$
	$\Delta G_{\text{decharge}}$	$\Delta G_{\text{recharge}}$	ΔG_{VdW}	ΔG_{Total}	$\Delta G_{\text{decharge}}$	$\Delta G_{\text{recharge}}$	ΔG_{VdW}	ΔG_{Total}	
1a → 1b	-5.75 (0.01)	-8.03 (0.02)	-0.72 (0.04)	-14.49	-5.68 (0.01)	-8.49 (0.02)	-2.81 (0.04)	-16.98	-2.49
1a → 1c	-5.75 (0.01)	-7.23 (0.02)	-6.91 (0.04)	-18.88	-5.68 (0.01)	-7.65 (0.02)	-8.56 (0.04)	-21.89	-2.00
1a → 1d	-5.75 (0.01)	-8.00 (0.02)	-5.22 (0.05)	-18.96	-5.68 (0.01)	-8.74 (0.02)	-6.40 (0.06)	-20.81	-1.85
1a → 1e	-6.39 (0.09)	-2.30 (0.01)	12.84 (0.03)	4.14	-5.93 (0.01)	-2.15 (0.01)	11.36 (0.04)	3.28	-0.86
2a → 2b	5.61 (0.05)	3.68 (0.01)	20.31 (0.04)	29.60	4.86 (0.06)	3.75 (0.01)	19.74 (0.10)	28.34	-1.26
2a → 2c	19.04 (0.02)	-24.07 (0.08)	-96.69 (0.08)	-101.72	18.72 (0.02)	-23.46 (0.08)	-96.97 (0.09)	-101.71	0.01

Table S2: Outcome of all FEP transformation steps of the ligand-hCatL covalent complex. All values in kcal/mol. The standard deviation is represented in parenthesis.

Transformation	Ligand in Water				Complex				$\Delta\Delta G_{\text{Total}}$
	$\Delta G_{\text{decharge}}$	$\Delta G_{\text{recharge}}$	ΔG_{VdW}	ΔG_{Total}	$\Delta G_{\text{decharge}}$	$\Delta G_{\text{recharge}}$	ΔG_{VdW}	ΔG_{Total}	
1a → 1b	-6.15 (0.01)	-8.98 (0.02)	4.92 (0.04)	-10.21	-5.78 (0.01)	-8.74 (0.02)	5.61 (0.04)	-8.91	1.30
1a → 1c	-6.15 (0.01)	-8.44 (0.02)	1.33 (0.04)	-13.25	-5.78 (0.01)	-8.23 (0.03)	0.34 (0.04)	-13.67	-0.42
1a → 1d	-6.15 (0.01)	-8.56 (0.02)	10.42 (0.05)	-4.29	-5.78 (0.01)	-8.39 (0.03)	9.01 (0.05)	-5.16	-0.87
1a → 1e	-6.17 (0.01)	-1.98 (0.01)	10.91 (0.03)	2.76	-5.73 (0.01)	-1.83 (0.01)	10.06 (0.04)	2.49	-0.27
2a → 2b	7.14 (0.04)	3.50 (0.01)	10.94 (0.05)	21.59	6.37 (0.05)	3.61 (0.01)	14.70 (0.05)	24.69	3.10
2a → 2c	19.63 (0.02)	-25.52 (0.08)	-109.80 (0.11)	-115.69	19.26 (0.02)	-25.39 (0.08)	-108.07 (0.09)	-114.20	1.49

Table S3: Parameters used to include the Extra-Point (EP) in MD simulations and alchemical free energy calculations.

Parameters	Values
Mass (EP)	0.00 amu
r^* (EP)	1.00 Å
r_{eq} (Cl—EP)	1.40 Å
r_{eq} (Br—EP)	1.55 Å
r_{eq} (I—EP)	1.60 Å
K_r (X—EP) ^a	600.0
Θ_{eq} (A—X—EP) ^a	180.0°
Cl—EP charge ^b	0.076407
Br—EP charge ^b	0.100989
I—EP charge ^b	0.134570

^aX = Cl, Br or I.
^bThe average value of the EP charge in the noncovalent and covalent structure

Table S1: Outcome of all FEP transformation steps performed for the relative hydration free energies of compounds **2a**, **2b** and **2c**. All values in kcal/mol. The standard deviation is represented in parenthesis.

Transformation	Ligand in Water			Vacuum			$\Delta\Delta G_{\text{Hyd}}$
	$\Delta G_{\text{decharge}}$	$\Delta G_{\text{recharge}}$	ΔG_{VdW}	$\Delta G_{\text{decharge}}$	$\Delta G_{\text{recharge}}$	ΔG_{VdW}	
2a → 2b	5.61 (0.05)	3.68 (0.01)	20.31 (0.04)	19.48 (0.04)	4.09 (0.01)	28.79 (0.01)	-22.74
2a → 2c	19.04 (0.02)	-24.07 (0.08)	-96.69 (0.08)	20.57 (0.02)	-46.59 (0.02)	-65.66 (0.01)	-10.04

Table S2: Outcome of all FEP transformation steps performed for the relative hydration free energies of pyridine, benzene and pyrimidine. All values in kcal/mol. The standard deviation is represented in parenthesis.

Transformation	Ligand in Water			Vacuum			$\Delta\Delta G_{\text{Hyd}}$
	$\Delta G_{\text{decharge}}$	$\Delta G_{\text{recharge}}$	ΔG_{VdW}	$\Delta G_{\text{decharge}}$	$\Delta G_{\text{recharge}}$	ΔG_{VdW}	
Pyridine → Benzene	29.51 (0.05)	-0.04 (0.01)	59.97 (0.04)	47.65 (0.06)	0.84 (0.01)	36.92 (0.10)	4.03
Pyridine → Pyrimidine	42.25 (0.02)	-49.34 (0.08)	-110.04 (0.08)	46.03 (0.02)	-78.17 (0.08)	-74.87 (0.09)	-10.12

3. Analysis of the discharge, vdW and recharge steps in alchemical transformations.

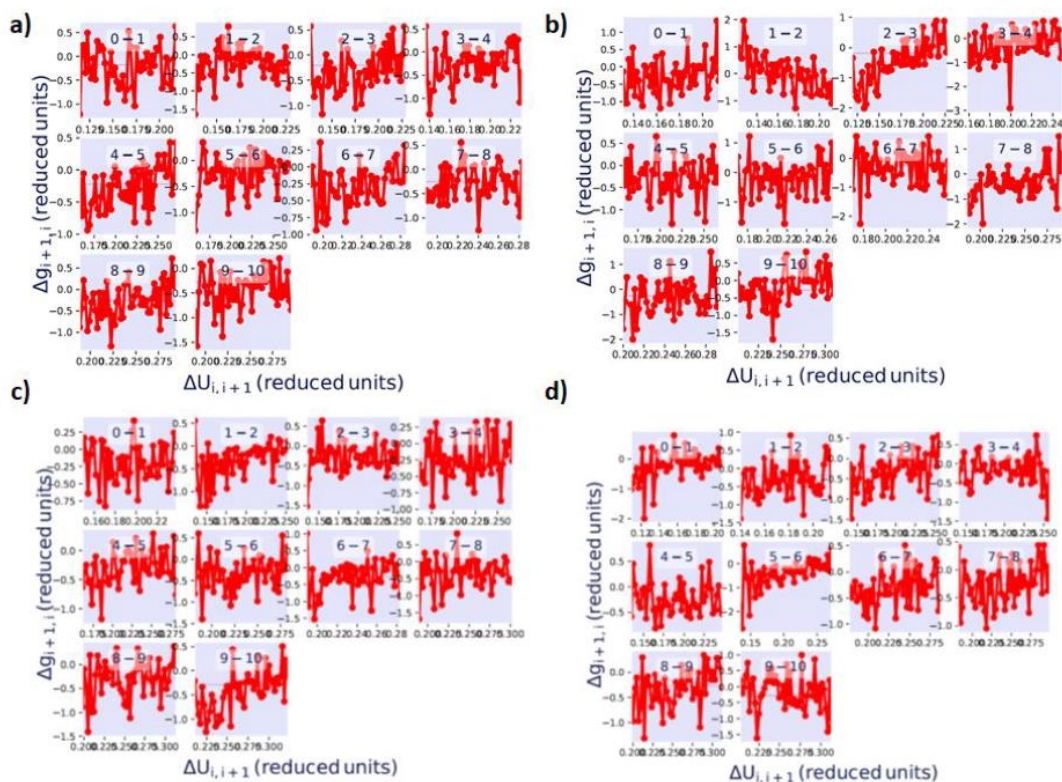


Figure S17: The overlapping distribution method for the discharge step of compound **1a** for the *meta* Hydrogen in a) noncovalent state in water, b) noncovalent state in protein, c) covalent state in water and d) covalent state in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

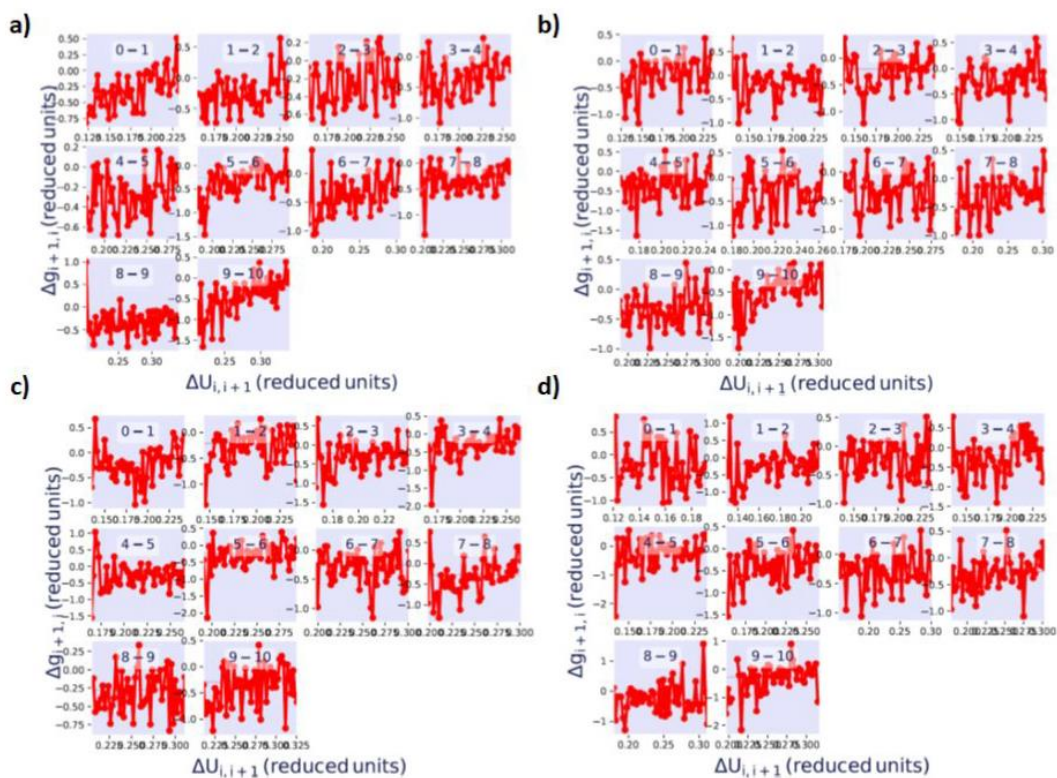


Figure S20: The overlapping distribution method for the discharge step of compound **1a** for the *para* Hydrogen in a) noncovalent state in water, b) noncovalent state in protein, c) covalent state in water and d) covalent state in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

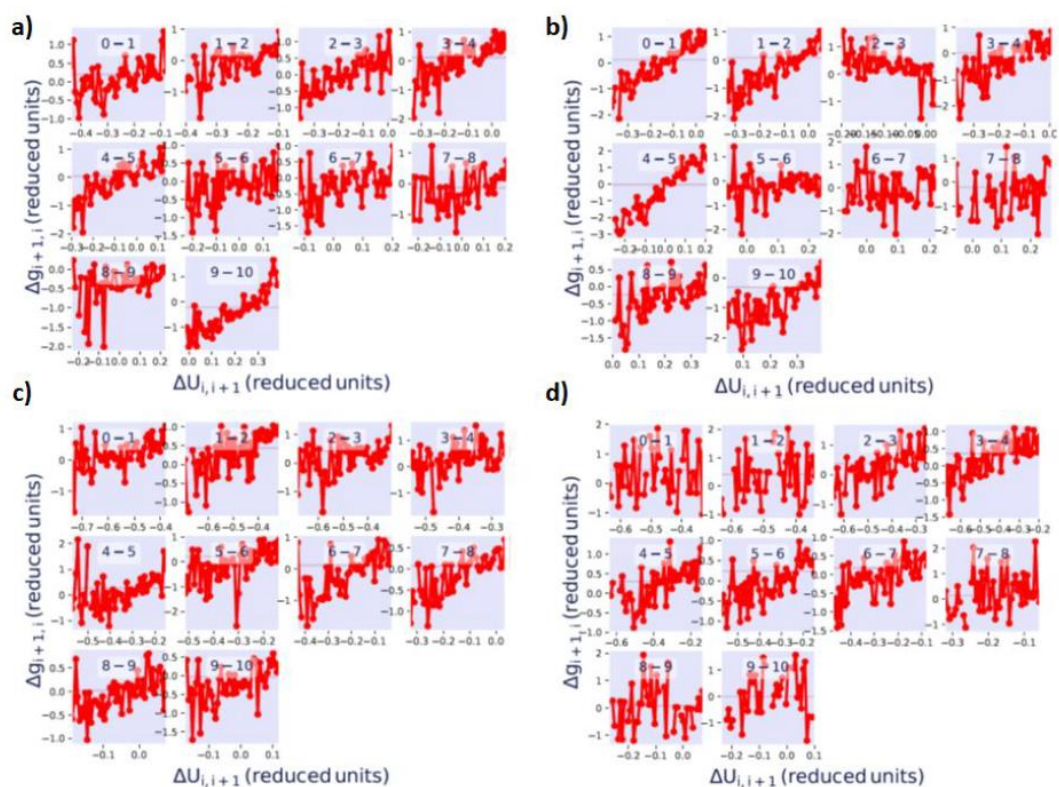


Figure S21: The overlapping distribution for $1a \rightarrow 1b$ transformation for noncovalent state a) vdW for ligand in water, and b) vdW for ligand in protein; and in covalent state c) vdW for ligand in water and d) vdW for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+i} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

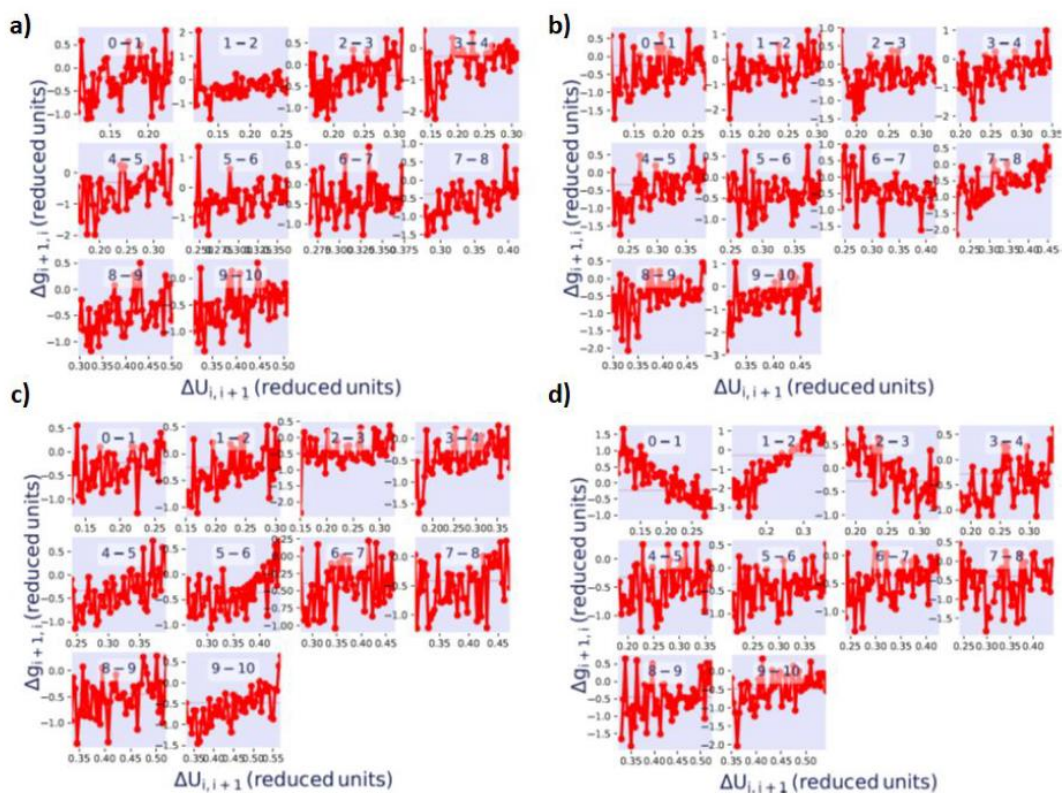


Figure S18: The overlapping distribution for $1a \rightarrow 1b$ transformation for noncovalent state a) recharge for ligand in water and b) recharge for ligand in protein; and covalent state c) recharge for ligand in water and d) recharge for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

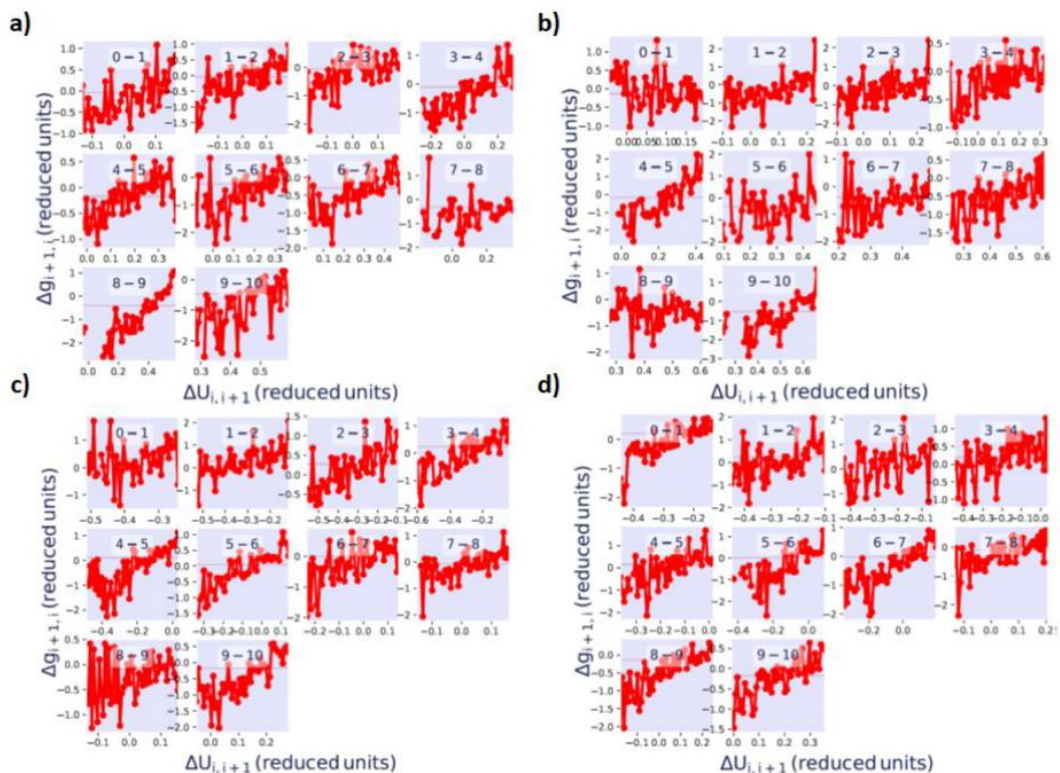


Figure S19: The overlapping distribution for $1a \rightarrow 1c$ transformation for noncovalent state a) vdW for ligand in water, and b) vdW for ligand in protein; and in covalent state c) vdW for ligand in water and d) vdW for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

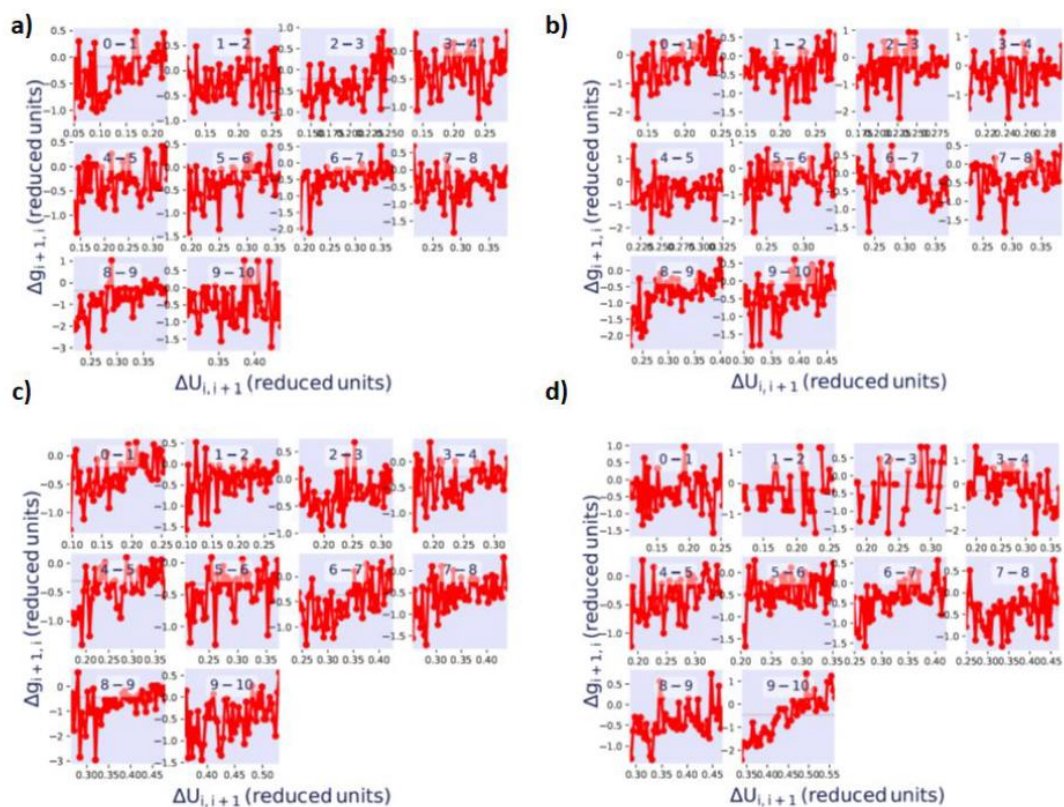


Figure S20: The overlapping distribution for $1a \rightarrow 1c$ transformation for noncovalent state a) recharge for ligand in water and b) recharge for ligand in protein; and covalent state c) recharge for ligand in water and d) recharge for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

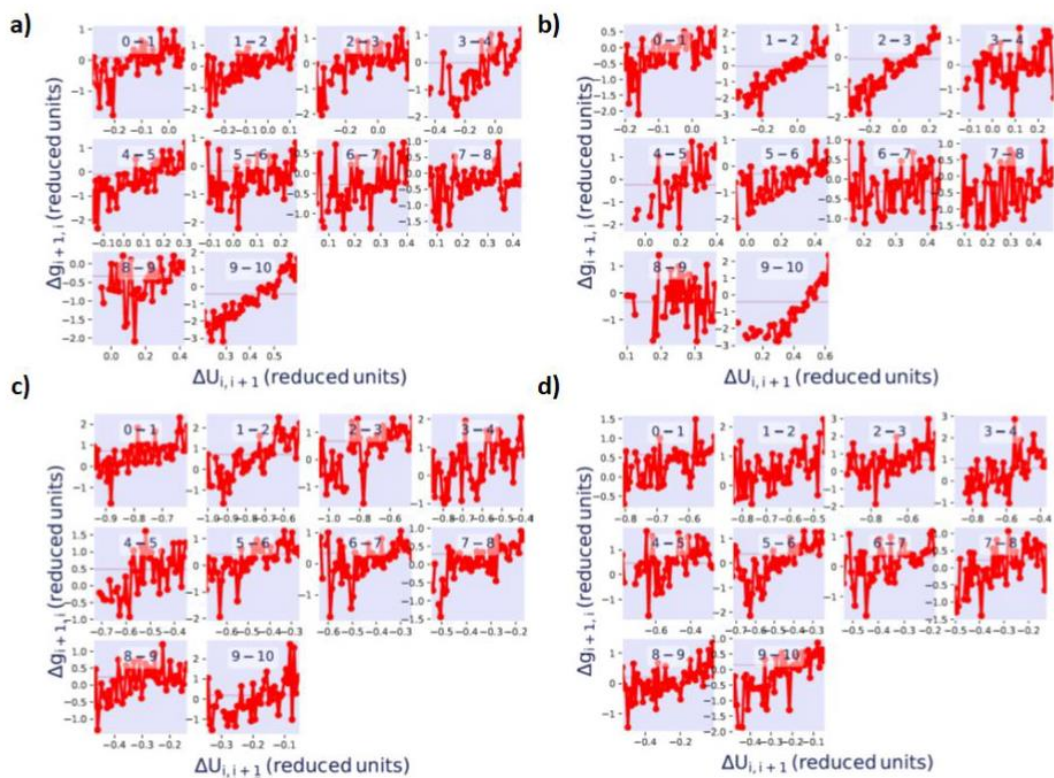


Figure S21: The overlapping distribution for $1a \rightarrow 1d$ transformation for noncovalent state a) vdW for ligand in water, and b) vdW for ligand in protein; and in covalent state c) vdW for ligand in water and d) vdW for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

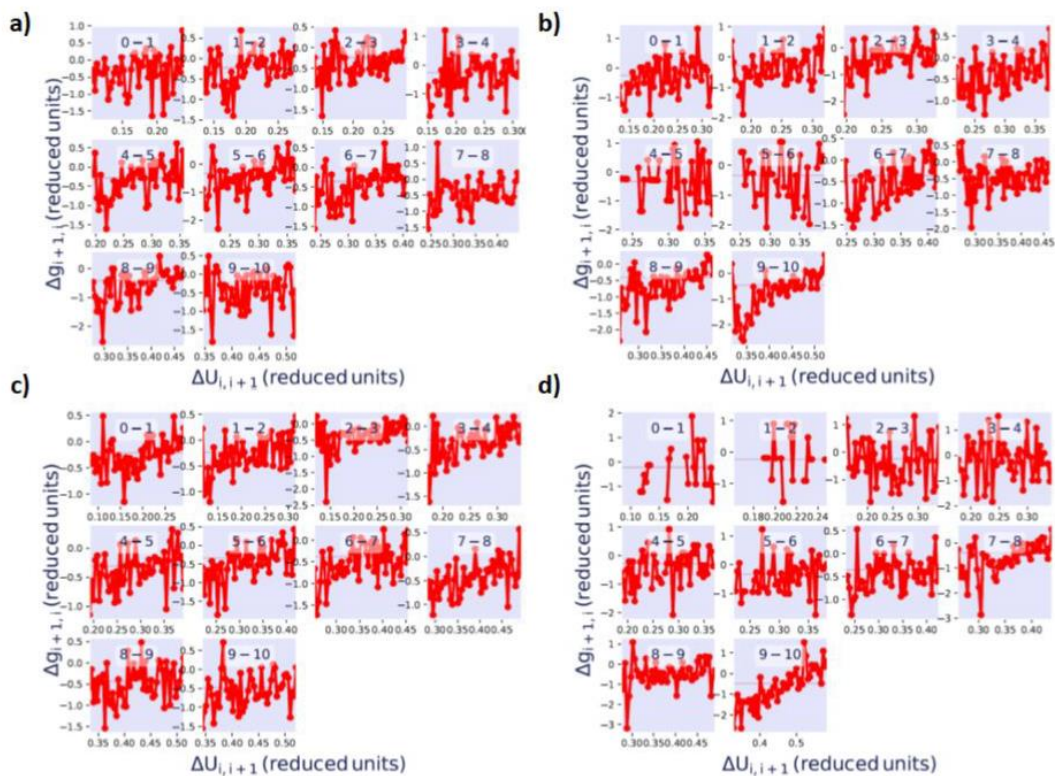


Figure S22: The overlapping distribution for $1a \rightarrow 1d$ transformation for noncovalent state a) recharge for ligand in water and b) recharge for ligand in protein; and covalent state c) recharge for ligand in water and d) recharge for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

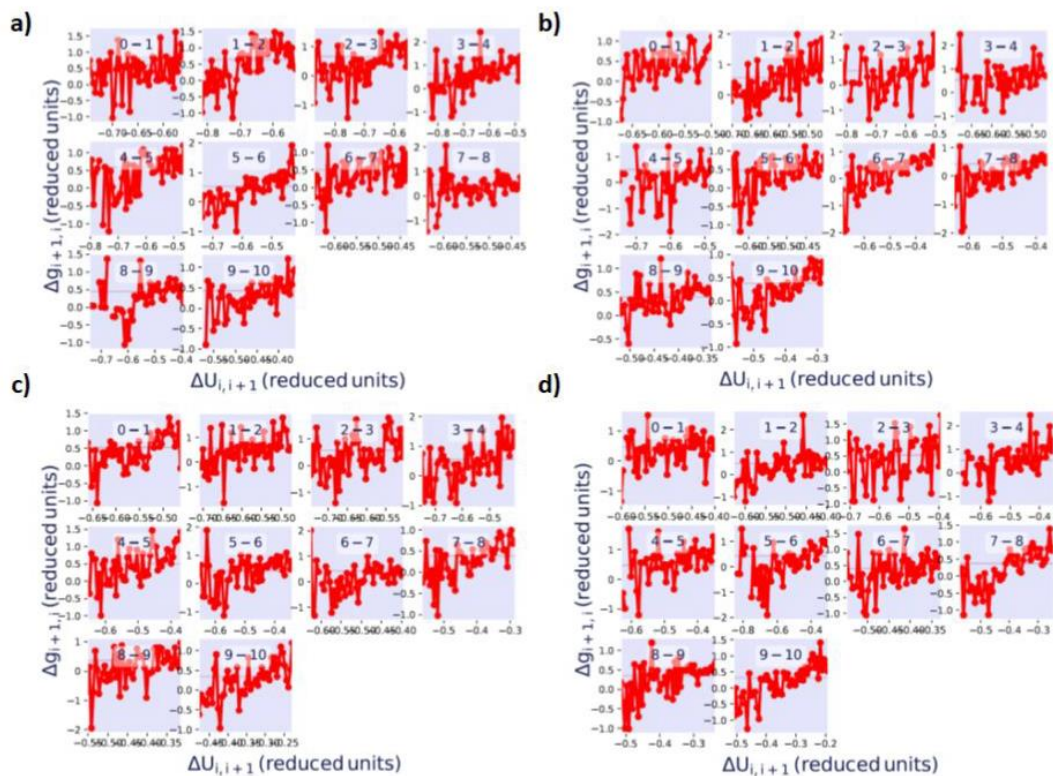


Figure S23: The overlapping distribution for $1a \rightarrow 1e$ transformation for noncovalent state a) vdW for ligand in water, and b) vdW for ligand in protein; and in covalent state c) vdW for ligand in water and d) vdW for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

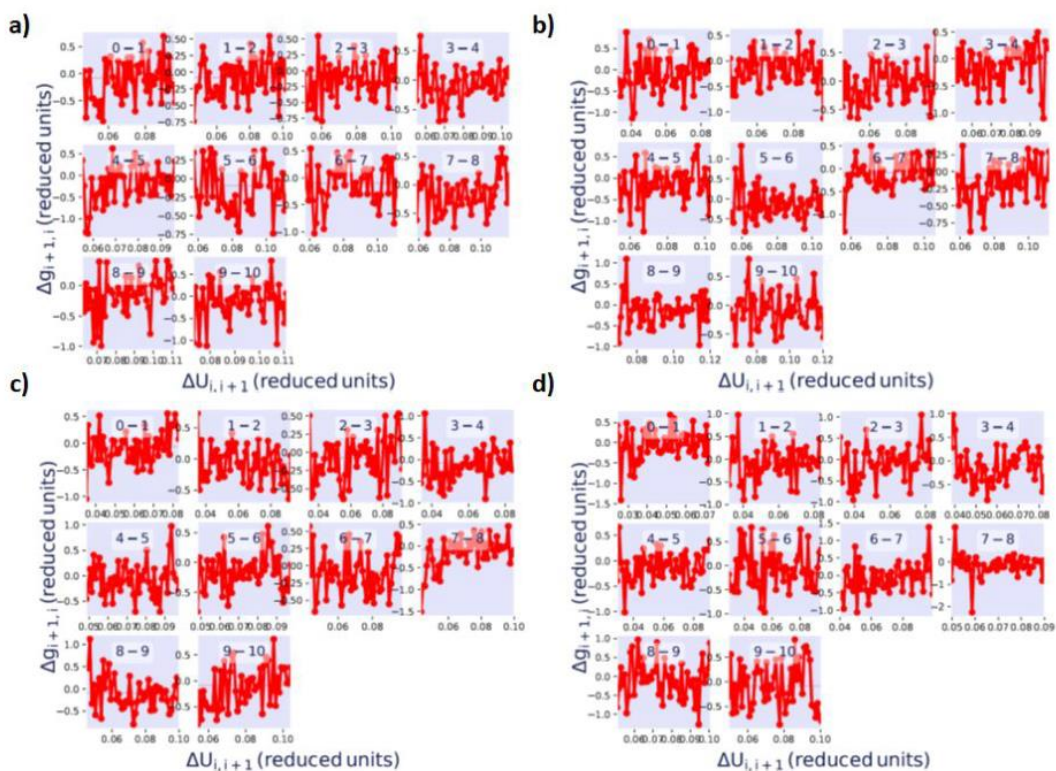


Figure S24: The overlapping distribution for $1a \rightarrow 1e$ transformation for noncovalent state a) recharge for ligand in water and b) recharge for ligand in protein; and covalent state c) recharge for ligand in water and d) recharge for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

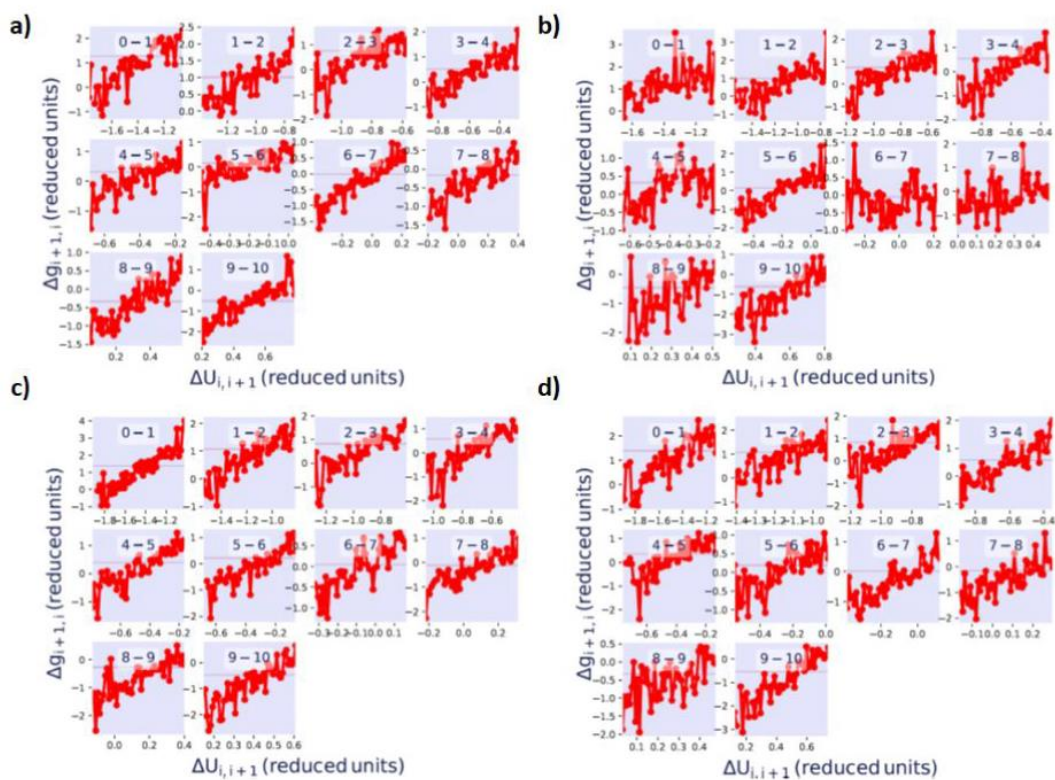


Figure S25: The overlapping distribution for $2a \rightarrow 2b$ transformation for noncovalent state a) discharge for ligand in water and b) discharge for ligand in protein; and covalent state c) discharge for ligand in water and d) discharge for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

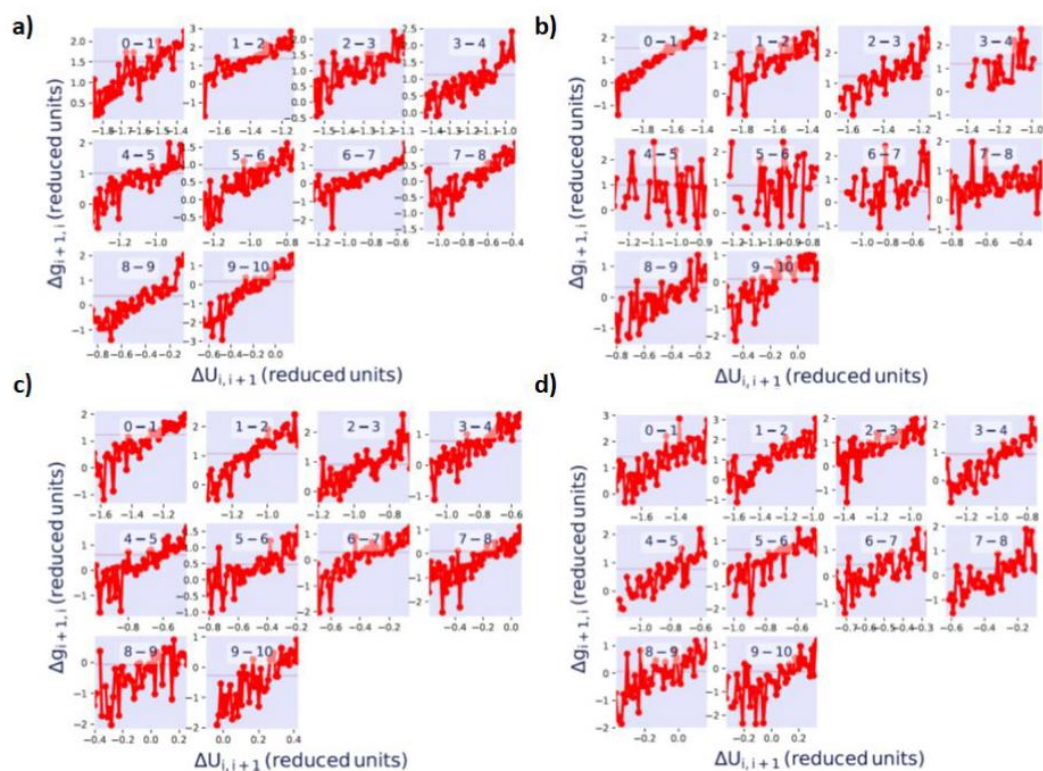


Figure S30: The overlapping distribution for $2a \rightarrow 2b$ transformation for noncovalent state a) vdW for ligand in water, and b) vdW for ligand in protein; and in covalent state c) vdW for ligand in water and d) vdW for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

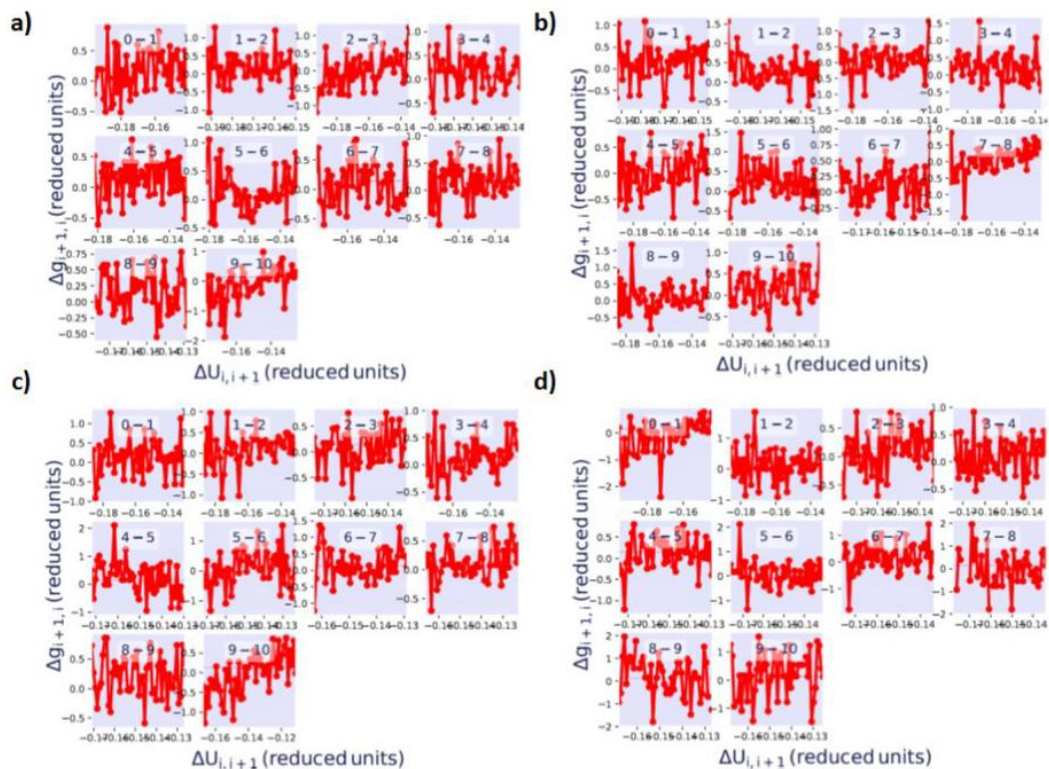


Figure S31: The overlapping distribution for $2a \rightarrow 2b$ transformation for noncovalent state a) recharge for ligand in water and b) recharge for ligand in protein; and covalent state c) recharge for ligand in water and d) recharge for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

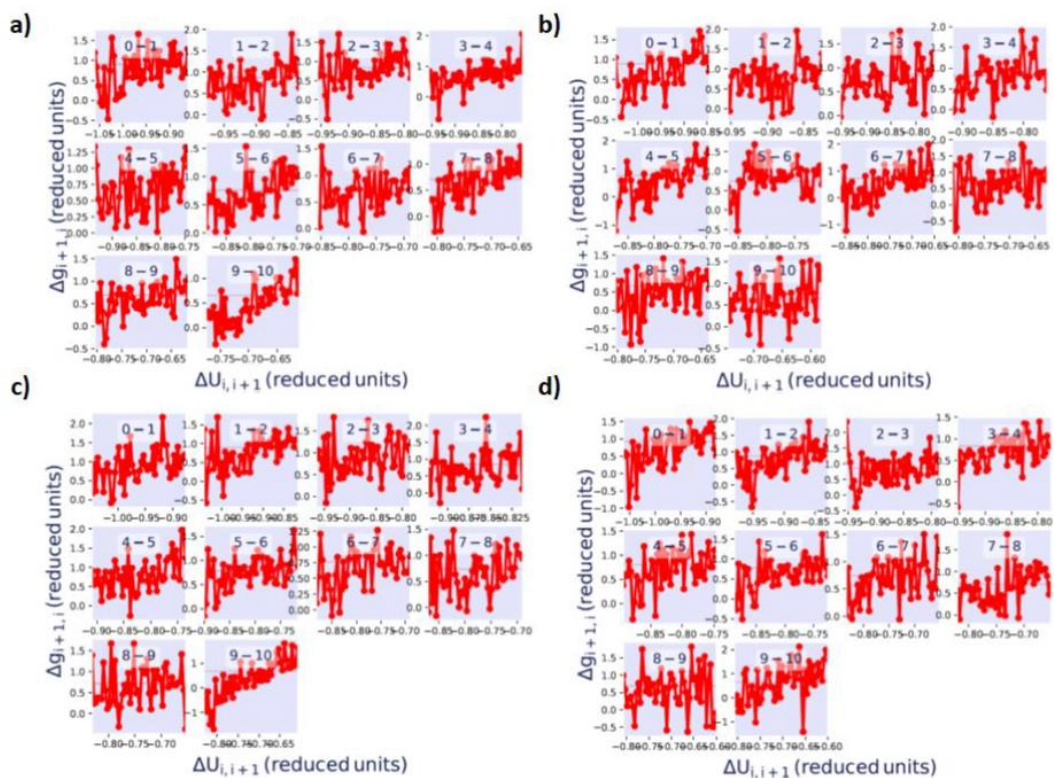


Figure S26: The overlapping distribution for $2a \rightarrow 2c$ transformation for noncovalent state a) decharge for ligand in water and b) decharge for ligand in protein; and covalent state c) decharge for ligand in water and d) decharge for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

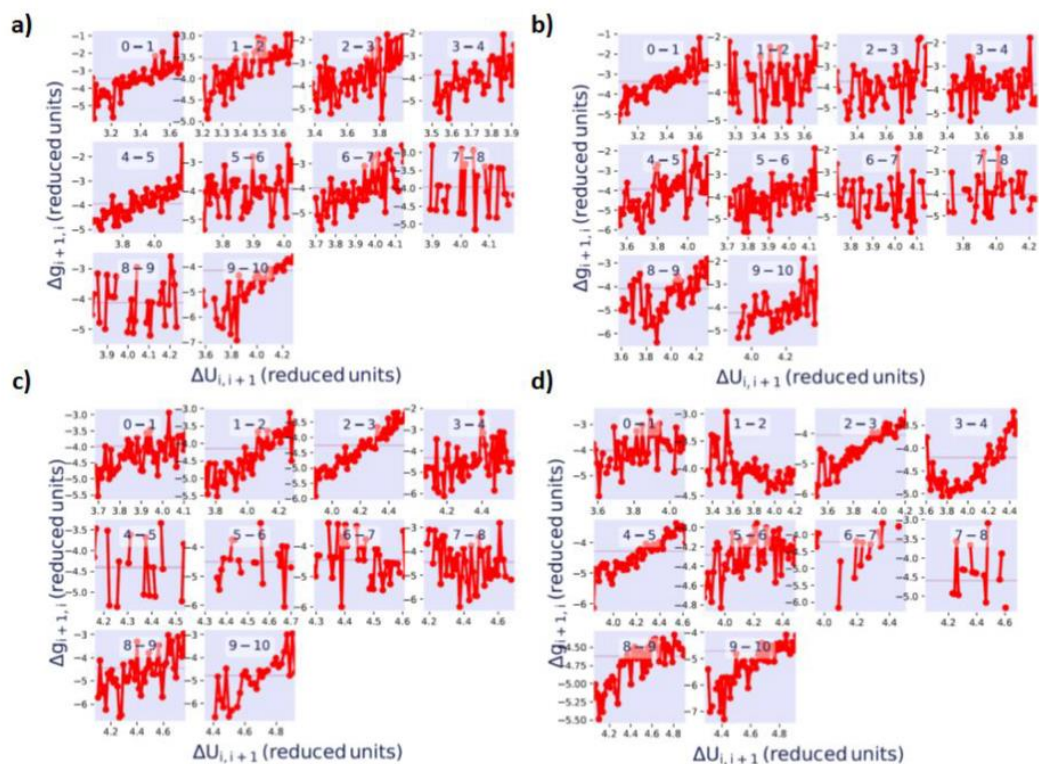


Figure S27: The overlapping distribution for $2a \rightarrow 2c$ transformation for noncovalent state a) vdW for ligand in water, and b) vdW for ligand in protein; and in covalent state c) vdW for ligand in water and d) vdW for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+1} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.

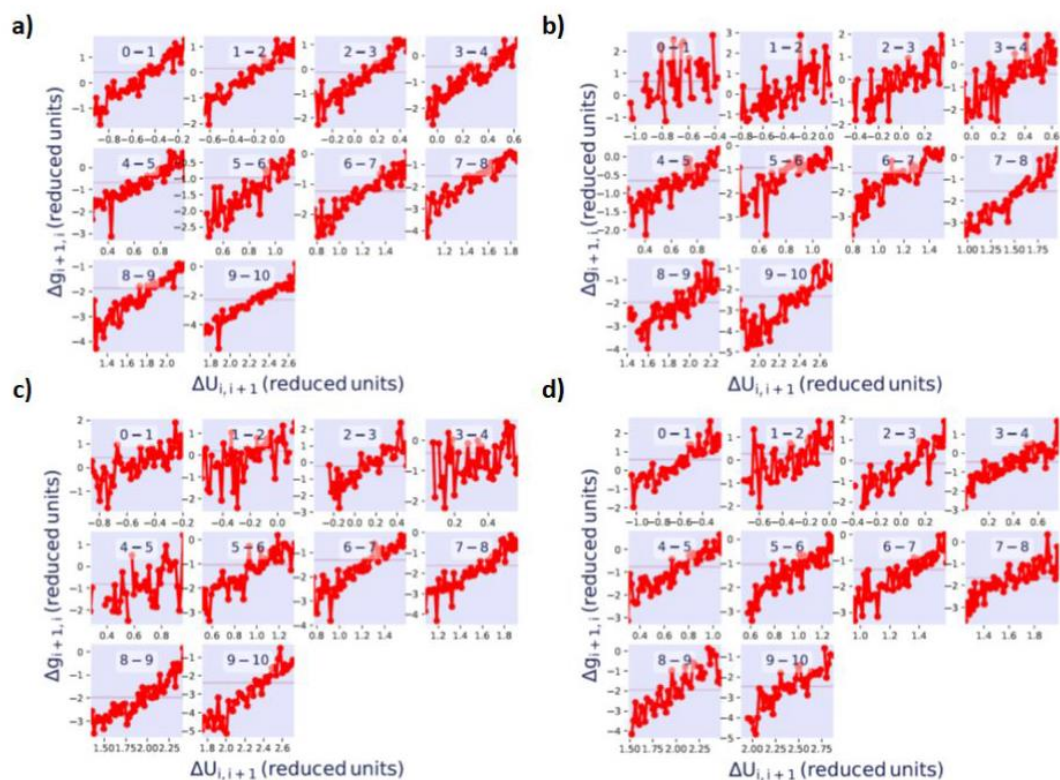


Figure S28: The overlapping distribution for $2a \rightarrow 2c$ transformation for noncovalent state a) recharge for ligand in water and b) recharge for ligand in protein; and covalent state c) recharge for ligand in water and d) recharge for ligand in protein. Note that this step is the same for all transformations studied. Here, the difference $\Delta g_{i,i+1} = g_{i+i} - g_i$ is plotted as a function of $\Delta U_{i,i+1}$ to evaluate the overlap between the ΔU distributions of two adjacent states.