



**HAL**  
open science

# Multiscale approaches to help clarify mechanisms of activation/inhibition for important biological systems

Angela Parise

► **To cite this version:**

Angela Parise. Multiscale approaches to help clarify mechanisms of activation/inhibition for important biological systems. Biochemistry, Molecular Biology. Université Paris-Saclay; Università degli studi della Calabria, 2022. English. NNT : 2022UPASF025 . tel-03714141

**HAL Id: tel-03714141**

**<https://theses.hal.science/tel-03714141>**

Submitted on 5 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Multiscale Approaches To Help Clarify Mechanisms of Activation/Inhibition for Important Biological Systems

*Approches multi-échelles pour aider à clarifier les mécanismes impliqués  
dans les processus d'activation/inhibition de systèmes biologiques  
importants*

**Thèse de doctorat de l'université Paris-Saclay et de l'Univerità della Calabria**

École doctorale n°571 : sciences chimiques : molécules, matériaux,  
instrumentation et biosystèmes (2MIB)  
Spécialité de doctorat: Chimie  
Graduate school: Chimie  
Référent: Faculté des sciences d'Orsay

Thèse préparée dans les unités de recherche **Institut de Chimie Physique** (Université Paris-Saclay, CNRS) et **Progettazione Molecolare e Chimica dei Sistemi Complessi** (Università della Calabria), sous la direction d'**Aurélien de la LANDE**, Directeur de recherche, et la co-direction de **Tiziana MARINO**, Professeur

**Thèse soutenue à Rende (Italie), le 29 mars 2022, par**

**Angela PARISE**

## Composition du Jury

**Dennis SALAHUB**

Professeur, Université de Calgary

Président & Rapporteur

**Fahmi HIMO**

Professeur, Stockholm University

Rapporteur & Examineur

**Nathalie BASDEVANT**

Maîtresse de conférences, Université Paris-Saclay

Examinatrice

**Cesare INDIVERI**

Professeur, Università della Calabria

Examineur

**Aurélien de la LANDE**

Directeur de recherche CNRS, Université Paris-Saclay  
(UMR 8000)

Directeur de thèse

**Titre :** Approches multi-échelles pour aider à clarifier les mécanismes impliqués dans les processus d'activation/inhibition de systèmes biologiques importants

**Mots clés :** agrégats fer-soufre, antiviraux naturels, enzymes à radical SAM, activation et inhibition enzymatique, chimie computationnelle.

**Résumé :** Les processus biomoléculaires sont difficiles à étudier en raison de l'organisation intrinsèquement multi-échelle et complexe des systèmes dans lesquels ils se produisent. Les méthodes théoriques sont désormais largement utilisées dans la compréhension des processus biochimiques. La précision croissante des approches computationnelles et des architectures informatiques permettent de décrire correctement les voies de réaction impliquées dans les processus biologiques, sans beaucoup de limitations autres que celles strictement liées au protocole théorique adopté. La biochimie computationnelle peut arriver à un niveau prédictif et fournir des données très détaillées sur des processus difficiles à analyser expérimentalement. Une application très importante de la modélisation moléculaire est la conception rationnelle de médicaments fondée sur la connaissance d'une cible biologique (conception de médicaments assistée par ordinateur). L'objectif de cette thèse est d'élucider les mécanismes d'activation et d'inhibition impliqués dans l'activité de certaines enzymes en utilisant des approches computationnelles multi-échelles originales. Tous les systèmes étudiés sont des processus biologiques multi-étapes impliqués dans des pathologies, telles que le cancer ou des infections virales. La clarification des mécanismes (réaction ou inhibition) est donc d'une importance fondamentale pour le développement de stratégies thérapeutiques.

Le **chapitre 1** fournit le contexte biochimique général permettant de comprendre les systèmes biologiques étudiés dans cette thèse, tandis que dans le **chapitre 2**, nous donnons un bref aperçu du contexte théorique et des méthodologies utilisées dans ce travail. Dans le **chapitre 3**, nous rendons compte de la production scientifique de la thèse, résumée ci-dessous. L'article I rapporte l'étude du mécanisme de réaction d'un catalyseur non hémique mononucléaire à base de fer biomimétique lié à l'activation de l'oxygène par la réduction de 2 électrons dans la décarboxylation oxydative

d'hydroxyacides comme réducteurs. La partie suivante porte sur les mécanismes d'inhibition des enzymes impliqués dans la réplication cellulaire. En particulier, le processus de métallisation de la ribonucléase A par un complexe contenant du Pt-As a été élucidé par des calculs DFT à l'aide d'un modèle capable de représenter de manière assez réaliste l'environnement du site de liaison. Le document III présente le mécanisme d'inhibition de la protéase principale (Mpro) du SRAS-CoV-2 en appliquant des méthodes de calcul allant de la dynamique moléculaire classique et des simulations de *docking* moléculaire aux calculs DFT et propose deux molécules comme inhibiteurs prometteurs de Mpro. Le document IV se concentre sur l'étude de l'enzyme antivirale naturelle, la viperin, une protéine inhibitrice du virus inducible par l'interféron, associée au réticulum endoplasmique, qui possède dans son site actif un cluster fer-soufre impliqué dans le processus catalytique. Il s'agit d'un membre de la superfamille des enzymes SAM radicales qui utilisent la SAM comme substrat pour convertir la SAM de façon stœchiométrique en Met et en 5'-désoxyadénosine. Les caractéristiques structurales de l'enzyme ont été analysées en profondeur par des simulations de dynamique moléculaire. Le potentiel d'ionisation a été évalué par des approches computationnelles avancées. En outre, comme le produit catalytique de la vipérine est une nucléobase modifiée (ddhCTP) basée sur une structure pyrimidine, elle a été testée en tant qu'inhibiteur dépendant de l'ARN-polymérase du SRAS-CoV-2 avec une nature de terminateur de chaîne (Wp).

La section B, qui comprend les articles V et VI, concerne l'étude théorique des composés naturels ayant des propriétés antioxydantes essentielles et les mécanismes de réaction impliqués dans le piégeage des radicaux libres pour minimiser le stress oxydatif.

**Title :** Multiscale Approaches to Help Clarify Mechanisms of Activation/Inhibition for Important Biological Systems

**Keywords :** iron-sulfur cluster, natural antivirals, radical SAM enzymes, enzyme inhibition and activation, computational chemistry

**Abstract :** Investigating biomolecular processes is challenging because of the intrinsic multi-scale hierarchical organization in system size and time in which chemical and chemical physical phenomena occur. Theoretical methods applied to biomolecules deal with the use of computation to obtain information about biological systems and their many interactions. The increasing accuracy of computational approaches and of computer architectures, allow one to properly describe reaction paths involved in both activation/inhibition enzymes, calculate electronic properties and structural dynamics conformations of biomolecules without any limitation except those connected with the adopted theoretical protocol. Currently, computational biochemistry has predictive power and provides very insightful data that are sometimes tedious to obtain from experiment. In addition to the clarification of biochemical processes at the molecular and atomistic level, a very important application of biomolecular modeling is to guide the rational drug design process based on knowledge of a biological target (Computer-Aided Drug Design). The objective of this PhD thesis is to elucidate activation or inhibition mechanisms involved in enzyme activity by using multiscale and original computational approaches. All the systems studied are multistep biological processes implicated in diseases, such as cancer and viral infections, so elucidating the mechanisms (reaction or inhibition) is of crucial importance to develop new therapeutic strategies.

**Chapter 1** provides general background to understand the biological systems studied in this thesis. In **Chapter 2**, we give a brief overview of the theoretical background and methodologies used in this work. In **Chapter 3**, we report the scientific production, summarised below.

Paper I reports the study of reaction mechanism of a mononuclear non-heme biomimetic iron catalyst related to the activation of oxygen by the reduction of two electrons in the oxidative decarboxylation of hydroxy acids as reductants.

The next part deals with the mechanisms of enzyme inhibition in cell replication. In particular, the process of metallization of ribonuclease A Pt-As-containing complex has been elucidated by DFT calculations using a large cluster model capable of representing the binding site environment in a fairly realistic way. Paper III discusses the mechanism of inhibition of the main protease (Mpro) of SARS-CoV-2 by applying computational methods ranging from classical molecular dynamics and molecular docking simulations to DFT calculations and proposes two molecules as promising inhibitors of Mpro. Paper IV focuses on the study of the natural antiviral enzyme, viperin, an endoplasmic reticulum-associated interferon-inducible virus inhibitory protein, which has in its active site an iron-sulphur cluster involved in the catalytic process. It is a member of the superfamily of radical SAM enzymes that use SAM as a substrate to convert SAM stoichiometrically to Met and 5'-deoxyadenosine. The structural characteristics of the enzyme were analysed in depth by molecular dynamics simulations. The ionisation potential was evaluated by advanced computational approaches. In addition, as the catalytic product of viperin is a modified nucleobase (ddhCTP) based on a pyrimidine structure, it was tested as a SARS-CoV-2 RNA polymerase dependent inhibitor with a chain terminator (Wp) nature.

Section B, which includes papers V and VI. Both concerns the theoretical study of natural compounds with essential antioxidant properties and the reaction mechanisms involved in free radical scavenging to minimise oxidative stress.





UNIVERSITÀ DELLA  
CALABRIA

UNIVERSITA' DELLA CALABRIA

Dipartimento di Chimica e Tecnologie Chimiche

**Dottorato di Ricerca in**

Medicina Traslazionale

**in co-tutela di tesi con**

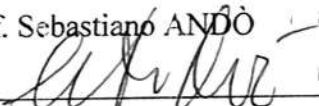
*Institut de Chimie Physique - Université Paris-Saclay*

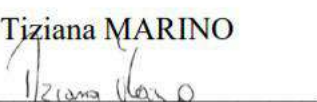
**CICLO**


**XXXIV**

**MULTISCALE APPROACHES TO HELP CLARIFY MECHANISMS OF  
ACTIVATION/INHIBITION FOR IMPORTANT BIOLOGICAL SYSTEMS**

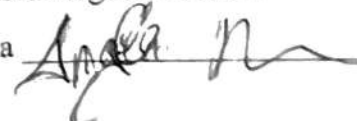
**Settore Scientifico Disciplinare CHIM/03**

**Coordinatore:** Ch.mo Prof. Sebastiano ANDÒ  
Firma 

**Supervisore/Tutor:** Prof.ssa Tiziana MARINO  
Firma 

**Supervisore/Tutor:** Prof. Aurélien DE LA LANDE  
Firma 

**Dottorando:** Dott.ssa Angela PARISE

Firma 

# Contents

Sintesi in italiano .....	1
Synthèse en français.....	2
<b>Acknowledgements</b> .....	<b>4</b>
<b>List of publications</b> .....	<b>5</b>
<b>Introduction</b> .....	<b>6</b>
<b>1. Biological Background</b> .....	<b>9</b>
<b>1.1 Enzyme catalysis</b> .....	<b>9</b>
1.1.1 <i>Michaelis Menten kinetics</i> .....	10
1.1.2 <i>Cofactors</i> .....	11
<b>1.2 Inhibition mechanisms</b> .....	<b>13</b>
1.2.1 <i>Reversible inhibition</i> .....	14
1.2.2 <i>Irreversible inhibition</i> .....	15
1.2.3 <i>Inhibitors as potential drugs</i> .....	16
<b>1.3 Enzymes as technological inspiration: biomimetic systems</b> .....	<b>17</b>
<b>1.4 Prevention of oxidative damage to bio-molecules</b> .....	<b>17</b>
<b>References</b> .....	<b>19</b>
<b>2. Computational Methods</b> .....	<b>22</b>
<b>2.1 Computational methods based on classical physics</b> .....	<b>23</b>
2.1.1 <i>Force Fields</i> .....	23
2.1.1 <i>FF: inclusion of induction</i> .....	25
2.1.2 <i>Molecular Dynamics simulations</i> .....	26
<b>2.2 Molecular Docking</b> .....	<b>28</b>
<b>2.3 From computational methods based on wave functions to Density Functional Theory</b> .....	<b>29</b>
2.3.1 <i>QM cluster models</i> .....	32
2.3.2 <i>Hybrid QM/MM</i> .....	34
2.3.3 <i>Environment effects</i> .....	35
2.3.4 <i>QM/MM<sup>POL</sup></i> .....	36
<b>References</b> .....	<b>37</b>
<b>3. Results and Discussion</b> .....	<b>41</b>
<b>Section A</b> .....	<b>42</b>
3.1 <b>Iron biomimetics (Paper I)</b> .....	<b>42</b>

<b>3.2 Inhibition mechanisms of enzymes involved in replication .....</b>	<b>43</b>
3.2.1 <i>RNase A inhibition by arsenoplatin-1 (Paper II)</i> .....	43
3.2.2 <i>Main protease SARS-CoV-2 inhibitors (Paper III)</i> .....	45
<b>3.3 Viperin: a radical-SAM enzyme with antiviral activity .....</b>	<b>47</b>
3.3.1 <i>Structural investigation and ionization potential determination (Paper IV)</i> .....	48
3.3.2 Viperin product (ddhCTP) as inhibitor of SARS-CoV-2 (Wp ).....	49
<b>Section B.....</b>	<b>51</b>
<b>3.1 Antioxidant (Paper V and VI).....</b>	<b>51</b>
<b>References .....</b>	<b>53</b>

### Section A

Paper I.....	i
Paper II .....	ii
Paper III.....	iii
Paper IV.....	iv
Wp (Work in progress).....	v

### Section B

Paper V.....	vi
Paper VI.....	vii
Acronyms and abbreviations.....	viii

### Addendum

Supporting information – Paper I.....	ix
Supporting information – Paper II.....	x
Supporting information – Paper III.....	xi
Supporting information – Paper IV.....	xii
Supporting information – Wp (Work in progress).....	xiii
Supporting information – Paper VI.....	xiv

## Sintesi in italiano

I processi biochimici avvengono su diverse scale di lunghezza e tempo che vanno da pochi angstrom, la dimensione del sito attivo delle proteine, dove le fasi di innesco ultrarapido di solito hanno luogo, fino alle cellule e agli organi, dove i loro effetti macroscopici sono rilevabili a occhio nudo. Questa organizzazione gerarchica inerente è responsabile della complessità della materia vivente: un singolo processo coinvolge una cascata multiscala di eventi la cui descrizione richiede la combinazione di diverse metodologie.

Le simulazioni multiscala forniscono un potente approccio per risolvere in modo computazionale la dinamica funzionale, l'energetica e i cambiamenti strutturali a livello di singola molecola e quindi per sondare i meccanismi molecolari nella conversione energetica biologica. Mentre i metodi della chimica quantistica (QM) sono necessari per studiare l'energetica e le dinamiche legate alle trasformazioni chimiche, i cambiamenti conformazionali accoppiati a queste ultime su una scala temporale più lunga richiedono l'integrazione di simulazioni di dinamica molecolare atomistica (MD) più approssimative.

Nel presente lavoro di tesi di dottorato, i sistemi studiati sono stati trattati applicando l'approccio computazionale più adeguato alla loro complessità strutturale e funzionale. Così sistemi molecolari coinvolti in reazioni che si svolgono in tempi molto brevi (dell'ordine dei picosecondi) sono stati investigati prevalentemente applicando metodi basati sulla Teoria Funzionale della Densità (DFT). Questo è stato il caso del sistema biomimetico contenente ferro non-eme e delle molecole antiossidanti coinvolte in multipli meccanismi: trasferimento di atomo di idrogeno e trasferimento di un singolo elettrone.

Con la DFT, infatti i processi chimici che coinvolgono la riorganizzazione degli elettroni, come il trasferimento di carica o la formazione/rottura di legami covalenti, possono essere trattati in tempi di calcolo ragionevoli.

Per i sistemi più complessi di natura proteica, come i quattro enzimi studiati, in cui si è voluto prendere in considerazione anche l'effetto della matrice proteica sul processo chimico in atto nel sito attivo dovuto alle interazioni elettrostatiche e alla polarizzazione tra il sito attivo e gli atomi circostanti, sono stati adoperati approcci MD e QM.

Metodi ibridi quantistici/classici (QM/MM) sono stati usati nella descrizione dell'enzima viperina. Gli approcci QM/MM permettono di trattare una piccola porzione del sistema con metodi quantomeccanici e quindi un'accuratezza molto elevata, mentre, i contributi dell'ambiente circostante sono considerati con calcoli MM più semplici e veloci.

L'obiettivo di questa tesi di dottorato è quello di chiarire i meccanismi di attivazione ed inibizione coinvolti nell'attività di sistemi semplici e complessi utilizzando consolidati approcci computazionali multi-scala.

## Synthèse en français

Les processus biomoléculaires sont difficiles à étudier en raison de l'organisation hiérarchique multi-échelle intrinsèquement complexe dans laquelle ils se produisent. Les méthodes théoriques sont désormais largement utilisées dans la compréhension des processus biochimiques. La précision croissante des approches computationnelles et des architectures informatiques permettent de décrire correctement les voies de réaction impliquées dans les processus biologiques, sans beaucoup de limitations autres que celles strictement liées au protocole théorique adopté. La biochimie computationnelle a aujourd'hui un pouvoir prédictif et fournit des données très détaillées sur des processus difficiles à analyser expérimentalement. Une application très importante de la modélisation moléculaire est la conception rationnelle de médicaments basée sur la connaissance d'une cible biologique (conception de médicaments assistée par ordinateur). Les méthodes de dynamique moléculaire (MD) basées sur les principes classiques de la mécanique moléculaire ont permis d'améliorer les notions et de s'éloigner de la rigidité irréaliste imposée par la détermination structurale cristallographique. La théorie de la fonction de densité (DFT) est une méthode théorique largement utilisée pour effectuer des calculs quantiques dans la recherche biochimique. Avec la DFT, les processus chimiques impliquant une réorganisation des électrons, comme le transfert de charge ou la formation/rupture de liaisons covalentes, peuvent être traités dans des temps de calcul raisonnables. Aujourd'hui, des systèmes comprenant jusqu'à des milliers d'atomes sont étudiés par des simulations *in silico*. Des méthodes hybrides quantiques/classiques (QM/MM) sont en effet utilisées pour reproduire de grands systèmes moléculaires. Les approches QM/MM permettent de traiter une petite partie du système avec des méthodes de mécanique quantique et donc une très grande précision, tandis que les contributions du milieu environnant sont prises en compte avec des calculs MM plus simples et plus rapides.

L'objectif de cette thèse est d'élucider les mécanismes d'activation et d'inhibition impliqués dans l'activité de certaines enzymes en utilisant des approches computationnelles multi-échelles originales. Tous les systèmes étudiés sont des processus biologiques multi-étapes impliqués dans des pathologies, telles que le cancer ou des infections virales. La clarification des mécanismes (réaction ou inhibition) est donc d'une importance fondamentale pour le développement de stratégies thérapeutiques. Le **chapitre 1** fournit le contexte biochimique général permettant de comprendre les systèmes biologiques étudiés dans cette thèse, tandis que dans le **chapitre 2**, nous donnons un bref aperçu du contexte théorique et des méthodologies utilisées dans ce travail. Dans le **chapitre 3**, nous rendons compte de la production scientifique de la thèse, résumée ci-dessous. L'article I rapporte l'étude du mécanisme de réaction d'un catalyseur non hémérique mononucléaire à base de fer biomimétique lié à l'activation de l'oxygène par la réduction de 2 électrons dans la décarboxylation oxydative d'hydroxyacides comme réducteurs. La partie suivante porte sur les mécanismes d'inhibition des enzymes impliqués dans la réplication cellulaire. En particulier, le processus de métallisation de la

ribonucléase A par un complexe contenant du Pt-As a été élucidé par des calculs DFT à l'aide d'un modèle capable de représenter de manière assez réaliste l'environnement du site de liaison. Le document III présente le mécanisme d'inhibition de la protéase principale (Mpro) du SRAS-CoV-2 en appliquant des méthodes de calcul allant de la dynamique moléculaire classique et des simulations de *docking* moléculaire aux calculs DFT et propose deux molécules comme inhibiteurs prometteurs de Mpro. Le document IV se concentre sur l'étude de l'enzyme antivirale naturelle, la viperin, une protéine inhibitrice du virus inducible par l'interféron, associée au réticulum endoplasmique, qui possède dans son site actif un cluster fer-soufre impliqué dans le processus catalytique. Il s'agit d'un membre de la superfamille des enzymes SAM radicales qui utilisent la SAM comme substrat pour convertir la SAM de façon stœchiométrique en Met et en 5'-désoxyadénosine. Les caractéristiques structurales de l'enzyme ont été analysées en profondeur par des simulations de dynamique moléculaire. Le potentiel d'ionisation a été évalué par des approches computationnelles avancées. En outre, comme le produit catalytique de la vipérine est une nucléobase modifiée (ddhCTP) basée sur une structure pyrimidine, elle a été testée en tant qu'inhibiteur dépendant de l'ARN-polymérase du SRAS-CoV-2 avec une nature de terminateur de chaîne (Wp).

La section B, qui comprend les articles V et VI, concerne l'étude théorique des composés naturels ayant des propriétés antioxydantes essentielles et les mécanismes de réaction impliqués dans le piégeage des radicaux libres pour minimiser le stress oxydatif.

## Acknowledgements

First of all, I would like to thank the PROMOCS members of the Department of Chemistry and Chemical Technologies, University of Calabria, for their moral support and friendship.

I sincerely thank past and present members of the TheoSim group of Institut de Chimie Physique, University of Paris-Saclay, for welcoming me and making me feel at home.

I'd want to thank Prof. Tiziana Marino for supervising my work during these three years, for giving me friendly advice and opportunities of growth.

My sincere thanks to Aurelien de la Lande for the good teachings, stimulating discussions and for his patience.

I thank the Ministry of Foreign Affairs and International Cooperation of Italy and Professor Annia Galano of the Universidad Autonoma Metropolitana Iztapalapa for give me the opportunity to participate at Italy-Mexico Joint Mobility Project (ID MX18MO10).

My gratitude to the Italo-French University (UIF/UFI) for financing to support academic bi-national projects between France and Italy within the program VINCI cap. II - Mobility grants for co-tutored PhD theses.





## List of publications

The results in the present thesis are based on the following papers:

**I. The generation of the oxidant agent of a mononuclear nonheme Fe(II) biomimetic complex by oxidative decarboxylation. A DFT investigation**

A. Parise, M.C. Muraca, N. Russo, M. Toscano, T. Marino

*Molecules*, **2020**, 25 (2), 328.

**II. The platination mechanism of RNase A by arsenoplatin: insight from the theoretical study**

A. Parise, N. Russo, T. Marino

*Inorganic Chemistry Frontiers*, **2021**, 8 (7), 1795.

**III. The Se-S Bond Formation in the Covalent Inhibition Mechanism of SARS-CoV-2 Main Protease by Ebselen-like Inhibitors: A Computational Study**

A. Parise, I. Romeo, N. Russo, T. Marino

*International journal of molecular sciences*, **2021**, 22 (18), 9792.

**IV. New Insights on the redox Properties of Viperin from Advanced Numerical Simulations**

A. Parise, N. Russo, T. Marino, A. de la Lande

*Manuscript in submission to Biomolecules*

**Wp. (Work in progress) The catalytic product of the natural antiviral viperin ddhCTP as inhibitor of SARS-CoV-2 RNA-polymerase-RNA-dependent: a computational insight**

A. Parise, G. Ciardullo, M. Prejanò, A. de la Lande, T. Marino

*Manuscript in preparation*

**V. The antioxidant capability of higenamine: Insights from theory**

I. Romeo, A. Parise, A. Galano, N. Russo, J.R. Alvarez-Idaboy, T. Marino

*Antioxidants*, **2020**, 9 (5), 358.

**VI. Quantum mechanical predictions of the antioxidant capability of moracin C isomers**

A. Parise, B.C. De Simone, T. Marino, M. Toscano, N. Russo

*Frontiers in Chemistry*, **2021**, 9, 232.

## Introduction

Investigating biomolecular processes is challenging because of the intrinsic multi-scale hierarchical organization in system size and time in which chemical and chemical physical phenomena occur. Theoretical methods applied to biomolecules deal with the use of computation to obtain information about biological systems and their many interactions. The increasing accuracy of computational approaches and of computer architectures, allow one to properly describe reaction paths involved in both activation/inhibition enzymes, calculate electronic properties and structural dynamics conformations of biomolecules without any limitation except those connected with the adopted theoretical protocol. Currently, computational biochemistry has predictive power and provides very insightful data that are sometimes tedious to obtain from experiment. In addition to the clarification of biochemical processes at the molecular and atomistic level, a very important application of biomolecular modeling is to guide the rational drug design process based on knowledge of a biological target (Computer-Aided Drug Design). Molecular Dynamics (MD) methods based on classical Molecular Mechanics principles have improved notions of the dynamics of molecular systems and helped to move away from the unintentional rigidity imposed by crystallographic structural determination. Density Functional Theory (DFT) is a widely used theoretical method for carrying out quantum calculations in chemistry and biology research. With DFT, chemical processes involving electron reorganization such as charge transfer or formation/cleavage of covalent bonds can be addressed in more reasonable computation times than a few decades ago. Nowadays, systems encompassing up to thousands of atoms are amenable to computer simulations. Hybrid quantum/classical methods (QM/MM) are used for modeling large molecular systems. In this way, a smaller region of a larger system is treated with the most accurate QM methods, while environmental contributions are still taken into account with simpler and faster MM calculations. The objective of this PhD thesis is to elucidate activation or inhibition mechanisms involved in enzyme activity by using multiscale and original computational approaches.

This manuscript is organized into three Chapters. **Chapter 1** provides a general biochemical background to understand the actual biological systems investigated in this thesis. In **Chapter 2**, we provide a brief outline of theoretical background and methodologies involved in this work.

In **Chapter 3** we aim at elucidating the mechanisms of important biological processes, organized as follows:

### Section A

- We report a mechanistic DFT study of the reaction promoted by a non-heme iron-based biomimetic catalyst (**Paper I**);

- In the second part, inhibition mechanisms of enzymes involved in cell replication are disclosed and shown. In particular, the metalation process of Ribonuclease-A by a new anticancer drug, containing arsenic (III) and platinum (II), has been elucidated by DFT calculations (**Paper II**) and the inhibition of the main protease ( $M^{pro}$ ) of SARS-CoV-2 has been studied by using a multiscale combination of classical Molecular Dynamics simulations, Molecular Docking and DFT computations (**Paper III**);
- the third part is devoted to the study of the natural antiviral enzyme viperin (virus inhibitory protein, endoplasmic reticulum-associated, interferon-inducible). Structural features of viperin have been analyzed by MD simulations and then the ionization potential of an iron-sulfur cluster, present in the active site and directly involved in the catalytic process, has been evaluated by advanced computational approaches (**Paper IV**). In addition, since the catalytic product of viperin is a modified nucleobase (ddhCTP) having a structure-like antiviral drug, it has been tested to predict its potential chain terminator activity towards RNA-polymerase-RNA-dependent of SARS-CoV-2 (**Wp**).

## Section B

- In the last section, we report DFT calculations to predict the main antioxidant properties and reaction mechanisms involved in free radical scavenging of some naturally occurring chemical compounds (Higenamin and Moracin) to minimize the oxidative stress (**Paper V and VI**).

All the investigated systems are engaged in multistep biological processes so elucidating the mechanisms (reaction or inhibition) is of crucial importance.



## 1. Biological Background

Complex biosystems represent a large family of molecules with different members, ranging from proteins, enzymes to nucleic acids, lipids, sugars, that have attracted the study of numerous scientists for decades. The term "enzyme" was coined in 1878 by Fredrich Wilhelm Kuhne, to emphasize that in yeast "something" catalyzes the reactions of fermentation.<sup>1</sup> Enzymes have the characteristic of acting as biological catalysts, they have the ability to reduce the activation energy of a reaction, modifying its path to make a kinetically slow process faster.

In 1963, the first amino acid sequence of an enzyme, ribonuclease A from bovine pancreas,<sup>2</sup> was known in its entirety and in 1965 an X-ray structure of an enzyme, lysozyme from egg white, was obtained for the first time. Since then, thousands of enzymes have been purified and structures were characterized and the number is rapidly increasing. Since their discovery, several theories have been proposed in order to rationalize, explain and predict enzymatic activity and its peculiarities. Catalysts allow a faster reaction through alternative pathways of intermediates and transition states with different energies, significantly lower than the non-enzyme-assisted reaction. The catalytic action of enzymes differs from that of non-enzymatic catalysts because of their very high catalytic efficiency, their specificity towards a restricted group of molecules and their possibility of variation from a state of low or no activity to one of maximum activity. All of these characteristics are linked to the fascinating and complex enzyme structures. Dealing with enzymes is a challenge: they have structures made of tens of hundreds of atoms with a complex three-dimensional organization that act on transition states with very short life times.

Nowadays, with the advent of high-performance computers and the development of dedicated algorithms, computational molecular simulations and modelling are changing the science of enzymology. *In silico* studies can provide detailed, atomic-level insight into the fundamental mechanisms of biological catalysts. This chapter aims at illustrating the background knowledge that aids understanding of the work described in the thesis.

### 1.1 Enzyme catalysis

Enzymes can be defined, with some exceptions (Ribozymes), as specific catalysts, predominantly proteins in Nature. These molecular machines catalyze highly chemo-, stereo- and regio-selective reactions. A protein is characterized by complex primary, secondary, ternary and quaternary structures. Protein scaffolds have evolved to allocate the active site, the catalytic center where the reactions takes place. According to the reactions catalyzed, it is customary to distinguish six main enzymes classes, defining EC (Enzyme Commission):<sup>3</sup>

- EC 1, oxidoreductases catalyze oxidation/reduction reactions;
- EC 2, transferases transfer functional groups;
- EC 3, hydrolases catalyze the hydrolysis reaction;

- EC 4, lyases catalyze breaking bond reactions, different to hydrolysis and oxidation/reduction;
- EC 5, isomerases favor isomerization reactions;
- EC 6, ligases catalyze bond formation reactions.

Other identification specifications are added to the EC, to recognize the class of the enzyme considering also the nature of the substrate. EC 1.1 is a subclass of oxidoreductases in which the number after the dot refers to the fact that enzymes act on CH-OH type donors (alcohols).

Any enzymatic reaction can be described by the reaction coordinate which passes through different intermediate steps. The equilibrium between substrate (S) and product (P), in the absence of enzyme, is governed by thermodynamic laws and it is represented by the ratio of direct and reverse reactions ( $S \rightarrow P / P \rightarrow S$ ). The intervention of the enzyme allows the alteration of the reaction stabilizing the Transition States (TS); providing alternative mechanism pathways and destabilizing substrate ground state.<sup>4-6</sup>

The process of enzymatic catalysis is reported schematically as a 3-step mechanism:



where E, S, and P represent the enzyme, substrate, and product, respectively. ES and EP are the transient enzyme complexes with the substrate (Michaelis-Menten complex) and with the product. The reaction mainly involves a molecular recognition process, the formation of the ES complex, then a series of reactions leading to the formation of the catalytic product. Interactions between the substrate functional groups and the enzyme are essential to catalyze the reaction.

### 1.1.1 Michaelis Menten kinetics

The enzymatic kinetics provides important information about the mechanism of the reaction and is useful and sometimes necessary for understanding enzyme activity. The reaction rate is derived from the rate of product formation or substrate consumption in a time unit. Michaelis and Menten, pioneers of this theory, obtained simple equations describing the pseudo-first-order kinetics of enzyme-catalyzed reactions by adopting an approach, now really well-known, where the rate of a reaction catalyzed by an enzyme is proportional to the concentration of the enzyme-substrate complex,<sup>7</sup> as shown in equation 1.2.

$$v = \frac{d[P]}{dt} = \frac{V_{max}}{K_M + [S]} \frac{k_2 [E]_0 [S]}{K_M + [S]} \quad (1.2)$$

where velocity ( $v$ ) is defined as the product concentration [P] derived with respect to time.  $v$  depends on the kinetic constant  $k_2$ , the initial enzyme concentration  $[E]_0$ , [S] the substrate concentration and the quantity  $K_M$ , that is defined as:

$$K_M = \frac{k_{-1} + k_2}{k_1} \quad (1.3)$$

where  $k_{-1}$ ,  $k_2$  and  $k_1$  refer to the reaction reported in (1.1). The *Michaelis-Menten constant* ( $K_m$ ) is a measure of the affinity of the enzyme for a specific substrate.

For each fixed enzyme concentration, the variation of the concentration of substrate S causes the initial velocity  $v$  to change according to the curve shown in figure 1.1

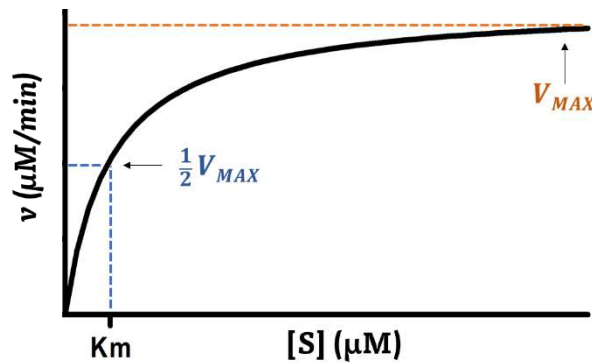


Figure 1.1 Representation of the Michaelis-Menten curve. The graph shows the trend of the initial velocity as a function of substrate concentration.

At low substrate concentrations,  $v$  is directly proportional to  $[S]$ . As the substrate concentration increases, the increment of  $v$  decreases progressively, until the asymptotically reached maximum velocity  $V_{MAX}$ . Above this limit, a further increase in  $[S]$  no longer modifies  $v$ . In equation 1.2,  $V_{MAX}$  is equal to  $k_2[E]_0$  denoting that  $V_{MAX}$  value is an important kinetic characteristic of a given "enzyme-substrate" system, dependent only on the enzyme concentration.  $K_M$  and  $V_{MAX}$  values can vary greatly from one enzyme to another; they are typical functions of each enzyme and can be used to evaluate and compare the catalytic efficiency of different enzymes.

### 1.1.2 Cofactors

In order to operate some enzymes require the presence of so-called cofactors, that are non-protein molecules. The cofactors, that can be organic units or metal ions, assist the enzyme to enable its catalytic activity.<sup>8</sup> Coenzymes are typically small and simple non-protein organic compounds, such as vitamin-derivatives, which bind specifically to enzymatic macromolecules and actively participate in catalytic biotransformations.<sup>9,10</sup> Coenzymes, as well as some metal-ion cofactors, can bind to the enzyme covalently and are classified as prosthetic groups or through weak interactions termed cosubstrates. For example, S-Adenosyl-methionine (SAM) is a common cosubstrate, utilized by three key metabolic pathways: transmethylation, transulfuration, and polyamine synthesis. SAM is defined as one of the most versatile and common cofactors involved in a wide variety of enzymatic reactions.<sup>11</sup>



The most known general catalytic mechanisms of SAM are: methylation, decarboxylation, cyclization, ring-opening reaction and radical-based reactions in which at the start of the enzymatic reactions, one coordinated SAM molecule is cleaved to yield the 5'-deoxyadenosyl radical (5'-dAdo·) as a potent oxidant.<sup>12</sup>

The structure of SAM is shown in Figure 1.2 and will be discussed in more detail in chapter 3 since it is part of the viperin enzyme.

Inorganic cofactors are essentially metal ions, stand-alone or organized in polyatomic clusters, bound to amino acid residues or coenzymes, inside the active site.<sup>13,14</sup> Almost half of all enzymes require the presence of a metal atom to function.<sup>15</sup> Specific functional groups of some amino acids, such as the imidazole of histidine, the carboxylate of glutamate/aspartate, and the thiolate groups of cysteine, are the most commonly observed ligands in metalloproteins (see examples in figure 1.2).

Metal elements perform very disparate functions within biological systems. Ions, in fact, can be directly involved in the catalytic mechanism for their ability to transfer electrons, show relevant contributions to the structural stability of the protein, or can correctly orient the substrate by facilitating the interaction with another catalytic agent that doesn't participate actively in the reaction.<sup>16,17</sup> Obviously, the activity is directly related to the nature of the metal.

Small metal cations, such as  $Mg^{2+}$  possess a particularly high affinity for areas with high negative charge density; in fact, magnesium ions can be found in the active sites of RNA and DNA polymerases (Section A 3.3.2).  $Mg^{2+}$  is present within enzymes to catalyze peptide bond formation, self-cleavage and transesterifications of phosphodiester.<sup>18</sup>

A metallic structural motif, the iron-sulfur cluster, is present in a wide range of metalloproteins such as ferredoxin, NADH dehydrogenase and Coenzyme Q-cytochrome c reductase of the electron transport chain. It is there in cells of all species and probably derives from primitive life forms but its versatility has allowed it to be maintained as part of the metabolism of more evolved organisms.<sup>19</sup> The majority of proteins with Fe-S centers are involved in charge transfer processes. Iron is a transition metal, which can form cations with incomplete valence orbitals. Thus, iron has variable valence, a property favorable to the mechanism of electron transfer in proteins.

Iron and sulfur atoms are grouped in different stoichiometric ratios. 4Fe-4S centers have four iron atoms with four sulfur atoms bridging them. The iron is generally coordinated by the side chains of four cysteine residues.

We take the example of viperin. This enzyme, topic of the next section A 3.3, presents a central domain where the catalytic pocket consists of a CxxxCxxC motif, characteristic of radical S-adenosylmethionine enzymes with cysteine residues within the CxxxCxxC motif ligating three iron of the [4Fe-4S] cluster, the fourth iron of the cluster coordinates an oxygen and a nitrogen atom of the SAM methionine part.

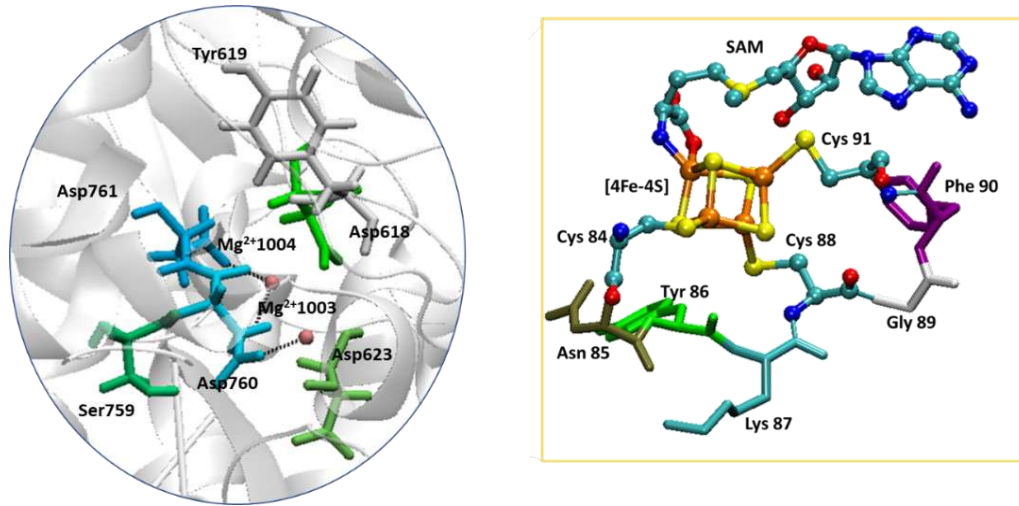


Figure 1.2 Examples of inorganic and organic cofactors. We show in the circle the catalytic site of the RNA polymerase RNA-dependent enzyme, which contains 2 catalytic magnesium ions. In the square is shown the active site of the enzyme viperin: it contains the 4Fe-4S cluster (inorganic cofactor) and S-Adenosylmethionine (coenzyme).

## 1.2 Inhibition mechanisms

Catalyst inhibition or more specifically enzymatic inhibition takes place when the catalytic agent is unable to process substrate molecules as usual. Inhibition occurs when a compound capable of interfering with the enzyme, induces various rearrangements in the structure of the enzyme itself. There are different modes of inhibition (see figure 1.3) depending on when the enzyme-inhibitor (EI) complex is formed and where the interaction occurs.<sup>20</sup>

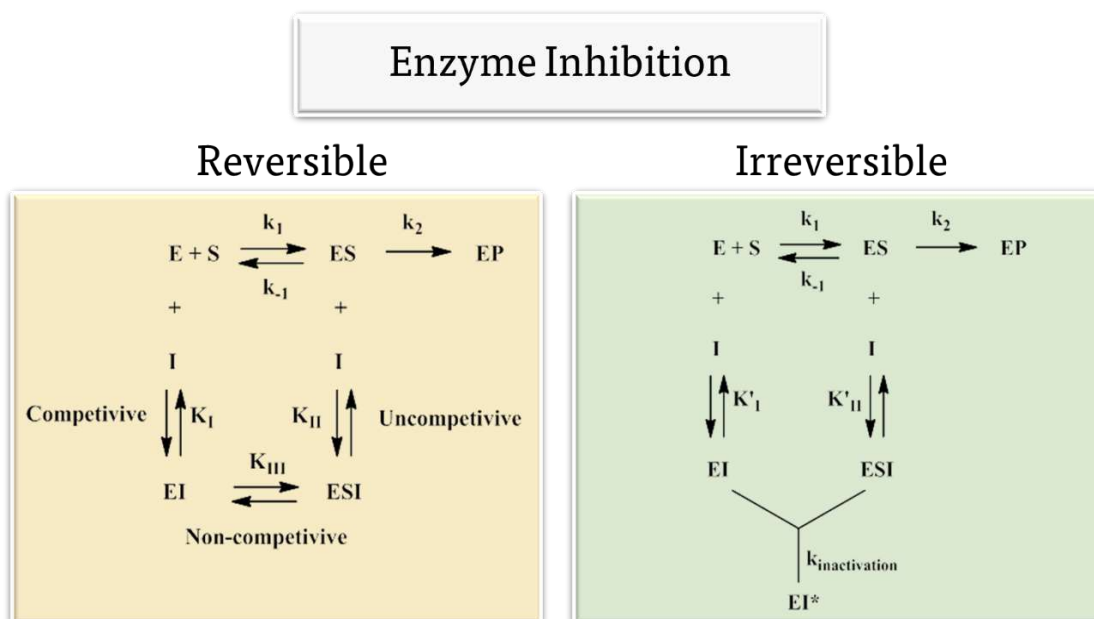


Figure 1.3 Representation of reversible and irreversible inhibition in yellow and green respectively.

### 1.2.1 Reversible inhibition

Reversible inhibitors bind to enzymes with non-covalent intermolecular forces. Since weak interactions are formed, reversible inhibitors do not permanently modify the enzyme and they can be removed by dilution.<sup>21</sup> In reversible inhibition, the inhibitor binds the enzyme (competitive mechanism) or the enzyme-substrate complex (non-competitive mechanism) forming the enzyme-inhibitor (EI) or enzyme-substrate-inhibitor (ESI) complex, respectively.<sup>22,23</sup>

Competitive inhibitor is so defined because it competes with the substrate for access to the active site of the enzyme. Binding of S or I to the enzyme is a competitive and mutually exclusive process. The inhibitor occupies the active site because it has a structural similarity to the substrate.

In the case of competitive reversible inhibition, the Michaelis-Menten equation is as follows:

$$v = \frac{V_{MAX}[S]}{\alpha K_m + [S]} \quad (1.4)$$

where,

$$\alpha = 1 + \frac{[I]}{K_I} \quad ; \quad k_I = \frac{[E][I]}{[EI]} \quad (1.5)$$

$\alpha K_m$ , defined as apparent-constant, corresponds to the experimentally observed  $K_m$  in the presence of enzyme, substrate, and inhibitor. Once  $V_{MAX}$  (Figure 1.1) is reached, the increase in apparent  $K_m$  confirms that the inhibition is competitive.<sup>22</sup>

Uncompetitive inhibition is observed when the inhibitor binds a site, defined as allosteric, that is not the catalytic pocket. An uncompetitive inhibitor is able to bind only to the ES complex leading to the IES formation.<sup>22</sup>

For the uncompetitive mechanism, the Michaelis-Menten equation is:

$$v = \frac{V_{MAX}[S]}{K_m + \alpha'[S]} \quad (1.6)$$

$$\alpha' = 1 + \frac{[I]}{k_{II}} \quad ; \quad k_{II} = \frac{[ES][I]}{[ESI]} \quad (1.7)$$

equation (1.6) shows with high substrate concentrations that  $v$  is very close to  $V_{MAX}/\alpha$ .

Inhibition caused by uncompetitive inhibitors with a single substrate is rare, but is seen more often in reactions with multiple substrates.

In non-competitive inhibition, the binding of the inhibitor to the enzyme reduces its activity but does not affect the binding of substrate, the particularity of non-competitive inhibition mechanism is that inhibitor binds to the allosteric site independently of substrate binding.<sup>24</sup>

Finally, I mention the case of the mixed inhibition that could be considered as a combination of all mechanisms previously described. As with uncompetitive inhibition, mixed inhibition cannot be removed by increasing the concentration of substrate.

Up to now, can be easily accessed using the well known double reciprocal diagram, a graphical representation of the Lineweaver-Burk equation of enzyme kinetics, described by Hans Lineweaver and Dean Burk.<sup>25</sup>

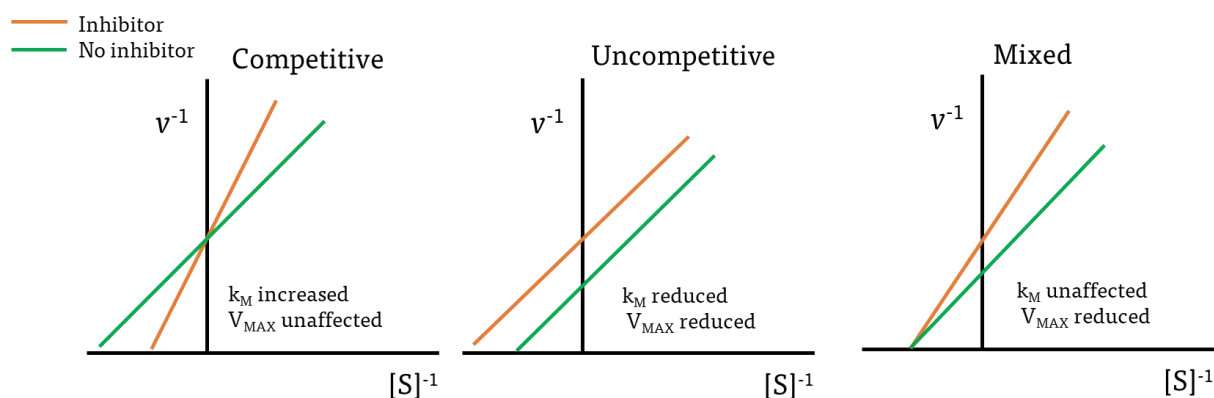


Figure 1.4 Enzyme inhibition displayed using double reciprocal Lineweaver-Burk plots.

The Lineweaver-Burk plot is correct when the enzyme kinetics obey ideal second-order kinetics as shown in Figure 1.4, however, nonlinear regression is necessary for systems that do not behave ideally, hence real systems. Despite this it is a useful technique to distinguish competitive, uncompetitive and mixed inhibition.<sup>26</sup>

### 1.2.2 Irreversible inhibition

The irreversible inhibitor binds to the enzyme covalently, transforming the functional groups essential for the activity of the enzyme. Irreversible inhibitors usually permanently modify an enzyme, thus resulting in an inactivated protein. Irreversible inhibitors often contain reactive functional groups such as mustards nitrogen, aldehydes, haloalkanes, alkenes, Michael acceptors, phenyl sulfonates or fluorophosphates.<sup>27</sup> These nucleophilic groups react with amino acid side chains to form covalent adducts. Irreversible inhibitors are also used as a means of characterizing reaction mechanisms and identifying functional groups involved in enzyme catalysis.

In general, irreversible enzyme inhibitors are designed to react with key residues in the catalytic pocket, however, there are also excellent examples of irreversible allosteric inhibitors.<sup>9,28,29</sup>

In general, covalent inhibition is a two-step process (scheme 1.1): the first step sees the reversible association of the inhibitor with its target macromolecule in such a way that a weakly electrophilic “warhead”, is brought into the vicinity of an appropriately positioned nucleophilic residue on the protein; in the second step the spontaneous reaction occurs between the participating functional groups of ligand and protein to produce the inactivated and covalently modified protein.<sup>30</sup>



*Scheme 1.1 General scheme of covalent interaction between a small molecule and its target.*

In section A 3.2 I will discuss covalent mechanisms of inhibition at the atomistic level. A special class of irreversible inhibitors are the suicide inactivators.<sup>31</sup> These inhibitors are stable until they bind to the active site of a particular enzyme. They carry on the first steps of the canonical enzymatic reaction, but instead of being converted into the normal product, the inactivator is converted into a very reactive compound that binds irreversibly with the enzyme. This type of inhibitor is widely used in the pharmaceutical field. Suicide inhibitor must be specific for a particular enzyme and becomes reactive only when it has reached the active site of the enzyme. Thus, new therapeutic agents are obtained that are specific, effective and free of side effects.

### 1.2.3 Inhibitors as potential drugs

Enzymes are used in several fields of research and at different levels of production: are used in industry, food processing and their inhibition/activation could be involved in treatments of several diseases, like cancer and viral infection. Effectively, time and costs associated with new drug discovery have been significantly reduced since using enzyme structure-based approaches for the discovery of new pharmacologically active molecules.

<sup>32</sup> Inhibitors of enzymes have been extremely successful in the war against infections. In fact, blocking the activity of a crucial enzyme for a microorganism clearly means killing the cell or preventing it from growing. This aspect has been abundantly used in the field of antivirals, an example of which are the drugs against HIV that inhibit the activity of an enzyme necessary for the replication of the virus, the protease (HIV-1 protease).<sup>33,34</sup>

Another example is given by the “chain terminators”, molecules that act as ‘false substrate’ and are incorporate into the growing of RNA or DNA chain during replication by polymerase. Once they have been added, the chain can no longer be extended and chain growth is stopped, causing the blockage of cell duplication of the pathogen. Nucleoside analogs that inhibit polymerases represent an important group of antiviral agents.<sup>35</sup>

In this thesis, the concept of inhibition as a strategy for designing of pharmacologically active molecules or understanding their action mechanism will be explored in detail. In fact, sections A 3.2 and 3.2 will focus on *in silico* study of inhibition mechanisms (covalent, non-covalent and chain terminator type) as strategies for developing and

understanding the mechanism of actions of molecules that have shown *in vitro* anticancer or antiviral activity.

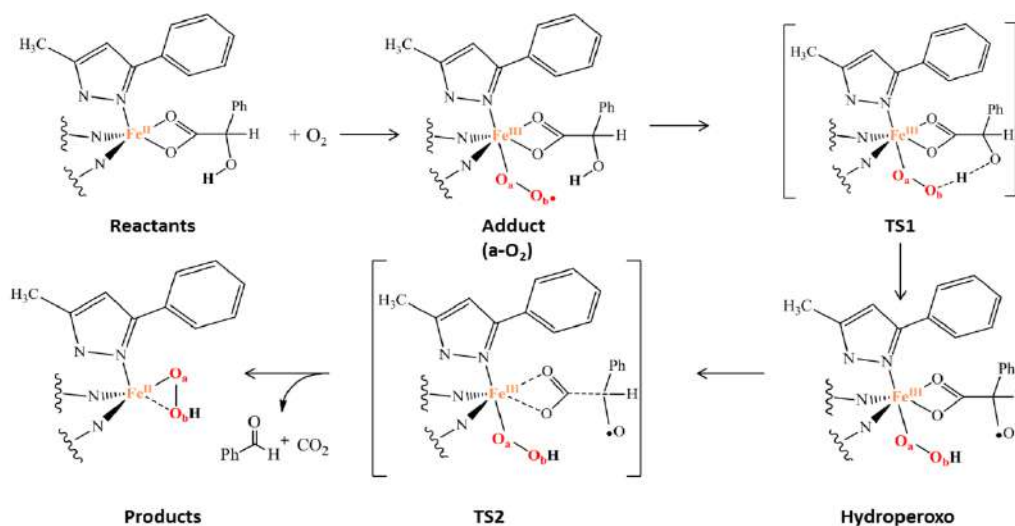
### 1.3 Enzymes as technological inspiration: biomimetic systems

Bioinspiration refers to a process by which one imitates and reproduces a phenomenon observed in biological systems. It is used in particular with reference to processes characterized by strong properties of self-organization, stereospecificity, molecular recognition, such as are often found in biological systems, as shown previously, especially in enzymes. In recent years, the use of biomimetic catalysts has increased significantly not only because of their relative high selectivity, but because they can operate under very mild conditions and are adapted to different types of substrates.<sup>36,37</sup>

The possibility of realizing even more sophisticated biomimetic processes has greatly increased with the development of supramolecular chemistry and is considered of interest for applications in the pharmacological field.<sup>38</sup>

For example, bioinspired catalysis using peroxides or peracids for the oxidation of hydrocarbon substrates is very common and well characterized.

Encouraged by the new reactions catalyzed by Rieske dioxygenases<sup>39,40</sup> sustainable mononuclear nonheme iron complexes have been proposed for bioinspired oxidation catalysis (see scheme 1.2). Dioxygenases are involved in the degradation of a wide range of natural and synthetic compounds by incorporating both atoms of O<sub>2</sub> into substrates. In chapter 3.1 (section A) is reported the study of the generation of the oxidant agent of a mononuclear non-heme Fe(II) biomimetic complex by oxidative decarboxylation.



*Scheme 1.2 Proposed mechanism for the oxidative decarboxylation promoted by the non-heme biomimetic complex model.*

### 1.4 Prevention of oxidative damage to bio-molecules

Oxidative stress is a redox disorder that includes all the alterations that are produced in tissues, cells and biomolecules when they are exposed to an excess of oxidizing agents.

A state of oxidative stress results from the action of highly reactive unstable chemicals (prevalently reactive oxygen species and reactive nitrogen species, ROS and RNS respectively), non-radical pro-oxidant agents (such as hydrogen peroxide) and ionizing radiation. If the antioxidant defenses of the cell and the body are insufficient to maintain the redox state in balance and the stress situation is prolonged, the excess of ROS and RNS can generate vital alterations that in the long run become irreversible.

Substances that can contrast, slow down or neutralize the formation of free radicals are the antioxidants.

There is a large variety of natural antioxidants present in food. Among them, those often called phenolics include more than 8,000 compounds, and they all share the structural feature of presenting a phenol moiety. A classification of such compounds depends on the number of phenol subunits: simple phenols contain only one phenol functionality (e.g., phenolic acids), whereas polyphenols contain two (e.g., flavonoids and stilbenes) or more phenol subunits (e.g., tannins).

Antioxidants have been reported to exert numerous beneficial effects on human health, including anticancer<sup>41</sup>, antibacterial<sup>42</sup> and antiviral<sup>43</sup> properties.

Quantum mechanics-based studies (described in Chapter 2) have become an important part of current chemical investigations because they can provide valuable physicochemical insights that help in the understanding of chemical processes at the molecular level.<sup>44</sup>

The possible action mechanisms of the antioxidants analyzed are: hydrogen-atom transfer (HAT), proton-coupled electron transfer (PCET), radical adduct formation (RAF), single electron transfer (SET), sequential proton-loss electron transfer (SPLET). Examples of the computational chemistry used to investigate reaction mechanisms by which antioxidants can act to contrast the effect of free radicals are reported in section B. Comparison between theoretical and experimental data confirm that modern theoretical tools are able to explain experimental facts and to predict the antioxidant activity of some molecules based on the comparison.



## References

- (1) Heckmann, C. M.; Paradisi, F. Looking Back: A Short History of the Discovery of Enzymes and How They Became Powerful Chemical Tools. *ChemCatChem* **2020**, *12*(24), 6082–6102.
- (2) Crestfield, A. M.; Stein, W. H.; Moore, S. On the Preparation of Bovine Pancreatic Ribonuclease A. *J. Biol. Chem.* **1963**, *238*, 618–621.
- (3) Thompson, R. H. S. Classification and Nomenclature of Enzymes. *Science* **1962**, *137*(3528), 405–408.
- (4) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106*(8), 3210–3235.
- (5) Fisher, M. Lehninger Principles of Biochemistry, 3rd Edition; By David L. Nelson and Michael M. Cox. *Chem. Educ.* **2001**, *6*(1), 69–70.
- (6) Benkovic, S. J.; Hammes-Schiffer, S. A Perspective on Enzyme Catalysis. *Science* **2003**, *301*(5637), 1196–1202.
- (7) Michaelis, L.; Menten, M. M. L. The Kinetics of Invertase Action. *The Kinetics of Invertase Action*. 49th ed. **1913**, 333.
- (8) Nagel, B.; Dellweg, H.; Gierasch, L. M. Glossary for chemists of terms used in biotechnology (IUPAC Recommendations 1992). *Pure Appl. Chem.* **1992**, *64*(1), 143–168.
- (9) Weisner, J.; Gontla, R.; van der Westhuizen, L.; Oeck, S.; Ketzer, J.; Janning, P.; Richters, A.; Mühlenberg, T.; Fang, Z.; Taher, A.; Jendrossek, V.; Pelly, S. C.; Bauer, S.; van Otterlo, W. A. L.; Rauh, D. Covalent-Allosteric Kinase Inhibitors. *Angew. Chem. Int. Ed.* **2015**, *54*(35), 10313–10316.
- (10) Fischer, J. D.; Holliday, G. L.; Rahman, S. A.; Thornton, J. M. The Structures and Physicochemical Properties of Organic Cofactors in Biocatalysis. *J. Mol. Biol.* **2010**, *403*(5), 803–824.
- (11) Lu, S. C. S-Adenosylmethionine. *Int. J. Biochem. Cell Biol.* **2000**, *32*(4), 391–395.
- (12) Sun, Q.; Huang, M.; Wei, Y. Diversity of the Reaction Mechanisms of SAM-Dependent Enzymes. *Acta Pharm. Sin. B* **2021**, *11*(3), 632–650.
- (13) Andreini, C.; Bertini, I.; Rosato, A. Metalloproteomes: A Bioinformatic Approach. *Acc. Chem. Res.* **2009**, *42*(10), 1471–1479.
- (14) Thomson, null; Gray, null. Bio-Inorganic Chemistry. *Curr. Opin. Chem. Biol.* **1998**, *2*(2), 155–158.
- (15) Finkelstein, J. Metalloproteins. *Nature* **2009**, *460*(7257), 813–813.
- (16) Waldron, K. J.; Rutherford, J. C.; Ford, D.; Robinson, N. J. Metalloproteins and Metal Sensing. *Nature* **2009**, *460*(7257), 823–830.
- (17) Mikšovská, J.; Larsen, R. W. Structure-Function Relationships in Metalloproteins. In *Methods in Enzymology*; Biophotonics, Part A; **2003**; 360, 302–329.
- (18) Pilchova, I.; Klacanova, K.; Tatarkova, Z.; Kaplan, P.; Racay, P. The Involvement of Mg<sup>2+</sup> in Regulation of Cellular and Mitochondrial Functions. *Oxid. Med. Cell. Longev.* **2017**, *2017*, e6797460.
- (19) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. Structure, Function, and Formation of Biological Iron-Sulfur Clusters. *Annu. Rev. Biochem.* **2005**, *74*(1), 247–281.
- (20) Cleland, W. W. The Kinetics of Enzyme-Catalyzed Reactions with Two or More Substrates or Products: II. Inhibition: Nomenclature and Theory. *Biochim. Biophys. Acta BBA - Spec. Sect. Enzymol. Subj.* **1963**, *67*, 173–187.
- (21) Winteringham, F.; Fowler, K. Substrate and Dilution Effects on the Inhibition of Acetylcholinesterase by Carbamates. *Biochem. J.* **1966**, *101*(1), 127–134.
- (22) Cornish-Bowden, A. A Simple Graphical Method for Determining the Inhibition Constants of Mixed, Uncompetitive and Non-Competitive Inhibitors. *Biochem. J.* **1974**, *137*(1), 143–144.
- (23) Srinivasan, B.; Tondast-Navaei, S.; Skolnick, J. Ligand Binding Studies, Preliminary Structure-Activity Relationship and Detailed Mechanistic Characterization of 1-Phenyl-6,6-Dimethyl-1,3,5-Triazine-

- 2,4-Diamine Derivatives as Inhibitors of Escherichia Coli Dihydrofolate Reductase. *Eur. J. Med. Chem.* **2015**, *103*, 600–614.
- (24) Delaune, K. P.; Alsayouri, K. Physiology, Noncompetitive Inhibitor. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), **2021**.
- (25) Lineweaver, H.; Burk, D. The Determination of Enzyme Dissociation Constants. *J. Am. Chem. Soc.* **1934**, *56*(3), 658–666.
- (26) Hamilton, B. K.; Gardner, C. R.; Colton, C. K. Effect of Diffusional Limitations on Lineweaver-Burk Plots for Immobilized Enzymes. *AIChE J.* **1974**, *20*(3), 503–510.
- (27) 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? *J. Med. Chem.* **2019**, *62*(23), 10497–10525.
- (28) Magnaghi, P.; D'Alessio, R.; Valsasina, B.; Avanzi, N.; Rizzi, S.; Asa, D.; Gasparri, F.; Cozzi, L.; Cucchi, U.; Orrenius, C.; Polucci, P.; Ballinari, D.; Perrera, C.; Leone, A.; Cervi, G.; Casale, E.; Xiao, Y.; Wong, C.; Anderson, D. J.; Galvani, A.; Donati, D.; O'Brien, T.; Jackson, P. K.; Isacchi, A. Covalent and Allosteric Inhibitors of the ATPase VCP/P97 Induce Cancer Cell Death. *Nat. Chem. Biol.* **2013**, *9*(9), 548–556.
- (29) Cross, D. A. E.; Ashton, S. E.; Ghiorghiu, S.; Eberlein, C.; Nebhan, C. A.; Spitzler, P. J.; Orme, J. P.; Finlay, M. R. V.; Ward, R. A.; Mellor, M. J.; Hughes, G.; Rahi, A.; Jacobs, V. N.; Brewer, M. R.; Ichihara, E.; Sun, J.; Jin, H.; Ballard, P.; Al-Kadhimi, K.; Rowlinson, R.; Klinowska, T.; Richmond, G. H. P.; Cantarini, M.; Kim, D.-W.; Ranson, M. R.; Pao, W. AZD9291, an Irreversible EGFR TKI, Overcomes T790M-Mediated Resistance to EGFR Inhibitors in Lung Cancer. *Cancer Discov.* **2014**, *4*(9), 1046–1061.
- (30) Baillie, T. A. Targeted Covalent Inhibitors for Drug Design. *Angew. Chem. Int. Ed.* **2016**, *55*(43),
- (31) Walsh, C. T. Suicide Substrates, Mechanism-Based Enzyme Inactivators: Recent Developments. *Annu. Rev. Biochem.* **1984**, *53*(1), 493–535.
- (32) Craig, S. P.; Eakin, A. E. Structure-Based Inhibitor Design. In *Vitamins & Hormones*; Academic Press, **2000**, *58*, 149–169.
- (33) Deeks, S. G.; Smith, M.; Holodniy, M.; Kahn, J. O. HIV-1 Protease Inhibitors: A Review for Clinicians. *JAMA* **1997**, *277*(2), 145–153.
- (34) Deeks, S. G.; Smith, M.; Holodniy, M.; Kahn, J. O. HIV-1 Protease Inhibitors: A Review for Clinicians. *JAMA* **1997**, *277*(2), 145–153.
- (35) Chien, M.; Anderson, T. K.; Jockusch, S.; Tao, C.; Li, X.; Kumar, S.; Russo, J. J.; Kirchdoerfer, R. N.; Ju, J. Nucleotide Analogues as Inhibitors of SARS-CoV-2 Polymerase, a Key Drug Target for COVID-19. *J. Proteome Res.* **2020**, *19*(11), 4690–4697.
- (36) Costas, M.; Mehn, M. P.; Jensen, M. P.; Que, L. Dioxygen Activation at Mononuclear Nonheme Iron Active Sites: Enzymes, Models, and Intermediates. *Chem. Rev.* **2004**, *104*(2), 939–986.
- (37) Talsi, E. P.; Bryliakov, K. P. Chemo- and Stereoselective CH Oxidations and Epoxidations/Cis-Dihydroxylations with H<sub>2</sub>O<sub>2</sub>, Catalyzed by Non-Heme Iron and Manganese Complexes. *Coord. Chem. Rev.* **2012**, *256*(13), 1418–1434.
- (38) Boldyreva, E. High-Pressure Studies of Pharmaceuticals and Biomimetics. Fundamentals and Applications. A General Introduction. In *High-Pressure Crystallography*; Boldyreva, E., Dera, P., Eds.; NATO Science for Peace and Security Series B: Physics and Biophysics; Springer Netherlands: Dordrecht, **2010**; 533–543.
- (39) Paria, S.; Chatterjee, S.; Paine, T. K. Reactivity of an Iron-Oxygen Oxidant Generated upon Oxidative Decarboxylation of Biomimetic Iron(II)  $\alpha$ -Hydroxy Acid Complexes. *Inorg. Chem.* **2014**, *53*(6), 2810–2821.
- (40) Chen, K.; Costas, M.; Kim, J.; Tipton, A. K.; Que, L. Olefin Cis-Dihydroxylation versus Epoxidation by Non-Heme Iron Catalysts: Two Faces of an Fe(III)-OOH Coin. *J. Am. Chem. Soc.* **2002**, *124*(12), 3026–3035.
- (41) Roleira, F. M. F.; Tavares-da-Silva, E. J.; Varela, C. L.; Costa, S. C.; Silva, T.; Garrido, J.; Borges, F. Plant Derived and Dietary Phenolic Antioxidants: Anticancer Properties. *Food Chem.* **2015**, *183*, 235–258.
- (42) Widsten, P.; Cruz, C. D.; Fletcher, G. C.; Pajak, M. A.; McGhie, T. K. Tannins and Extracts of Fruit Byproducts: Antibacterial Activity against Foodborne Bacteria and Antioxidant Capacity. *J. Agric. Food Chem.* **2014**, *62*(46), 11146–11156.

- (43) Panchal, R. G.; Reid, St. P.; Tran, J. P.; Bergeron, A. A.; Wells, J.; Kota, K. P.; Aman, J.; Bavari, S. Identification of an Antioxidant Small-Molecule with Broad-Spectrum Antiviral Activity. *Antiviral Res.* **2012**, *93*(1), 23-29.
- (44) Galano, A.; Mazzone, G.; Alvarez-Diduk, R.; Marino, T.; Alvarez-Idaboy, J. R.; Russo, N. Food Antioxidants: Chemical Insights at the Molecular Level. *Annu. Rev. Food Sci. Technol.* **2016**, *7*(1), 335-352.

## 2. Computational Methods

Computational modeling of biological molecules is challenging because of the complexity of the systems (i.e. proteins, metalloenzymes and enzymes or DNA). In general, living matter has enormous complexity compared to even the most complex examples of inanimate matter. Biological processes occur on different length, from pm to mm and beyond and time scales, from fs to h and beyond. Understanding of chemical reactions (i.e., processes that involve substantial changes in the electronic structure of the system mainly due to cleavage/formation of bonds) are essential for development in several scientific areas. Computational chemistry has become an indispensable tool that is nowadays commonly used together with state-of-the-art experimental techniques to investigate biochemical reactions. For *in silico* insight, a “type of modeling” that explicitly represents the electronic degrees of freedom to be treated with quantum mechanical methods is mandatory. Fundamental properties without using empirical parameters and simplifications are studied by a so called *ab initio* or also first principles based methodologies. These calculations consider the electronic and structural features of the molecular systems at the absolute zero temperature and without considering the evolution in time. However, the fact to solve complex equations describing the electronic structure in an almost “exact” way makes these methods very expensive from the computational time point of view. So, today, *ab initio* calculations are still limited to systems containing at most some hundreds of atoms. When it is necessary to simulate very big and complex molecules, the use of first principles methods proves impractical. The problem can be faced through the introduction of approximations, such as the use of empirical parameters arising from experimental measurements in calculations that are, therefore, called semi-empirical or more drastically abandoning the quantum treatment of the problem by resorting to laws based on classical physics.

Computational chemistry methods that leverage our understanding of classical and quantum mechanics are now increasingly being used, also in a combined way under the so-called multiscale approach, to analyze the fundamental properties of many fascinating biological reactions,<sup>1</sup> such as electron transfer in proteins or bond breaking and formation along a catalytic reaction path.

In this chapter, following a multi-scale guideline, different theoretical approaches are described without going into deep detail, but simply to show how computational models applied to complex biological systems can be used as a powerful investigative tool. In section 2.1 methods based on classical physics are described. The molecular docking method to predict the interaction between biomolecule and ligand is described in section 2.2. The last section, 2.3, is dedicated to quantum mechanical (QM) and hybrid quantum mechanical/molecular mechanical (QM/MM) methods.

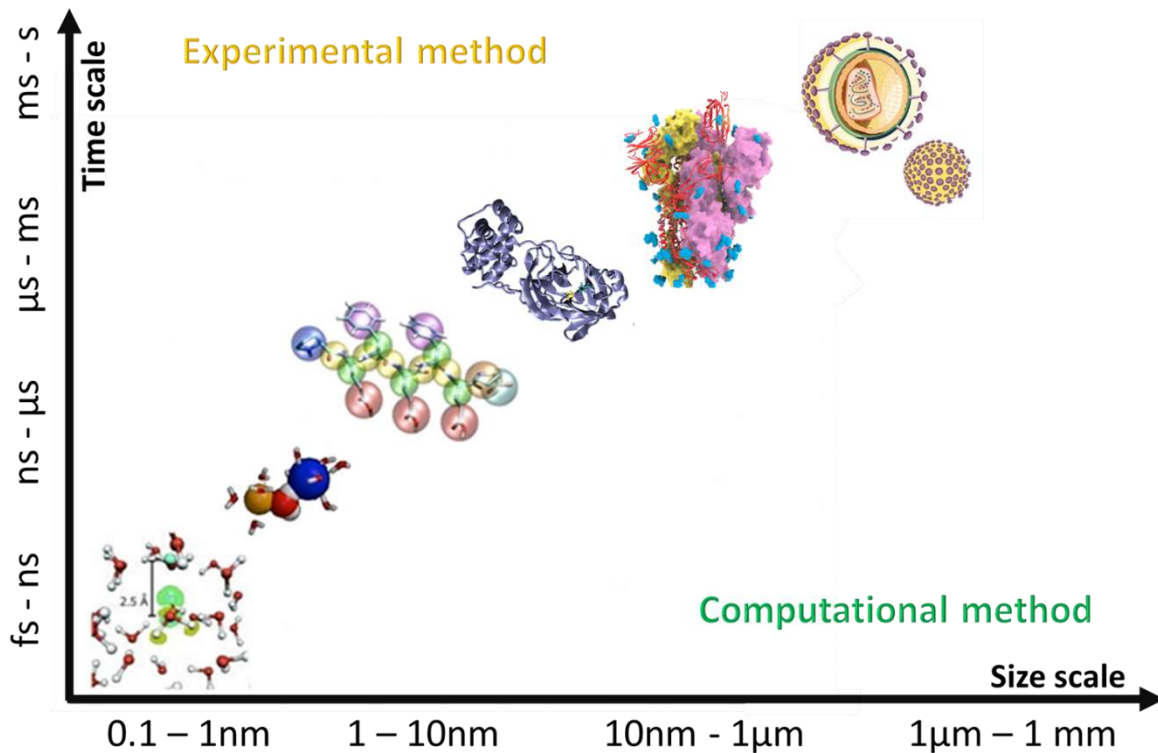


Figure 2.1 Hierarchical organization of bio-matter, from bottom left to top right, in ascending size and time magnitude, in a multi-scale modelling scheme.

## 2.1 Computational methods based on classical physics

Between 1 and 10 nm size system and on time scales between ns and  $\mu\text{s}$ , structural transitions take place involving entire functional biomolecules, like proteins. In some cases it is necessary, in order to have a simulations near as possible to reality, taking account of the solvent-dependent geometrical characteristics of protein, that the biomolecule is modeled in its aqueous environment. This involves the addition of water molecules to the system simulated. So systems that have a number of atoms that exceed tens of thousands are simulated.

With a number of  $10^3$ - $10^4$  atoms, it is impossible to consider explicitly the electronic degrees of freedom in reasonable computation time. For this reason, approaches based on the empirical description of the interatomic interactions for which electron effects are classically implicitly treated, are used. These methods are globally called all-atom Force Field based simulations and are often indicated also as Molecular Mechanics (MM) or FF based simulations. <sup>2</sup>

### 2.1.1 Force Fields

The potential function used in MM to describe the interactions between atoms in a system is called a Force Field (FF). The basic functional form of the potential function in the MM approach includes terms related to atoms that are covalently bound, bonding terms, and

terms to describe non-bonding properties (non-covalent interactions), these refer to long-range electrostatic and van der Waals forces. In this view, the form of the total energy in an additive force field is the resultant of the energy term associated with the covalent contribution, and the energy related to the non-covalent interactions, as reported in equation 2.1.

$$E_{TOT} = E_{Bonded} + E_{Non-Bonded} \quad (2.1)$$

In which:

$$E_{Bonded} = E_{bond} + E_{angle} + E_{dihedral} \quad (2.2)$$

$$E_{Non-Bonded} = E_{electrostatic} + E_{van\ der\ Waal} \quad (2.3)$$

Molecules are considered to be composed of a set of atoms held together by simple elastic forces (harmonic oscillator approximation) and the FF replaces the real potential with a simplified model valid in the region to be simulated. In this view, by replacing in equation 2.1 the explicit forms of eqs. 2.2 and 2.3 is obtained. For example, a simple first generation FF expression for the energy:

$$E_{TOT} = \sum_{(bonds)} \frac{1}{2} K_b (r - r_0)^2 + \sum_{(angles)} \frac{1}{2} K_a (\vartheta - \vartheta_0)^2 + \sum_{(torsions)} \frac{V_n}{2} [1 + \cos(n\vartheta - \delta)] + \sum_{ij(LJ)} 4\epsilon_{bij} \left( \frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) + \sum_{ij(elec)} \frac{q_i q_j}{r_{ij}} \quad (2.4)$$

where the first four terms refer to intramolecular or local contributions to the total energy (bond stretching, angle bending, and dihedral and improper torsions), owing to two, three, four and four body interactions, respectively. The last two terms serve to describe the repulsive and van der Waals interactions (in this case by means of a 12-6 Lennard-Jones potential) and the Coulombic interactions between particle  $i$  and  $j$ .<sup>3</sup>

The parameters for a chosen energy function can be obtained from experiments, quantum mechanical calculations, or both. Ideally, parameters should be simple enough to be evaluated quickly, but at the same time detailed enough, to reproduce the properties of interest of the systems studied. However, this approach suffers certain limitations,<sup>4</sup> such as high computational costs, generation of parameters for so defined non-standard residues and approximations of molecular forces required.

There is no universal FF to accurately describe all molecules; there are many force fields available in the literature characterized by different degrees of complexity and oriented to treat different kinds of systems. Among the most widely used FF can be included, AMBER (Assisted Model Building and Energy Refinement) used for proteins and DNA,<sup>5</sup> CHARMM (Chemistry at HARvard Macromolecular Mechanics) used for both small and macro molecules<sup>6</sup> and GROMOS (GRONingen MOlecular Simulation) a general-purpose molecular dynamics computer simulation package for the study of biomolecular systems.<sup>7</sup>



The universal force field (UFF) is a broadly applicable classical force field containing parameters for almost every atom type of the periodic table <sup>8</sup> mainly used in hybrid methods (i.e. QM/MM).

It is necessary to remember that the possibility to observe certain properties is directly related to the quality of the force field and, whether or not it has been parameterized for the system under analysis.

### 2.1.1 FF: inclusion of induction

It is worth noting that in all the force fields listed above, the effect of atoms, polarization by their environment is not taken into account. Polarization represents the modification of the electron clouds of an atom or a molecule by the effect of an external electric field. Deformation can be caused by physical chemical changes, such as changing the redox state of a cofactor for example, within the molecule itself.

PIPF (polarizable intermolecular potential function), <sup>9</sup> DRF90, <sup>10</sup> and AMOEBA <sup>11</sup> are good examples of force fields that consider polarization. It has been shown by several research groups that polarization must be taken into account. <sup>12-15</sup> There are several theoretical models to include polarization effects in FF, below are the three most popular.

In the *Drude oscillator model* <sup>16,17</sup> electronic polarization is based on the presence of a Drude particle attached to a harmonic spring with a defined force constant. This force is associated with the electric field perceived by the Drude particle. Through the development of parameters based on this model, called "Drude-2013",<sup>18</sup> many interesting simulations have been performed that have produced quantitative improvements over the first generation FF.<sup>18</sup> To date, there are still few simulations using the Drude model or proteins. <sup>19,20</sup> The advantage of this model is that it is easier to implement and less computationally demanding than the induced dipole model explained below.

The *fluctuating charge model* uses the same partial charge description as the traditional non-polarizable FF <sup>21,22</sup>, however, this time, the partial charges on each atom can change to accommodate different electrostatic environments during the simulation. This fluctuation is based on electronegativity equalization and thus can account for the polarization effect. <sup>23</sup> Polarization only takes into account charge fluxes along the bond directions, so out-of-plane effects are ignored. This aspect makes the use of a FF based on the fluctuating charge method unsuitable for some situations, such as simulation of redox potentials in proteins, where the effect of out-of-plane polarization is important.

In *the induced dipole model*, <sup>24</sup> a point classical dipole moment is induced at each polarizable atomic site according to the electric field sensed by this site. The induced dipole at each atomic site  $K$  is calculated as  $\mu_K = \alpha_K \phi_K$ , where  $\alpha_K$  represents the atomic polarizability and  $\phi_K$  refers to the electric field created by other atomic charges and other atomic induced dipoles. Since the induced dipoles affects the field at each atomic site, the procedure must be iterated to generate a self-consistent set of mutually influencing induced dipoles until convergence.



The iterative scheme is computationally expensive but is better able to reproduce the anisotropy and non-additivity of the molecular polarization response across many different compounds. The latter aspect makes it advantageous over the previously described models. An example of a point charge model FF is the so called AMBERff02.<sup>25</sup> In **Paper IV** AMBERff02 is used to determine the ionization potential of the 4Fe-4S cluster of the enzyme viperin.

### 2.1.2 Molecular Dynamics simulations

Classical Molecular Dynamics (MD)<sup>2</sup> is a computational methodology that allows one to study, through the integration of the equations of motion, the dynamical evolution of a chemical system at the atomistic level. In MD the time evolution of a set of  $N$  interacting particles is described by Newton's equations, in which the force acting on each atom  $i$  comes from:

$$F_i = -\nabla_{r_i} V(r_1, \dots, r_i, \dots, r_N) = m_i \frac{\partial^2 r_i}{\partial t^2} \quad (2.5)$$

where  $F_i$  is the gradient of the interaction potential with respect to the atomic displacements. This is a system of  $N$  coupled second order non linear differential equations that cannot be solved exactly, so equation (2.5) has to be solved numerically step by step using an appropriate integration algorithm. To deal with molecular dynamics simulation it is necessary to fix at least two aspects: a set of initial velocity and position of each  $i$ -particle in the system for a given time  $t$  and a time interval  $\delta t$  (time step) after which one wants to know the new velocities and positions. It is also necessary that a good model represents the forces acting between the particles (as the empirical force fields presented in the previous section) and to define the boundary conditions to be employed. In the case of a protein, a possible set of initial positions are available from crystallographic structure or more generally from experimental data. With regard to velocity of the  $i$ -particle it is attributed randomly from a Maxwell-Boltzmann distribution centered on the desired temperature and then adjusted in order to zero the angular momentum and the center of mass velocity of the total system.<sup>26</sup> So, through molecular dynamics microscopic information, such as velocity and position, is attributed to each particle. Via statistical mechanics it is possible to link microscopic information to macroscopic observables such as pressure, energy, heat capacity etc., remembering that to a given macroscopic state can correspond more than one microscopic states. The term *ensemble* is used to denote the idealization of a large number of copies of the system considered at one time ( $t_i$ ), each representative of a possible state in which the real system could be. Thus, an *ensemble* is defined as a collection of all the different microscopic states that correspond to the same macroscopic state of thermodynamic observables of the system such as temperature (T), pressure (P), and number of particles (N). Statistical mechanics is fundamental to the study of biomolecules by MD simulations. The hot particles of the system will have to give kinetic energy to the cooler particles and then forces will appear: there is a transfer from

one part to another of kinetic and potential energy. Based on this premise, the following ensembles can be simulated.

- *NVE ensemble* is defined as the microcanonical ensemble and it is associated with an adiabatic process in which there is no heat exchange. NVE is an isolated system in which the number of particles, velocity and total energy are conserved. Keeping in mind that if  $E_{tot}$  is constant, the instantaneous kinetic energy varies and the instantaneous temperature changes. This does not correspond to the conditions under which most experiments take place. In fact, constant-energy simulations are not recommended for equilibration because, without the energy flow facilitated by the temperature control methods, the desired temperature cannot be achieved. However, during the data collection phase, if there is an interest in exploring the constant-energy surface of the conformational space, or, for other reasons one does not want the perturbation introduced by temperature- and pressure-bath coupling, this is a useful ensemble.

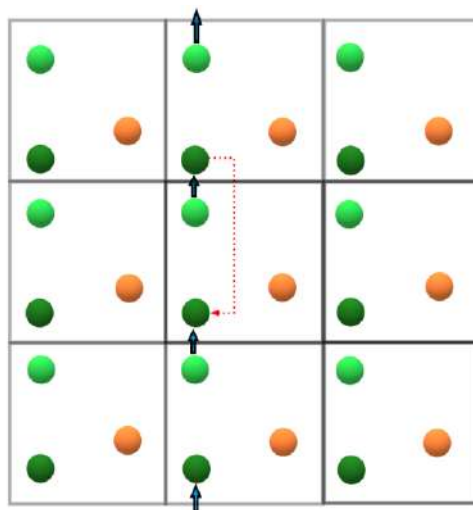
- *NVT ensemble* (canonical ensemble): because it is not possible to work at constant energy, constant temperature is used. In this way, the kinetic energy is constant but the potential energy and consequently the total energy, change. In order to keep the  $T$  constant, it is necessary to consider a heat exchange between the system and the universe (thermostatic bath). The NVT ensemble is essential to study unfolding protein that requires a precise control of temperature. This is the appropriate choice when conformational searches of molecules are carried out in vacuum without periodic boundary conditions.

- *NPT ensemble*: in this case  $P$  and  $T$  are kept constant. This ensemble plays an important role in chemistry because chemical reactions are usually carried out under constant pressure. This is the ensemble of choice when the correct pressure, volume, and densities are important in the simulation. This ensemble can also be used during equilibration to achieve the desired temperature and pressure before changing to the constant-volume or constant-energy ensemble when data collection starts.

For MD simulations of biomolecules, in addition to the choice of ensemble, the role of the solvent molecules is of crucial importance. The inclusion of an electrostatic constant ( $\epsilon$ ) in the electrostatic potential energy term, corresponds to the use of an implicit solvent in which solvent molecules are implicitly considered by modifying the interactions between the nuclei as if they suffered the average effect of a large number of solvent molecules.

The equilibrium position of many biochemical processes, including the recognition of a ligand/inhibitor by a receptor is strongly influenced by the solvent. It is essential that the water molecules, as well as any ions that guarantee electroneutrality, are explicitly present in the system under simulation. In this case, water molecules are added to the system in a solvation box of about 10 Å and they are classically described by employing a specific force field (e.g. TIP3P).<sup>27</sup> When explicit solvent is used, it is necessary to impose some boundary conditions. The boundary conditions are usually periodic (PBC: Periodic Boundary Condition) to minimize surface effects, whereby the system, near the edges, is less dense than inside the box. The particles are contained in a box and this box is replicated in all the

Cartesian directions, producing an infinite periodic system. Each particle inside the box will be affected by the interaction of the replicas in the surrounding identical volumes.



*Figure 2.2 Two-dimensional schematic representation of periodic boundary conditions with a particle, which exits the simulation box (central one) is reintroduced in replicas.*

## 2.2 Molecular Docking

Molecular interactions between protein-protein, enzyme-substrate, protein-nucleic acid and inhibitor-protein play important roles in many essential biological processes such as cellular transport and regulation, control of gene expression, enzymatic inhibition, and even assembly or cleavage of multi-domain proteins. In the field of computational chemistry applied to biological systems, if only substrate-free structure is available in the bioinformatics databases for a given protein, it is possible to predict preferred orientation of the substrate in a macromolecule by Molecular Docking.<sup>28</sup> Interaction between two molecules can be evaluated through the study of the hydrophobic interactions, electrostatic forces, hydrogen bonds and Van der Waals forces involved in the intermolecular association. The accuracy of prediction of the binding mode between ligand and protein is of essential. The docking protocol can be described as the combination of two algorithms: *search*<sup>29</sup> and *scoring functions*<sup>30</sup>. Docking can be rigid, where both the ligand and the protein are considered rigid<sup>31</sup> or flexible where the ligand flexibility within the receptor cannot be neglected.<sup>32</sup> All primary *search algorithms* calculating intermolecular interactions of molecular complexes considered molecules as rigid bodies. According to classical physics, the rigid body is an idealization of a solid body whose deformation is neglected and its position determined by the position of its center of mass. Therefore six degrees of freedom characterize the system (three rotational and three translational). These algorithms consider geometry matching between molecules and are based on Shape Matching (SM)<sup>33,34</sup> and Fast Fourier Transform (FFT)<sup>35</sup>. These algorithms were a starting point to the development of search algorithms available that consider

ligand and/or receptor as flexible bodies, then, conformational degrees of freedom are introduced. Flexible-ligand docking methods can better reproduce biological systems and can be divided in three categories:

- *Systematic methods*: where the ligand is divided into several fragments that are individually docked in the receptor site and covalently bound to reassemble the ligand, using the "anchoring and growth procedure".<sup>35,36</sup>
- *Random or stochastic methods*: seeks to generate and optimize low-energy conformations and orientations of the ligand and thus search for global energy minimum. Several docking programs apply the random approach: Monte Carlo, simulated annealing, evolutionary algorithms, Particle Swarm Optimization, etc.<sup>37-40</sup>
- *Deterministic methods*: produce time dependent conformations so the program uses molecular simulations to generate different poses.<sup>41</sup>

Once different ligands and ligand poses are determined, a *scoring function* is involved in prediction of the affinity of a ligand to bind to a protein target. The main goal of the scoring function is the calculation of the energy which can be expressed by the binding free energy ( $\Delta G_{\text{bind}}$ ).<sup>42</sup> The score function is the most critical element of docking and can be described, in general, by the relation:

$$S_{\text{core}} = C_s + C_i + E_{NB} + E_{Des} + (\dots) \quad (2.6)$$

where  $C_s$  represents the steric complementarity,  $C_i$  the hydrophobic contribution,  $E_{NB}$  the non-bonding energy, and  $E_{Des}$  is the desolvation energy. Other terms can be added to the formula (2.6) to increase the accuracy. For papers (**II**, **III** and **Wp**) present in this thesis the AutoDock 4 software was used,<sup>42</sup> where an empirical force field scoring function applies thermodynamic model of the binding process which includes intramolecular terms for the free energy estimation and also a full desolvation model. This provides terms for all atom types (favorable energetics of desolvating carbon atoms and unfavorable energetics of desolvating polar and charged atoms). This semiempirical approach estimates the energetics of complex formation using pair-wise terms to evaluate interactions between molecules and an empirical method to estimate the contribution of surrounding water. In this aspect, the semiempirical approach differs from traditional molecular mechanics force field, which uses explicit water molecules for evaluation of solvation contributions and does not include entropic terms that are important for the interaction energy assessment. Poisson-Boltzman-surface<sup>43</sup> and generalized Born-surface area<sup>44,45</sup> are examples of models that treat water as continuum dielectric medium. For atomic charges AutoDock uses the Gasteiger-Marsili charges<sup>46</sup> for the calculation of electrostatic interactions.

### 2.3 From computational methods based on wave functions to Density Functional Theory

The quantum-mechanical study of stationary properties of microscopic systems and their evolution over time is based on the Schrödinger equation. This equation can only be solved

for simple systems, while for multi-electronic systems it is practically impossible to trace the energy of the system itself by solving this equation, due to its explicit dependence on the number of particles present in the system. Based on the Heisenberg uncertainty principle, it is not possible, using a wavefunction-based method, to know exactly how a microscopic particle evolves over time because its position and momentum cannot be known at the same time. The analysis of complex systems can be approached with a quantum mechanical approach based on a functional not strictly related to the number of particles, so not based on the use of wave functions. The development of a "density matrix functional theory" or more simply a "density functional theory" (DFT) has been determinant for the theoretical study of multi-electronic systems. In this view we have a functional that depends on only three spatial coordinates ( $x, y, z$ ), regardless of the number of system particles. The use of a functional makes the numerical resolution of complex quantum mechanical systems simpler with respect to wavefunctions theory. A rigorous mathematical treatment for this approach was achieved through stepwise theoretical developments starting from the first attempt carried out with the introduction of the Thomas-Fermi model, by which the energy of a multi-electronic system is evaluated without the explicit use of the wavefunction. The use of this statistical model gives the possibility to derive, for a homogeneous electron gas, a medium potential expressible simply through the density  $\rho(r)$ , that can be expressed:

$$\rho(r) = N \int \dots \int |\psi(x_1, x_2, \dots, x_N)|^2 dx_1 dx_2 \dots dx_N \quad (2.7)$$

where, the electron density  $\rho(r)$ , excluding the spin variable, depends on three variables only, also for many-electron systems. In this view DFT allows one to deal with large molecular systems, including the electron correlation, with a computational demand much lower than methods based on wave functions.

Modern DFT methods are based on the two Hohenberg and Kohn (HK) theorems<sup>47</sup> and the Kohn-Sham (KS) formalism.<sup>48</sup> In 1964, Hohenberg-Kohn introduced the existence of a unique relationship between  $\rho(r)$  and all fundamental properties of a given system, including energy.<sup>47</sup> The HK theorem demonstrates that the only fundamental variable that determines the Hamiltonian, the eigenvalues correlated to it and the ground-state function, is the electron density. The second HK theorem demonstrates that for a trial electron density  $\tilde{\rho}(r)$ , the energy is higher or equal to the real energy of system, as shown in eq. 2.8:

$$E_0 \leq E_t[\tilde{\rho}] \quad (2.8)$$

where  $E_0$  is the real energy and  $E_t[\tilde{\rho}]$  is the energy, written as functional of the trial electron density  $\tilde{\rho}(r)$ , of a system with an external potential  $v(r)$ .

The explicit form of the functional is the major challenge of DFT.

In 1927 Thomas and Fermi provided the first example of density functional theory. However the performance of their model had a deficiency due to the poor approximation

of the kinetic energy. To solve this problem Kohn and Sham proposed, in 1965, a new approach. They considered a reference system with non-interacting electrons, having the same density of the real, interacting one.

The kinetic energy of the real system can be expressed as the sum of two contributions: the kinetic energy of the reference system  $T_s[\rho]$  and the kinetic energy that measures the electron correlation  $T_c[\rho]$ :

$$T[\rho] = T_s[\rho] + T_c[\rho] \quad (2.9)$$

consequently, a new redefinition of the universal functional can be introduced:

$$F[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] \quad (2.10)$$

where  $E_{xc}[\rho]$  is the exchange and correlation functional that represents the sum of the terms having an unknown analytical form.  $E_{xc}[\rho]$  includes the difference between the exact kinetic energy and the kinetic energy of the reference system, the non-classical electron-electron interaction and the self-interaction correction. It can be considered as the difference between  $T[\rho]$  and  $T_s[\rho]$  plus a non-classical part  $E_{ee}[\rho]$ .

$$E_{xc}[\rho] = (T[\rho] - T_s[\rho]) + (E_{ee}[\rho] - J[\rho]) \quad (2.11)$$

Now, considering a system with an external potential  $v(r)$  acting on the electrons due to the nuclear charges and omitting the interaction between nuclei,  $E_t[\rho]$  can be defined as a sum of four terms:

$$E_t[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int \rho(r)v(r)dr \quad (2.12)$$

$E_t[\rho]$  represents the expression of energy for the real, interacting system. In eq. 2.12 the only term for which no explicit form can be given is  $E_{xc}$ .

Applying the variational method, imposing the wave function orthogonality condition and using the Lagrange multipliers method, the Kohn-Sham equations are:

$$-\frac{1}{2} \nabla_i^2 + v_s(r)\psi_i^{KS} = \varepsilon_i \psi_i^{KS} \quad (2.13)$$

here  $v_s$  is the local potential for the single particle that includes the exchange and correlation potential,  $v_{xc}$ , defined as the functional derivative of  $E_{xc}$  with respect to  $\rho(r)$ . So  $v_s$  depends on the density and therefore the Kohn-Sham equations have to be solved iteratively.

$$v_{xc} = \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \quad (2.14)$$



Kohn-Sham equations led to a formalism that is exact and computationally accessible. The only drawback is the fact that the explicit form of the functional  $E_{xc}$  is unknown. Various DFT methods have been developed in order to find accurate expression for this functional. DFT functionals can be divided into two main groups, pure and hybrid functionals, according to whether they contain a component of Hartree-Fock exchange, or not. Considering  $E_{xc}[\rho(r)]$  as follow:

$$E_{xc}[\rho(r)] = E_x[\rho(r)] + E_c[\rho(r)] \quad (2.15)$$

Several exchange, correlation and exchange-correlation functionals were proposed. It is common in DFT to separate  $E_{xc}[\rho]$  into two parts, a pure  $E_x[\rho]$  and a correlation  $E_c[\rho]$  functionals. The simplest approach to calculate the exchange-correlation energy  $E_{xc}[\rho(r)]$  is based on assuming that the density  $\rho$  varies very slowly and locally with position and can thus be treated as a homogeneous electron gas. This is referred to as the *Local (Spin) Density Approximation* (L(S)DA).<sup>49</sup> The exchange energy of a uniform electron gas  $E_x[\rho(r)]$  can be computed exactly. *Gradient Corrected* or *Generalized Gradient Approximation* (GGA) makes a further step, utilizing not only the density  $\rho$ , but also the gradient.<sup>50</sup> Finally, some methods of modern DFT are hybrid type: they combine functionals LSDA, corrections from GGA, and also a piece of HF (Hartree-Fock) exchange  $E_x^{HF}$ , calculated via KS orbitals. Examples of correlation functionals are LYP,<sup>51</sup> PBE,<sup>52</sup> and P86,<sup>53</sup> some widely adopted exchange functional are PBEh,<sup>54</sup> mPW<sup>55</sup> and wPBEh.<sup>56</sup> M06-L<sup>57</sup> and B97D,<sup>58</sup> finally, are examples of exchange-correlation pure functionals. About hybrid functionals, the most famous and used are B3LYP,<sup>59</sup> M06,<sup>60</sup> PBE0<sup>61</sup> and wB97XD.<sup>62</sup> Anyway, many DFT methods fail to describe accurately long-range dispersion interactions, since they have local dependency on the electron density. This is why, over the years, several methods<sup>63</sup> have been developed to include dispersion interactions into density functional theory, according to the equation:

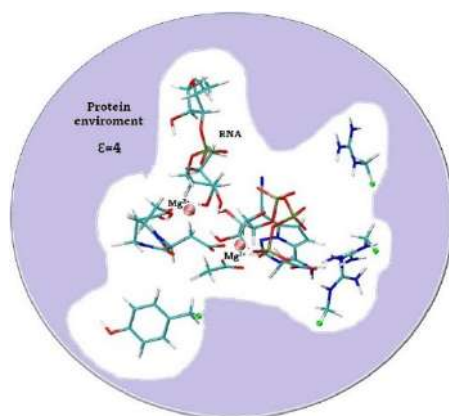
$$E_{DFT-D} = E_{DFT} + E_{disp} \quad (2.16)$$

A theoretical modeling of a biomolecules reactions mechanism can, in general, be accomplished at different levels of approximation. In the next sub-sections different methods to deal with mechanisms of inhibition/reaction and radical activity will be described.

### 2.3.1 QM cluster models

Most biochemical systems are too large to be entirely described at any level of wavefunctions methods or DFT theory in reasonable computational time. At the same time, it is well known that a quantum-mechanical approach allows one to calculate essential electronic properties in chemical reactions. Based on these two assumptions, the first model designed to achieve a mechanistic understanding of enzymes has been

developed and is known as the "QM-cluster" or "QM-only" approach.<sup>64</sup> In the cluster approach a model is built starting from the 3D structure of the enzyme by selecting a reasonable number of atoms, considering the available computer power and the quantum-chemical method employed, which obviously include the active site and the surrounding residues relevant to the enzymatic reactions. In the case of metalloenzymes, where the chemistry is dominated by electronic interactions with the metal ion, the choice of model is simplified by definition of first and second metal coordination spheres. A selection of atoms ranging from 100 to 300, depending on the system to be analyzed and the presence of any metals involved in the enzyme reactivity, allows one to obtain a reasonably accurate cluster.<sup>64</sup> This approach is not enough to account for steric hindrances and the electrostatic influence of the portion of the enzyme excluded from the QM selection. The conventional strategy to model the electrostatic influence of the protein environment is to assume a homogeneous polarizable medium using a dielectric constant,<sup>65-67</sup> taking into account that the choice of dielectric constant becomes less and less important as the cluster size increases.<sup>68-70</sup> Steric effects are simply simulated by fixing several atoms to their crystallographic positions, typically where truncation is made. It is important to determine where to cut and which atoms to fix, the side chain atoms or the  $\alpha$ -carbon atoms and two hydrogens along the backbone are the ones that are generally selected (see figure 2.3). It must be specified that the more the size of the QM-cluster increases the more accurate the description (model increases its flexibility).



*Figure 2.3. Scheme showing how to perform, generally, accurate cluster study for metalloenzyme. The green spheres indicate the frozen H atoms, the pink spheres indicated the two catalytic metal ions.*

This type of modeling is not always sufficiently accurate. Some times the protein environment imposes mechanical constraints that are required for the stability of the structure (such a situation occurs for the viperin enzyme, the object of study in chapter 3.3). Protein environment contributions are by definition absent in a QM cluster approach, leading us to consider an alternative methodology, QM/MM.



### 2.3.2 Hybrid QM/MM

The most obvious limitations of the cluster model is to neglect long-distance interactions and the conformational flexibility that arises from the complex large enzyme structure. These restrictions in some types of reaction cannot be ignored, for example in the cases where charge changes, such as in the calculations of redox potentials.<sup>67</sup> A hybrid quantum mechanics/molecular mechanics (QM/MM) approach was introduced by Warshel and Levitt in 1976.<sup>71</sup> It provides a treatment of a primary system at the QM level and its surroundings at the MM level. QM/MM combines the accuracy of QM calculations and the speed MM approaches, thus enabling the study of chemical processes in proteins.

Hybrid QM/MM methodologies differ, mainly, for the type of scheme (additive or subtractive), for the way in which interactions between QM and MM regions are considered and for the treatment of boundary region.<sup>64</sup>

General equations of additive and subtractive schemes are reported in figure 2.4.

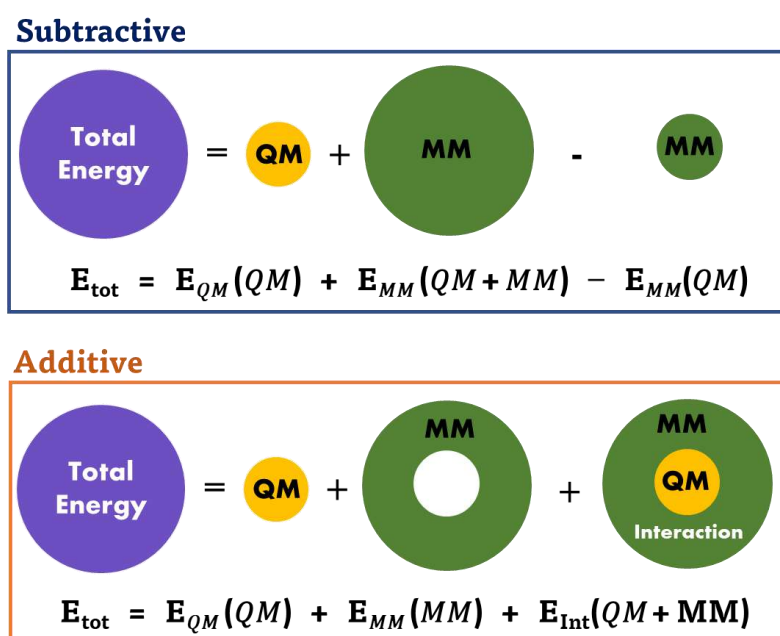


Figure 2.4. Representation of two-layer subtractive (top) and additive (bottom) schemes. The general formulation of the total energy of the entire system has been reported for each scheme.

The additive scheme does not require communication between the QM and MM calculations and mutual polarization of QM and MM parts can be taken into account at the MM level. ONIOM<sup>72</sup> is one of the most widely applied methods based on the subtractive scheme. The major limitations of this method stem from the need for a force field to describe the QM part and from the absence of polarization by the surroundings in the QM calculation. Additive QM/MM methods calculate explicitly the energy of the system as the sum of QM, MM and QM-MM contributions. In principle, this is the most accurate approach where the interaction term can be handled at the MM or QM level, the polarization of the QM region can be included in the QM calculation. The advantage of the additive QM/MM scheme is that no MM parameters for the QM atoms are needed, because

those energy terms are calculated by QM. However, dealing with the interactions at the border between QM and MM region can be tricky, no interactions must be omitted or double-counted.

In enzyme QM/MM calculations, there are in most of them covalent bonds across the QM and MM boundary. Possible strategies are used to treat it, among the most used, link atoms or frozen orbitals approaches. Briefly, in the first case link atoms are added to replace MM atoms, so as to saturate the dangling bond, considering the classical valence of QM atoms covalently bonded to the MM part. Usually, hydrogens are added and considered explicitly in the QM energy calculation. This method is simple to implement but there is a risk of charge overpolarization on the first MM atoms layer at the border.

In the second one, a frozen localized orbital is used to describe the interface between the two regions, assuming that the bonds have constant and well-defined properties. The bond is obtained from a linear combination between hybrid orbitals of the two atoms.

### 2.3.3 Environment effects

A critical point in QM/MM simulations is to treat the electrostatic interactions between the two regions. Mechanical, electrostatic or polarizable embedding are the three methodologies usually adopted to deal with QM/MM subsystems interaction, as shown respectively in the equations (2.17, 2.18, 2.19)

$$E_{int}(QM + MM) = \sum_{A=1}^{N_{QM}} \sum_{B=1}^{N_{MM}} V_{AB}^{Coul} + V_{AB}^{LJ} \quad (2.17)$$

$$E_i^{QM/MM} = E_i^{gas} - \sum_{B=1}^{N_{MM}} \frac{e \cdot q_B}{4\pi\epsilon_0 r_{iB}} \quad (2.18)$$

$$E_{ind}(MM) = -\frac{1}{2} \sum_{B=1}^{N_{MM}} \vec{\mu}_B \cdot \vec{\Phi}_B \quad (2.19)$$

$E_{int}(QM + MM)$  is the interaction energy between QM and MM regions obtained following a mechanical embedding scheme. Mechanical embedding treats electrostatic interactions at the MM level (equation 2.17) although simpler to implement with respect to other methods, some problems can arise. The definition of LJ parameters and partial charges for the QM region is not obvious and the limitation of the mechanical embedding scheme is that the MM subsystem has no direct influence on the electronic density of the QM subsystem. Therefore, all the properties of the latter are those corresponding to the isolated system not polarized by the surrounding. In the electrostatic embedding the MM environment is included in the QM Hamiltonian in terms of fixed point charges ( $q_B$  term in the eq. 2.18). This means that MM atoms polarize the QM density through electrostatic interactions. At the QM-MM boundary, where the MM charges are placed in close proximity to the electronic density, the MM charges can lead to overpolarization of the electronic density.<sup>73</sup> There is a risk of a polarization catastrophe and thus it is necessary to damp small-distance interactions. Electrostatic embedding represents the most popular

embedding scheme nowadays, due to its easy implementation into existing QM schemes and its efficiency to describe environmental effects by parametrized atomic charges.<sup>74,75</sup> This method does not consider the polarization of the MM system that, in some cases, such as redox reactions in proteins, affects the results.

### 2.3.4 QM/MM<sup>POL</sup>

While electrostatic embedding takes into account the polarization of the QM system by the MM system while neglecting the polarization of the MM system by the QM system, polarized embedding takes into account both contributions. These models allow flexible MM charges and fall into two categories. In the first category, the MM region is polarized by the QM electric field but then does not act on the QM system again. In the second category, there are fully self-consistent formulations that allow iterative mutual polarization between the QM and MM systems. These mentioned above are just some of the schemes, but nowadays there are more advanced ones available.<sup>76,77</sup>

When chemical processes in the region under QM investigation involve charge shifting, for example during electron or proton transfers, it becomes essential to capture the mutual polarization of the QM and MM regions.<sup>78</sup> In some types of reactions therefore electrostatic polarization defined by induction, must also be incorporated into the potential energy using polarizable QM/MMpol schemes and there are several ways to do this. The QM/MMpol calculations presented in this thesis (see **paper IV**) were carried out using deMon2k, a readily available program specialized in DFT simulations within the framework of Auxiliary DFT.<sup>79</sup> deMon2k implements the additive QM/MM scheme and thanks to the use of auxiliary basis sets allows the treatment of very large QM systems in a reasonable time. The program, to deal with polarization, implements a charge-dipole induction model,<sup>80</sup> the same used for the POL3 water model<sup>81</sup> and the AMBERff02 force field,<sup>82</sup> where the induction results:

$$\boldsymbol{\mu}_k = \alpha_k \mathbf{F}_k = \alpha_k \left( \mathbf{F}_k^{(0)} + \mathbf{F}_k^{(ind)} + \mathbf{F}_k^{(QM)} \right) \quad (2.19)$$

where  $(\alpha_k)$  is a force field parameter and  $F_k$  includes three terms, namely, the field created by other permanent MM charges ( $F_k^{(0)}$ ), the field created by other MM induced dipoles ( $F_k^{(ind)}$ ) and that created by the QM region  $F_k^{(QM)}$ . In this definition the interaction energy of the additive scheme, reported in figure 2.4, includes the energies arising from the interaction of induced dipoles with the electron density and the nuclei of the QM region, respectively.<sup>80</sup> The QM/MM<sup>pol</sup> scheme accounts for the mutual polarization between the QM and MM regions. The QM region contributes to the amplitude of the induced dipoles through the electric field  $F_k^{(QM)}$ , at the same time the induced dipoles are incorporated into the embedding Kohn–Sham Hamiltonian. This makes the self-consistent field (SCF) slightly more complicated because the Hamiltonian depends implicitly on the electron density through the induced dipoles ( $\mu_k$ ) and has to be recomputed at every SCF cycle.

## References

- (1) Warshel, A. Multiscale Modeling of Biological Functions: From Enzymes to Molecular Machines (Nobel Lecture). *Angewandte Chemie International Edition* **2014**, *53*(38), 10020–10031.
- (2) Frenkel, D.; Smit, B.; Ratner, M. A. *Understanding Molecular Simulation: From Algorithms to Applications. Physics Today* **1997**, *50*(7), 66–66.
- (3) Smit, B. Phase Diagrams of Lennard-Jones Fluids. *J. Chem. Phys.* **1992**, *96*(11), 8639–8640.
- (4) Karplus, M.; McCammon, J. A. Molecular Dynamics Simulations of Biomolecules. *Nat Struct Mol Biol* **2002**, *9*(9), 646–652.
- (5) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*(19), 5179–5197.
- (6) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry* **1983**, *4*(2), 187–217.
- (7) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *Journal of Computational Chemistry* **2004**, *25*(13), 1656–1676.
- (8) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*(25), 10024–10035.
- (9) Gao, J.; Habibollahzadeh, D.; Shao, L. A Polarizable Intermolecular Potential Function for Simulation of Liquid Alcohols. *J. Phys. Chem.* **1995**, *99*(44), 16460–16467.
- (10) Swart, M.; van Duijnen, P. Th. DRF90: A Polarizable Force Field. *Molecular Simulation* **2006**, *32*(6), 471–484.
- (11) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*(8), 2549–2564.
- (12) Blumberger, J.; Lamoureux, G. Reorganization Free Energies and Quantum Corrections for a Model Electron Self-Exchange Reaction: Comparison of Polarizable and Non-Polarizable Solvent Models. *Molecular Physics* **2008**, *106*(12–13), 1597–1611.
- (13) King, G.; Warshel, A. Investigation of the Free Energy Functions for Electron Transfer Reactions. *J. Chem. Phys.* **1990**, *93*(12), 8682–8692.
- (14) Ceccarelli, M.; Marchi, M. Simulation and Modeling of the Rhodobacter Sphaeroides Bacterial Reaction Center: Structure and Interactions. *J. Phys. Chem. B* **2003**, *107*(6), 1423–1431.
- (15) Ungar, L. W.; Newton, M. D.; Voth, G. A. Classical and Quantum Simulation of Electron Transfer Through a Polypeptide. *J. Phys. Chem. B* **1999**, *103*(34), 7367–7382.
- (16) van Maaren, P. J.; van der Spoel, D. Molecular Dynamics Simulations of Water with Novel Shell-Model Potentials. *J. Phys. Chem. B* **2001**, *105*(13), 2618–2626.
- (17) Lamoureux, G.; Roux, B. Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003**, *119*(6), 3025–3039.
- (18) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116*(9), 4983–5013. <https://doi.org/10.1021/acs.chemrev.5b00505>.
- (19) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9*(12), 5430–5449.

- (20) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. The Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J Chem Theory Comput* **2013**, *9*(9), 4046–4063.
- (21) Lucas, T. R.; Bauer, B. A.; Patel, S. Charge Equilibration Force Fields for Molecular Dynamics Simulations of Lipids, Bilayers, and Integral Membrane Protein Systems. *Biochim Biophys Acta* **2012**, *1818*(2), 318–329.
- (22) Rappe, A. K.; Goddard, W. A. Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* **1991**, *95*(8), 3358–3363. <https://doi.org/10.1021/j100161a070>.
- (23) Mortier, W. J.; Ghosh, S. K.; Shankar, S. Electronegativity-Equalization Method for the Calculation of Atomic Charges in Molecules. *J. Am. Chem. Soc.* **1986**, *108*(15), 4315–4320.
- (24) Applequist, J.; Carl, J. R.; Fung, K.-K. Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atom Polarizabilities. *J. Am. Chem. Soc.* **1972**, *94*(9), 2952–2960.
- (25) Cieplak, P.; Caldwell, J.; Kollman, P. Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *Journal of Computational Chemistry* **2001**, *22*(10), 1048–1057.
- (26) Silva, R.; Plastino, A. R.; Lima, J. A. S. A Maxwellian Path to the Q-Nonextensive Velocity Distribution Function. *Physics Letters A* **1998**, *249*(5), 401–408.
- (27) 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*(2), 926–935. <https://doi.org/10.1063/1.445869>.
- (28) Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr Comput Aided Drug Des* **2011**, *7*(2), 146–157.
- (29) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of Docking: An Overview of Search Algorithms and a Guide to Scoring Functions. *Proteins* **2002**, *47*(4), 409–443.
- (30) Jain, A. N. Scoring Functions for Protein-Ligand Docking. *Curr Protein Pept Sci* **2006**, *7*(5), 407–420.
- (31) Sauton, N.; Lagorce, D.; Villoutreix, B. O.; Miteva, M. A. MS-DOCK: Accurate Multiple Conformation Generator and Rigid Docking Protocol for Multi-Step Virtual Ligand Screening. *BMC Bioinformatics* **2008**, *9*(1), 184.
- (32) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking<sup>1</sup> Edited by F. E. Cohen. *Journal of Molecular Biology* **1997**, *267*(3), 727–748.
- (33) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A Geometric Approach to Macromolecule-Ligand Interactions. *Journal of Molecular Biology* **1982**, *161*(2), 269–288.
- (34) Shoichet, B. K.; Kuntz, I. D.; Bodian, D. L. Molecular Docking Using Shape Descriptors. *Journal of Computational Chemistry* **1992**, *13*(3), 380–397.
- (35) Katchalski-Katzir, E.; Shariv, I.; Eisenstein, M.; Friesem, A. A.; Aflalo, C.; Vakser, I. A. Molecular Surface Recognition: Determination of Geometric Fit between Proteins and Their Ligands by Correlation Techniques. *PNAS* **1992**, *89*(6), 2195–2199.
- (36) DesJarlais, R. L.; Sheridan, R. P.; Dixon, J. S.; Kuntz, I. D.; Venkataraghavan, R. Docking Flexible Ligands to Macromolecular Receptors by Molecular Shape. *J. Med. Chem.* **1986**, *29*(11), 2149–2153.
- (37) Caflisch, A.; Niederer, P.; Anliker, M. Monte Carlo Docking of Oligopeptides to Proteins. *Proteins: Structure, Function, and Bioinformatics* **1992**, *13*(3), 223–230.
- (38) Hart, T. N.; Read, R. J. A Multiple-Start Monte Carlo Docking Method. *Proteins: Structure, Function, and Bioinformatics* **1992**, *13*(3), 206–222.
- (39) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. Optimization by Simulated Annealing. *Science* **1983**, *220*(4598), 671–680.
- (40) Burton, R. M. Computational Laboratories for Organization Science: Questions, Validity and Docking. *Computational & Mathematical Organization Theory* **2003**, *9*(2), 91–108.



- (41) Wu, G.; Robertson, D. H.; Brooks, C. L.; Vieth, M. Detailed Analysis of Grid-Based Molecular Docking: A Case Study of CDOCKER-A CHARMM-Based MD Docking Algorithm. *J Comput Chem* **2003**, *24* (13), 1549–1562.
- (42) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *Journal of Computational Chemistry* **1998**, *19*(14), 1639–1662.
- (43) Grant, J. A.; Pickup, B. T.; Nicholls, A. A Smooth Permittivity Function for Poisson–Boltzmann Solvation Methods. *Journal of Computational Chemistry* **2001**, *22* (6), 608–640.
- (44) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* **1990**, *112*(16), 6127–6129.
- (45) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Pairwise Solute Descreening of Solute Charges from a Dielectric Medium. *Chemical Physics Letters* **1995**, *246*(1), 122–129.
- (46) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity—a Rapid Access to Atomic Charges. *Tetrahedron* **1980**, *36*(22), 3219–3228.
- (47) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*(3B), B864–B871.
- (48) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*(4A), A1133–A1138.
- (49) Nemeth, J.; Vautherin, D. Study of Finite Nuclei in the Local Density Approximation. *Physics Letters B* **1970**, *32*(7), 561–564.
- (50) Ziesche, P.; Kurth, S.; Perdew, J. P. Density Functionals from LDA to GGA. *Computational Materials Science* **1998**, *11* (2), 122–127.
- (51) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*(2), 785–789.
- (52) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*(18), 4.
- (53) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*(12), 8822–8824.
- (54) Ernzerhof, M.; Perdew, J. P. Generalized Gradient Approximation to the Angle- and System-Averaged Exchange Hole. *J. Chem. Phys.* **1998**, *109*(9), 3313–3320.
- (55) Adamo, C.; Barone, V. Exchange Functionals with Improved Long-Range Behavior and Adiabatic Connection Methods without Adjustable Parameters: The MPW and MPW1PW Models. *J. Chem. Phys.* **1998**, *108*(2), 664–675.
- (56) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid Functionals Based on a Screened Coulomb Potential. *J. Chem. Phys.* **2003**, *118*(18), 8207–8215.
- (57) Zhao, Y.; Truhlar, D. G. Hybrid Meta Density Functional Theory Methods for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions: The MPW1B95 and MPWB1K Models and Comparative Assessments for Hydrogen Bonding and van Der Waals Interactions. *J. Phys. Chem. A* **2004**, *108*(33), 6908–6918.
- (58) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry* **2006**, *27*(15), 1787–1799.
- (59) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *The Journal of Chemical Physics* **1993**, *98*(7), 5648–5652.
- (60) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor Chem Account* **2008**, *120*(1–3), 215–241.
- (61) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*(13), 6158–6170.
- (62) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Physical Chemistry Chemical Physics* **2008**, *10* (44), 6615–6620. <https://doi.org/10.1039/B810189B>.

- (63) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *The Journal of Chemical Physics* **2010**, *132* (15), 154104.
- (64) Ahmadi, S.; Barrios Herrera, L.; Chehelamirani, M.; Hostaš, J.; Jalife, S.; Salahub, D. R. Multiscale Modeling of Enzymes: QM-Cluster, QM/MM, and QM/MM/MD: A Tutorial Review. *International Journal of Quantum Chemistry* **2018**, *118* (9), e25558.
- (65) Liao, R.-Z.; Yu, J.-G.; Himo, F. Quantum Chemical Modeling of Enzymatic Reactions: The Case of Decarboxylation. *J. Chem. Theory Comput.* **2011**, *7* (5), 1494–1501.
- (66) Georgieva, P.; Himo, F. Quantum Chemical Modeling of Enzymatic Reactions: The Case of Histone Lysine Methyltransferase. *Journal of Computational Chemistry* **2010**, *31* (8), 1707–1714.
- (67) Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M. Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114* (7), 3601–3658.
- (68) Siegbahn, P. E. M.; Himo, F. The Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *WIREs Computational Molecular Science* **2011**, *1* (3), 323–336.
- (69) Siegbahn, P. E. M.; Himo, F. Recent Developments of the Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *J Biol Inorg Chem* **2009**, *14* (5), 643–651.
- (70) Ramos, M. J.; Fernandes, P. A. Computational Enzymatic Catalysis. *Acc. Chem. Res.* **2008**, *41* (6), 689–698.
- (71) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *Journal of Molecular Biology* **1976**, *103* (2), 227–249.
- (72) Morokuma, K. ONIOM and Its Applications to Material Chemistry and Catalyses. *Bulletin of the Korean Chemical Society* **2003**, *24* (6), 797–801.
- (73) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *Journal of Computational Chemistry* **1990**, *11* (6), 700–733.
- (74) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor Chem Acc* **2006**, *117* (2), 185. <https://doi.org/10.1007/s00214-006-0143-z>.
- (75) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew Chem Int Ed Engl* **2009**, *48* (7), 1198–1229.
- (76) Loco, D.; Lagardère, L.; Caprasecca, S.; Lipparini, F.; Mennucci, B.; Piquemal, J.-P. Hybrid QM/MM Molecular Dynamics with AMOEBA Polarizable Embedding. *J. Chem. Theory Comput.* **2017**, *13* (9), 4025–4033.
- (77) Loco, D.; Lagardère, L.; Adjoua, O.; Piquemal, J.-P. Atomistic Polarizable Embeddings: Energy, Dynamics, Spectroscopy, and Reactivity. *Acc. Chem. Res.* **2021**, *54* (13), 2812–2822.
- (78) de la Lande, A.; Alvarez-Ibarra, A.; Hasnaoui, K.; Cailliez, F.; Wu, X.; Mineva, T.; Cuny, J.; Calaminici, P.; López-Sosa, L.; Geudtner, G.; Navizet, I.; Garcia Iriepa, C.; Salahub, D. R.; Köster, A. M. Molecular Simulations with In-DeMon2k QM/MM, a Tutorial-Review. *Molecules* **2019**, *24* (9), 1653.
- (79) A.M. Koster, G. Geudtner, A. Alvarez-Ibarra, P. Calaminici, M.E. Casida, J. Carmona-Espindola, V.D. Dominguez, R. Flores-Moreno, G.U. Gamboa, A. Goursot, T. Heine, A. Ipatov, A. de La Lande, F. Janetzko, J.M. Del Campo, D. Mejia-Rodriguez, J. U. Reveles, J. Vasquez-Perez, A. Vela, B. Zuniga-Gutierrez, and D.R. Salahub, DeMon2k, Version 6, The DeMon Developers, Cinvestav, Mexico City (2018).
- (80) Wu, X.; Teuler, J.-M.; Cailliez, F.; Clavaguera, C.; Salahub, D. R.; de la Lande, A. Simulating Electron Dynamics in Polarizable Environments. *J Chem Theory Comput* **2017**, *13* (9), 3985–4002.
- (81) Caldwell, J. W.; Kollman, P. A. Structure and Properties of Neat Liquids Using Nonadditive Molecular Dynamics: Water, Methanol, and N-Methylacetamide. *J. Phys. Chem.* **1995**, *99* (16), 6208–6219.
- (82) Wang, Z.-X.; Zhang, W.; Wu, C.; Lei, H.; Cieplak, P.; Duan, Y. Strike a Balance: Optimization of Backbone Torsion Parameters of AMBER Polarizable Force Field for Simulations of Proteins and Peptides. *Journal of Computational Chemistry* **2006**, *27* (6), 781–790.

### 3. Results and Discussion

The results of this thesis derived, for the most part, from multiscale computational approaches are presented in two different sections according to the different nature of the studied systems. In fact, the first section (**section A**) includes **Paper I, Paper II, Paper III, Paper IV** and **Wp**, that refer to enzymatic systems or ascribable to them, whereas **section B**, that includes **Paper V** and **Paper VI**, to naturally occurring chemical compounds whose main antioxidant properties and reaction mechanisms involved in free radical scavenging in minimizing the oxidative stress are studied.

All the investigated systems are engaged in multistep biological processes so elucidating the mechanisms (reaction or inhibition) is of crucial importance. Computational methodologies have the privilege to provide mechanistic insights not accessible to the experimental ones but necessary to drive rational drug design.

In the first sub-section (3.1), the study of the reaction mechanism of a mononuclear non-heme iron-based biomimetic catalyst is reported, related to the O<sub>2</sub> activation by 2 electrons reduction in the oxidative decarboxylation of the hydroxy acids as sacrificial reductants (**Paper I**).

**Paper II** refers to inhibition mechanisms of enzymes involved in cell replication. In particular, the metalation process of Ribonuclease A by a Pt-As containing complex has been elucidated by DFT calculations using a large cluster model able to represent the environment of the binding site quite realistically. Since viral proteases are considered to be attractive drug targets because of their crucial function in the viral replication machinery, the inhibition mechanism of the main protease (M<sup>pro</sup>) of SARS-CoV-2 has been studied by applying computational methods ranging from classical Molecular Dynamics simulations and Molecular Docking to DFT calculations (**Paper III**).

**Paper IV** focuses on the study of the natural antiviral enzyme viperin, a virus inhibitory protein, endoplasmic reticulum-associated, interferon-inducible featuring in the active site an iron-sulfur cluster involved in the catalytic process. It is a member of the radical SAM superfamily of enzymes that use SAM as a substrate to convert SAM stoichiometrically to Met and 5'-deoxyadenosine. Structural characteristics of the enzyme has been deeply analyzed by Molecular Dynamics simulations. Ionization potential has been evaluated by advanced computational approaches. In addition, since the catalytic product of viperin is a modified nucleobase (ddhCTP) based on a pyrimidine structure, it has been tested as inhibitor of RNA-polymerase-RNA-dependent of SARS-CoV-2 with chain terminator nature (**Wp**).

**Section B** reports electronic structure calculations performed for some naturally occurring chemical compounds using DFT and solvent models. These calculations allow one to estimate reactivity indices used to identify the best candidates for free radical elimination via single electron transfer (SET) or hydrogen atom transfer (HAT) mechanisms. The examined species represent a “chemical tool” to fight the oxidative stress (OS) that is an elevated risk to human health as it is associated with the onset and development of numerous diseases including different types of cancer and cardiovascular



and neurodegenerative disorders. OS does not produce obvious symptoms, there are currently no routine medical tests implemented to detect it, so it could be labeled as a "silent chemical killer". Identifying possible antioxidant therapies that can reduce the damage caused by OS can represent an important and innovative solution (**Paper V and VI**).

## Section A

### 3.1 Iron biomimetics (Paper I)

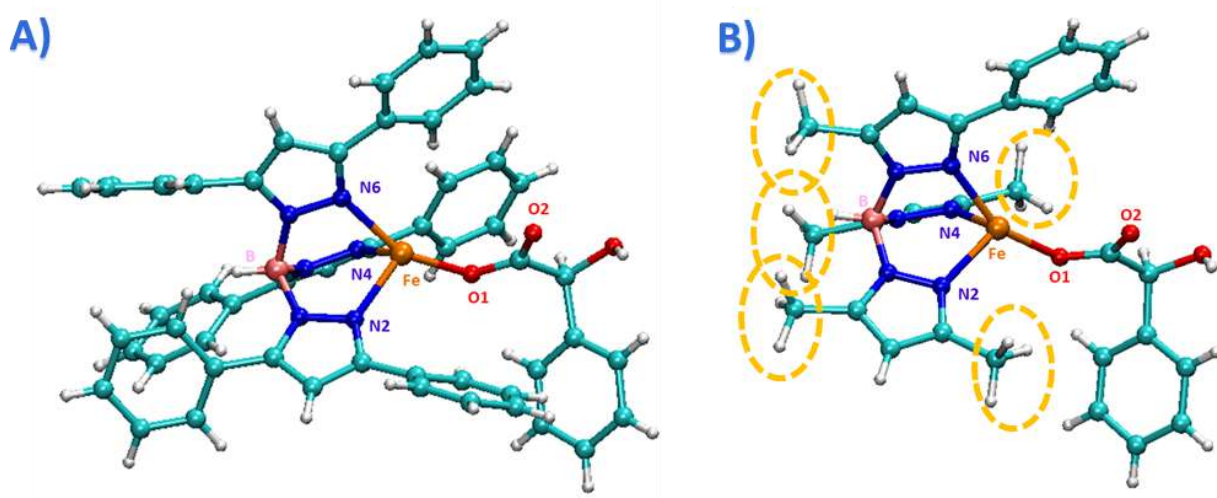
Biomimetic complexes that simulate reactivity properties of the metalloenzyme active sites can provide important mechanistic insights. In particular, the chosen example serves as model to reproduce the dioxygenase activity of  $\alpha$ -ketoglutarate-dependent enzymes. Such reactions occurring with the activation of dioxygen ( $O_2$ ) are extremely important in biology.<sup>1</sup>

Members of the non-heme iron enzyme family catalyse the incorporation of  $O_2$  into a wide range of biological molecules and use diverse strategies to activate their substrates.<sup>1</sup> The oxidative decarboxylation of the iron(II)  $\alpha$ -hydroxy acid (mandelic acid) complex model, biomimetic of Rieske's dioxygenase, was investigated at the density functional level.

The crystal structure (X ray) of the biomimetic catalyst<sup>2</sup> has been used as the starting point and, to reduce the computational time, five phenyl rings of the  $Tp^{Ph_2}$  ligand have been replaced by methyl groups (see Figure 3.1).

The proposed mechanism has been shown in Scheme 1.2 of chapter 1.

The potential energy surfaces have been explored in different electron spin states. The naked biomimetic iron(II)  $\alpha$ -hydroxy acid has been optimized in septet, quintet, triplet and singlet spin-states by using the M06L<sup>3,4</sup> exchange-correlation functional with the Gaussian09 software.<sup>5</sup> Spin contamination has also been considered.



*Figure 3.1 A) Crystal structure of hydrotris(3,5-diphenylpyrazole-1-yl)borate) Fe(II)-mandelate, B) computational model used in Paper I. Replaced methyl groups are highlighted in yellow.*

### 3.2 Inhibition mechanisms of enzymes involved in replication

Irreversible enzyme inhibitors can form a covalent bond to a key amino acid of the catalytic pocket and block irrevocably the enzyme.<sup>6</sup> The most effective irreversible inhibitors are those that contain an electrophilic group capable of receiving a pair of electrons by the nucleophilic group of the key amino acid to form a new covalent bond. Covalent inhibitors are recognized as an important component in drug development or repurposing. The many advantages of covalent inhibitors are counteracting initial concerns regarding potential off-target toxicity. Thus, research continues, especially for cancer targets or in the development of antiviral molecules.<sup>7</sup>

In the next two sub-sections (3.2.1 and 3.2.2) the computational study for the investigation of covalent inhibition to clarify and predict the anticancer activity of arsenoplatin-1 compound ( $[\text{Pt}(\mu\text{-NHC}(\text{CH}_3)\text{O})_2\text{ClAs}(\text{OH})_2]$ ) and the antiviral activity of ebselen (2-phenyl-1,2-benzoselenazol-3-one) against SARS-CoV-2 virus, is described and addressed in **Paper II** and **Paper III**, respectively.

#### 3.2.1 RNase A inhibition by arsenoplatin-1 (**Paper II**)

RNase (Ribonuclease A) is an endoribonuclease (EC 3.1) that cleaves the phosphodiester bonds of single strand RNA after pyrimidine nucleotides. It attacks at the 3' phosphate end.<sup>8</sup>

Cisplatin is the progenitor of the Pt-based antitumor family.<sup>9</sup>

The interaction of the cisplatin drug with ribonuclease A (RNase A) in aqueous solution at physiological pH has been well investigated.<sup>10</sup>

Recently, mixed systems of Pt (II) and As (III) have been synthesized and tested as potential anticancer agents. The first representative of this new class of anticancer agents, arsenoplatin-1, shows a superior activity profile compared to the parental drugs  $\text{As}_2\text{O}_3$  or cisplatin in most of the tested cancer cell lines. It has been demonstrated that arsenoplatin-1 complex (AP1) depicted in Figure 3.2, once hydrated, can interact not only with the DNA bases,<sup>11</sup> but also with specific amino acid residues of some proteins involved in RNA hydrolysis: Hen Egg White Lysozyme (HEWL) and Ribonuclease A (RNase A).<sup>12</sup> X-ray crystallographic structures evidenced that the preferred site for cisplatin and AP1 are different, in fact the platination site in the case of the cisplatin molecule is Met 29 whereas in the case of AP-1 are the residues His105 and His119 of the ribonuclease. Since His119 residues are involved in the catalytic activity of the RNase A, it results that the binding of AP1 inhibits such a function, as indicated by a catalytic activity assay.<sup>12</sup>

The crystallographic structure of platinated RNase (PDB code 5NJ7)<sup>12</sup> has been used as the starting point for the mechanistic hypothesis for a backward studied pathway, starting

from the product and arriving at the reactants (see scheme 3.1). A detailed metalation process of RNase A by novel multitarget anti-cancer AP1, has been investigated by using both B3LYP<sup>13</sup> and M062x<sup>3</sup> functionals and with a large QM model of the active site. The role of water molecules in the active site has been carefully monitored and analyzed during the Molecular Dynamics simulations.

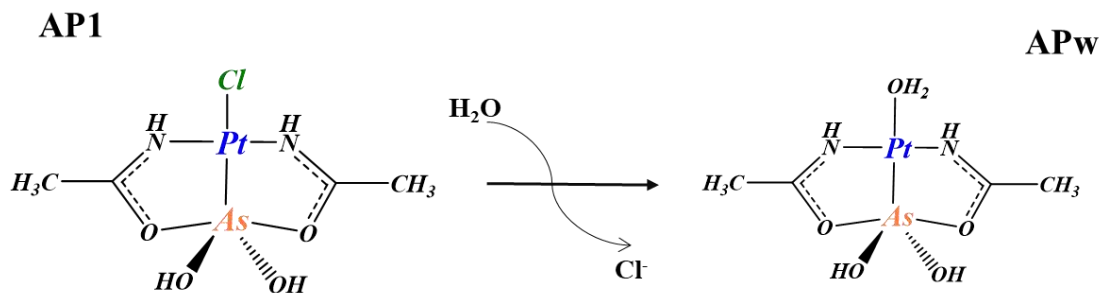
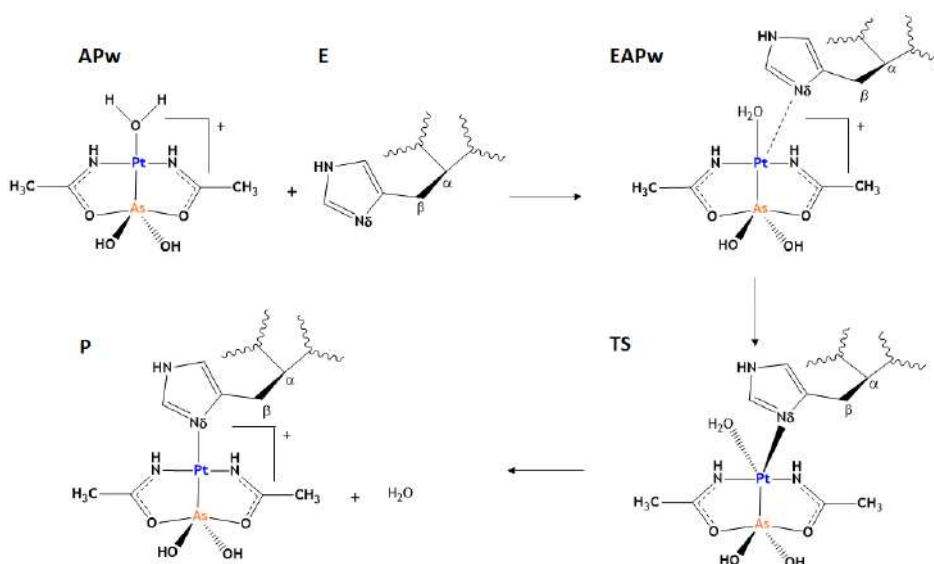


Figure 3.2 Structure of arsenoplatin-1 (AP1) and its aquated form (APw).



Scheme 3.1 The mechanism of the RNase A metalation process promoted by AP1. APw refers to aquated form of AP1.

To better rationalize the different behavior of AP1 (bound to His119) and the reference cisplatin (bound to Met29) towards the same target RNase A, a small model consisting of the single amino acid residue (His or Met) and Pt-based complex (cis-Pt or AP-1) has been employed. (see Figure 3.3) Results confirm that the more favorable energetics concern the aquated form APw and the histidine resulted in the preferred target in agreement with the experimental observations.<sup>12</sup>

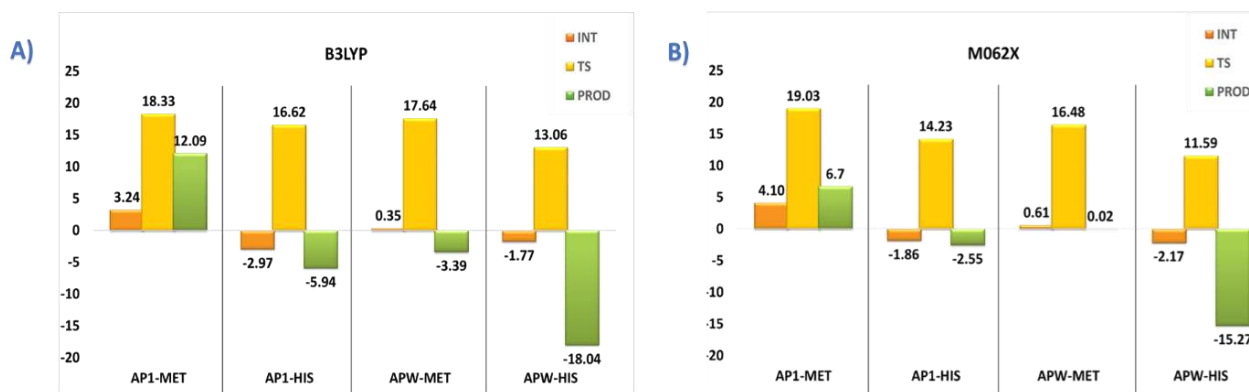


Figure 3.3 A) B3LYP/6-31+G(d,p) and B)M062x/6-31+G(d,p) free energy profiles of His and Met amino acid residues metalation process promoted by AP1 and APw, in gas phase. For both Pt and As atoms the effective core potential SDD coupled with their related orbital basis set has been selected.

All details about the theoretical study on platination mechanism of RNase A by AP1 are present in **Paper II**.

### 3.2.2 Main protease SARS-CoV-2 inhibitors (**Paper III**)

A vital step in the life cycle of coronaviruses is the proteolytic processing of virally expressed polyproteins into functional units by virus-encoded proteases.<sup>14</sup> Proteases (EC 3.4) are proteolytic enzymes that form one of the largest groups of enzymes and play a critical role in cell replication. Proteases can be divided into four major classes: aspartic, serine, cysteine, and metalloproteases. Protease inhibitors therefore have considerable potential utility for therapeutic intervention as antiviral agents.

The 21st century has been characterized by three outbreaks of coronavirus (CoVs) infections caused by severe acute respiratory syndrome (SARS)-CoV, Middle East respiratory syndrome (MERS)-CoV, and SARS-CoV-2. The morphology of virions is similar to a crown due to the mushroom-shaped protein called spike glycoproteins (S-proteins) responsible for the host tropism. Following host cell entry, the viral single-stranded positive RNA is released for replication and translation of virus polyproteins that are cleaved into two large viral polyproteins, processed by two cysteine proteases, such as papain-like protease (PL<sup>pro</sup>) and a 3-Chymotrypsin-like protease, also known as 3CL<sup>pro</sup> or main protease (M<sup>pro</sup>). Considering that M<sup>pro</sup> is only found in the virus rather than in the host cell, this protein represents an interesting target for the development of new promising anti-coronavirus therapeutic agents. As a cysteine protease, most of inhibitors covalently bind to the Cys residue of the catalytic site of M<sup>pro</sup>. From a screening of a wide range of approved drugs as possible M<sup>pro</sup> inhibitors, ebselen (EBS), a small selenium containing molecule was revealed to be a promising drugs lead to target M<sup>pro</sup>, exhibiting strong antiviral activity in cell based assays.<sup>15</sup>

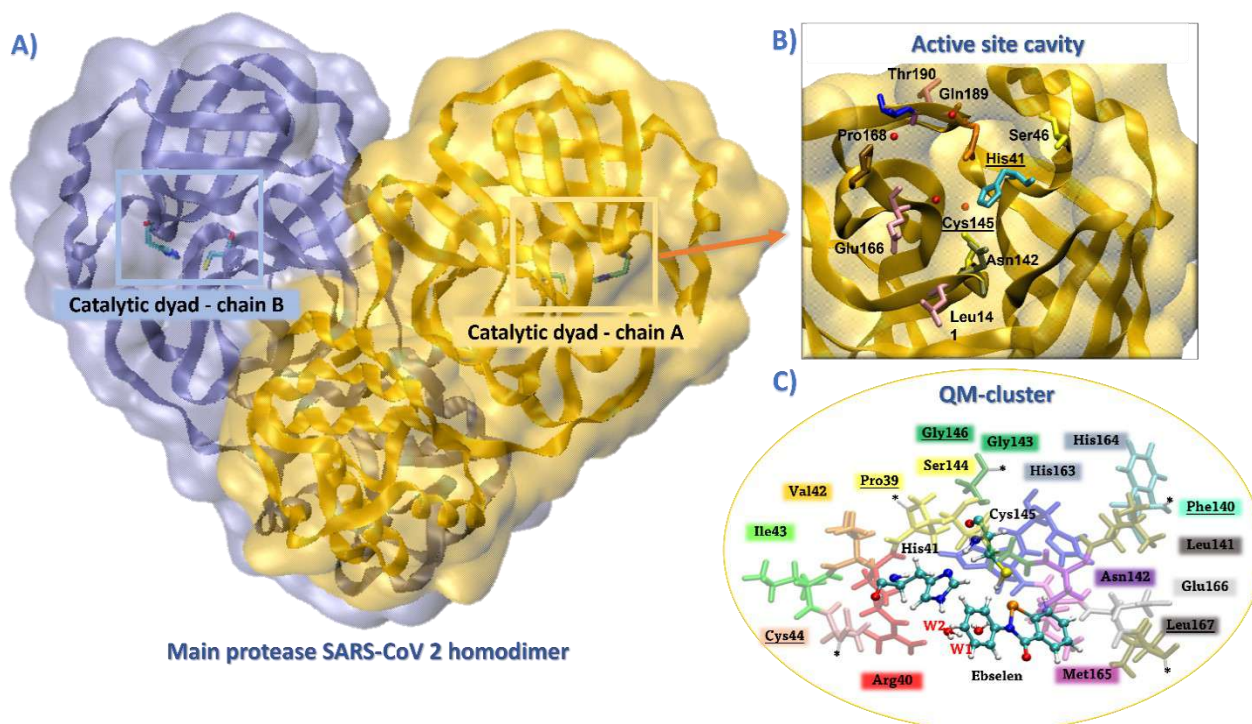
The crystal structure of main protease bound to potent broad-spectrum non-covalent inhibitor X77, with PDB code 6W63, has been used as starting point. The structure is characterized by one chain of 306 amino acids. The apo-enzyme has been obtained by removing the inhibitor from the crystallographic structure. Molecular dynamics simulations on the apo-form of M<sup>pro</sup> have been performed taking into account both hydrogen donor and acceptor natures of the N $\delta$  and N $\epsilon$  of His41, a residue of the M<sup>pro</sup> catalytic dyad (His41-Cys145).

Molecules with inhibiting activity proposed by Jin *et al.*,<sup>15</sup> have been docked into the active site of M<sup>pro</sup> of the most populated clusters arising from the clustering of MD trajectory of M<sup>pro</sup>, by using the Autodock software.<sup>16</sup> Results obtained from molecular docking are shown in Table 3.1.

*Table 3.1 Docking scores of protein-ligand binding affinity free energy expressed in kcal mol<sup>-1</sup>. Molecules column contains drug candidates proposed as inhibitors (the underlined molecules showed both covalent and noncovalent inhibition behavior).<sup>15</sup> Cluster0 to cluster3 refers to four most populated structures obtained by RMSD-based clustering MD simulation; w average is the average of the docking scores weighed by the population on total processed frames.*

<b>Molecule</b>	<b>cluster0</b>	<b>cluster1</b>	<b>cluster2</b>	<b>cluster3</b>	<b>w average</b>
<b>Amprenavir</b>	-5.32	-6.08	-5.90	-6.24	-5.81
<b>Indinavir</b>	-7.52	-7.95	-6.07	-8.16	-7.41
<b>Lopinavir</b>	-6.38	-6.26	-4.90	-6.93	-6.09
<b>Nelfinavir</b>	-7.85	-8.68	-6.13	-7.65	-7.66
<b>Ritonavir</b>	-5.89	-6.08	-3.59	-5.66	-5.38
<b><u>Disulfiram</u></b>	-3.71	-4.67	-3.10	-4.35	-3.94
<b><u>Ebselen</u></b>	-6.39	-6.14	-6.79	-6.10	-6.37
<b><u>Carmofur</u></b>	-6.03	-5.80	-5.91	-6.01	-5.99
<b><u>Px12</u></b>	-5.76	-5.81	-5.67	-5.88	-5.78
<b><u>N3</u></b>	-6.03	-6.11	-6.12	-6.00	-6.03





*Figure 3.4 A)  $M^{pro}$  homodimer cartoon representation, B) zoom into the active site including the catalytic dyad and some crystallographic water molecules. C) QM-cluster model of 312 atoms used for the DFT study in **Paper III**. Amino acid residues are shown in sticks and colored by residue name, the catalytic dyad (Cys145, His41), EBS and water molecules W1 and W2 are in balls and sticks. Stars on residues Pro39, Cys44, Phe140, Gly146 and Leu167 point frozen atoms and delimit the model truncations.*

Docking scores in Table 3.1 are related to the free energy of protein-ligand binding affinity expressed in  $\text{kcal mol}^{-1}$ , more negative docking score indicates a higher ligand-receptor affinity. Data obtained through docking simulations coupled with experimental data candidate ebselen as the most promising in the group of the covalent inhibitors tested. Therefore, ebselen provides all the promise of drug repurposing. In this regard, we focused the investigation on both non-covalent and covalent inhibition mechanisms of ebselen against  $M^{pro}$ , evaluating at the computational level its behaviour into binding pocket. Moreover, from the inhibitory potency studied for ebselen analog with a hydroxyl group at position 2 of benzisoxaselenazol-3(2H)-one ring (EBS-OH) with  $PL^{pro}$ , EBS-OH  $M^{pro}$  inhibition mechanism has been also investigated and compared with that of ebselen. To do this, molecular docking tools, Molecular Dynamics (MD) and QM-cluster (DFT) methods have been used. Complete details can be found in **Paper III**.

### 3.3 Viperin: a radical-SAM enzyme with antiviral activity

Oxidation processes in living cells are catalysed by the cooperation of a number of enzymes and coenzymes that successively transfer reducing equivalents, either hydrogen atoms or electrons. Some of the most studied redox metalloenzymes, e.g. cytochrome

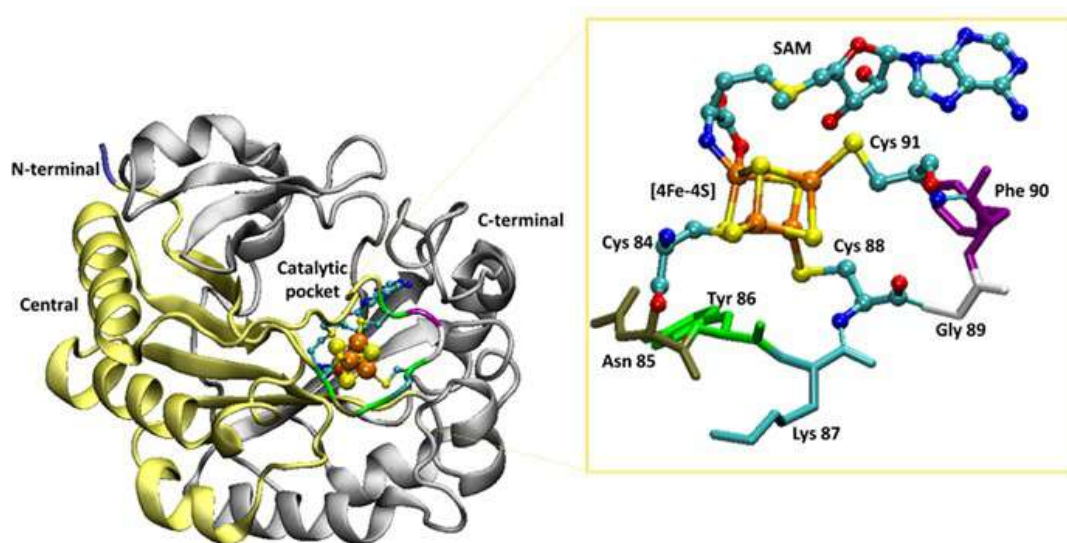
P450, methane monooxygenase, ribonucleotide reductase and adenosylcobalamin, involve in their catalytic cycle the formation of a one or more radical organic species. The main enzyme's superfamily related to radical reactions uses two cofactors widely present in all three kingdoms of life: iron-sulfur clusters and SAM (S-Adenosylmethionine).<sup>17</sup>

Viperin (Virus inhibitory protein, endoplasmic reticulum-associated, interferon-inducible) is a member of the radical SAM superfamily of enzymes found in mammals, including humans. It is defined as “natural antiviral” since it participates in the innate immune response of cells against a series of viruses. Its effect has been verified to inhibit a broad range of viruses, including HIV-1, West Nile virus, hepatitis C, dengue virus type 2, influenza A virus, and tick-borne encephalitis virus. A recent study<sup>18</sup> revealed that viperin catalyses the conversion reaction of cytidine triphosphate (CTP) to 3'-deoxy-3',4'-didehydro-CTP (ddhCTP). The catalytic product of viperin structurally resembles the antiviral drugs used in clinical treatment.

### 3.3.1 Structural investigation and ionization potential determination (Paper IV)

Human viperin is a single polypeptide of 361 amino acids with a predicted molecular weight of 42 kDa. The N-terminal 42 amino acids of viperin forms an amphipathic  $\alpha$ -helix, which is relatively less conserved in different species and has a minor effect on the antiviral activity of viperin. The N-terminal domain of viperin is required for its localization to the endoplasmic reticulum and lipid droplets.<sup>19</sup>

The C-terminal 218-361 amino acids of viperin are highly conserved in different species and essential for pathogens inhibition. The central domain contains the catalytic pocket that consists of a CxxxCxxC motif, characteristic of radical S-adenosylmethionine enzymes. Cys<sup>•</sup> residues within the characteristic motif ligate three of the iron atoms of the [4Fe-4S] cluster, the fourth iron of the cluster coordinates an oxygen and a nitrogen atom of the SAM methionine part (Figure 3.5).



*Figure 3.5 Cartoon diagram of viperin where N-terminal domain is in light blue, central domain in yellow, C-terminal domain in grey. [4Fe-4S] cluster shown as orange and yellow spheres, S-Adenosylmethionine and cysteines bond to cluster are in ball and sticks, residues X of CX<sub>3</sub>CX<sub>2</sub>C are in tube colored by residue. In the square zoom on viperin binding site. (PDB code 6Q2P)*

The active site architecture, together with multiple sequence alignments, show that vertebrate viperins are highly conserved and that fungi, bacteria, and archaeobacteria express viperin-like enzymes. The structural similarity among the radical SAM enzymes allows to put forward mechanistic hypotheses on the catalytic reaction involving a single-electron reductive cleavage of SAM by the iron-sulphur cluster and generation of a highly reactive radical intermediate, 5'-deoxyadenosyl radical (dAdo·).<sup>18</sup> In the presence of substrate, the radical abstracts a hydrogen atom from the substrate (CTP) and forms 5'-deoxyadenosine (dAdoH) as a by-product.

Knowing the electronic structure of the Fe-S cluster is of crucial importance to understand the mechanisms of electron transfer that activate the catalytic enzymatic reaction.

Our work focuses on the structural characteristics that favour housing of the CTP inside the catalytic cavity of viperin. To this end, we performed molecular dynamics simulations of viperin in its apo- and CTP-bound forms, at physiological pH and temperature. Considering the nature of the metal cofactor, electronic structure knowledge of the Fe-S cluster is of crucial importance to understand the mechanisms of electron transfer that activate the catalytic enzymatic reaction. The broken-symmetry density functional theory method is generally used to describe the various possible spin states of the [4Fe-4S] cluster.<sup>20-22</sup> We propose an original approach based on constrained Density Functional Theory (cDFT)<sup>23</sup> to describe the antiferromagnetically-coupled spin configuration of [4Fe-4S]<sup>+</sup> cluster in the diabatic approximation. Additionally the simplified QM + MM<sup>pol</sup> approach<sup>24</sup> has been used for 4Fe-4S cluster ionization potential calculation, taking into account the influence of protein environment. (**Paper IV**)

### 3.3.2 Viperin product (ddhCTP) as inhibitor of SARS-CoV-2 (**Wp**)

As already described in section 3.2.2, SARS-CoV-2 is positive-sense single-stranded RNA viruses. Proliferation of SARS-CoV-2 requires RNA-dependent RNA polymerase (RdRp), a key enzyme that regulates replication and transcription of the viral genome and it is thus the validated target for the development of therapies against COVID-19 disease. The RdRp complex has multiple non-structural protein (nsp) units. The structure of RdRp consists of the nsp7-nsp8 heterodimer, nsp8 subunit, and nsp12 core catalytic unit.<sup>25</sup> The complex structure of the RdRp is shown in the Figure 3.6.



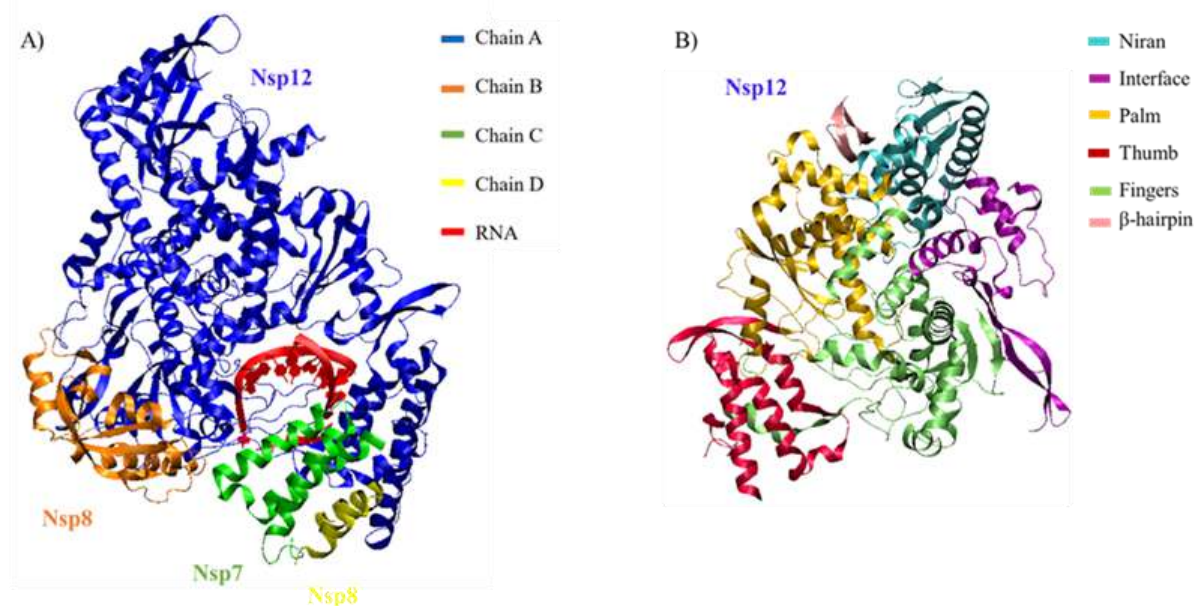


Figure 3.6 A) Structures of SARS-CoV-2 nsp7-nsp8-nsp12 complex. B) Structure of nsp12 domain organized by color.

The polymerase domain adopts the conserved architecture of the viral polymerase family<sup>26</sup> and is composed of three subdomains: a Finger subdomain (residues Leu 366 to Ala 581 and Lys 621 to Gly 679), a Palm subdomain (residues Thr 582 to Pro 620 and Thr 680 to Gln 815), and a Thumb subdomain (residues His 816 to Glu 920).

RdRp is proposed as a target of a class of antiviral drugs that are nucleotide analogues; this category includes remdesivir, nucleotide analogue mimicking the structure of adenosine, approved by the FDA on August 10, 2020, as a COVID-19 disease clinical treatment.

3'-deoxy-3',4'-didehydro-cytidine triphosphate (ddhCTP) is a natural antiviral nucleotide-like compound produced by enzyme viperin as part of the innate immune response.<sup>18</sup> ddhCTP moreover, has recently been shown to be easy to synthesize.<sup>27</sup> Its effect as a chain terminator for RNA-dependent polymerases of multiple members of the Flavivirus genus has been shown.<sup>18</sup> Recently it has been shown that SARS-CoV-2 polymerase incorporates this cytosine analogue well.<sup>28</sup> In addition, the -OH group loss in the 3' position, compared to its natural precursor (cytidine triphosphate, CTP) makes ddhCTP competitive with CTP at an intracellular concentration of  $\sim 100 \mu\text{M}$  and lower than purine analogues.<sup>28</sup>

Encouraged by experimental evidence on ddhCTP and the computational work on Remdesivir,<sup>29,30</sup> we perform an *in silico* study on inhibition of SARS-CoV-2 RdRp. Classical Molecular Dynamics (MD) simulations and free energy perturbation methods are used to study the RdRp inhibition promoted by ddhCTP. In addition, the behavior of the natural nucleotide cytidine triphosphate is investigated and compared with its analogue. Molecular docking to test the affinity between polymerase active site and the ligands. 300 ns of molecular dynamics simulation have been performed for the following systems: i) RdRp system that include nsp7, nsp8, nsp12, a primer nucleotide strand of 6 units and a template strand of 8 nucleotides; ii) RdRp system in complex with ddhCTP and iii) RdRp system with CTP. MD simulation analyses have been compared for the three systems. Finally, molecular mechanics energies combined with the Poisson-Boltzmann surface area

continuum solvation (MM-PBSA) has been used to understand the binding affinities between substrate or inhibitor and the biological target.

## Section B

### 3.1 Antioxidant (Paper V and VI)

As discussed in chapter 1 antioxidants have been reported to have numerous beneficial effects on human health.<sup>31</sup> Many of these beneficial effects are associated with the ability of antioxidants to inhibit oxidative stress and the resulting bio-molecular damage. Antioxidants are able to perform their function in different ways through a series of different mechanisms.<sup>32</sup> The studied antioxidants belong to type I,<sup>33</sup> those that react directly with free radicals, reducing significantly reactive species or turning off the chain of radicals reactions. In this context possible action mechanisms are: hydrogen-atom transfer (HAT), proton-coupled electron transfer (PCET), radical adduct formation (RAF), single electron transfer (SET), sequential proton-loss electron transfer (SPLET), sequential electron proton transfer (SEPT), and sequential proton-loss hydrogen-atom transfer (SPLHAT). The mechanisms explored in **papers V** and **VI** are shown in Figure 3.2.

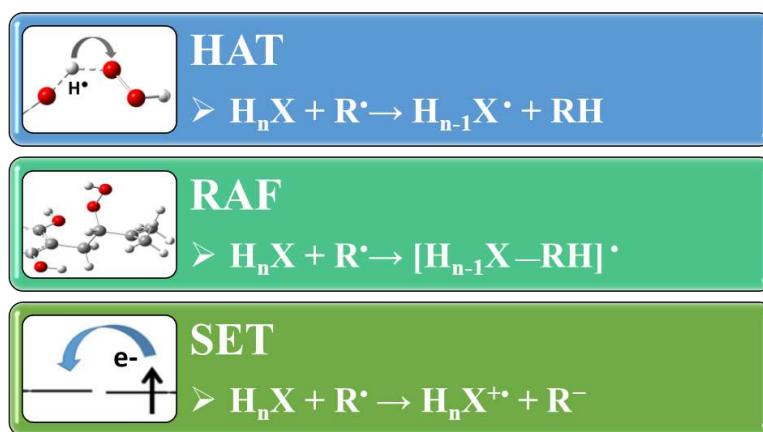


Figure 3.7 Possible reaction mechanisms involved in the free radical scavenging activity of phenolic compounds described in **papers V** and **VI**.

The important role of HAT in antioxidant activity has been proposed as a key reaction mechanism for polyphenols in general.<sup>34-36</sup>

In the case of the RAF, the key structural antioxidant feature is the presence of multiple bonds. Electrophilic radicals are more likely involved in RAF reactions than non-electrophilic radicals.<sup>37</sup> Steric effects can play an important role as they can limit the viability of RAF reactions.<sup>38</sup>

The transfer of a single electron is an important mechanism to be evaluated in the analysis of antioxidant activity of the molecule and generally the associated kinetics is studied by the Marcus theory.<sup>39,40</sup>

Free radicals are species containing one or more unpaired electrons present in cells. The most common ROS are oxygen-centered free radicals including the superoxide anion radical ( $O_2\cdot^-$ ), hydroxyl ( $\cdot OH$ ), alkoxyl ( $RO\cdot$ ), ( $ROO\cdot$ ), and hydroperoxyl ( $HOO\cdot$ ) radicals. Among oxygen-centered radicals,  $\cdot OH$  is the most electrophilic and reactive, with a half-life of  $\sim 10^{-9}$  s.<sup>41</sup>

In **paper V** calculations based on DFT have been carried out to highlight the mechanism of antioxidant functioning of higenamine in aqueous and lipid-like environments (Gaussian09 software)<sup>5</sup>. The mechanisms shown in Figure 3.7 between higenamine and the  $HOO\cdot$  radical have been studied by applying the Quantum Mechanics-based Test for Overall Free Radical Scavenging Activity (QM-ORSA).<sup>42</sup> It is a computational protocol designed to be a reliable tool in the study of radical-molecule reactions in solution. It provides a separated quantification of such activity in polar and non-polar media. It includes two different scales for quantification: (i) absolute, based on overall rate coefficients and (ii) relative, using Trolox as a reference antioxidant.<sup>42</sup> The molar fractions at physiological pH were determined in aqueous solution. All the species present in water derived from acid–base equilibria have been considered.

In **paper VI**, computational approach analogous to that of paper V has been applied to determine the antioxidant capability of moracin C and iso-moracin C isomers against the  $HOO\cdot$  free radical. In this case, besides evaluating the antioxidant activity of moracin, it was possible to perform a comparative evaluation of the reactivity of the two isomers.

## References

- (1) Que, L.; Ho, R. Y. N. Dioxygen Activation by Enzymes with Mononuclear Non-Heme Iron Active Sites. *Chem. Rev.* **1996**, *96*(7), 2607–2624.
- (2) Paria, S.; Chatterjee, S.; Paine, T. K. Reactivity of an Iron–Oxygen Oxidant Generated upon Oxidative Decarboxylation of Biomimetic Iron(II)  $\alpha$ -Hydroxy Acid Complexes. *Inorg. Chem.* **2014**, *53*(6), 2810–2821.
- (3) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor Chem Account* **2008**, *120*(1–3), 215–241.
- (4) Valero, R.; Costa, R.; de P. R. Moreira, I.; Truhlar, D. G.; Illas, F. Performance of the M06 Family of Exchange–Correlation Functionals for Predicting Magnetic Coupling in Organic and Inorganic Molecules. *J. Chem. Phys.* **2008**, *128*(11), 114103.
- (5) 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, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; 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, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision D.01; Gaussian, Inc.: Wallingford CT, **2010**.
- (6) Powers, J. C.; Asgian, J. L.; Ekici, Ö. D.; James, K. E. Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases. *Chem. Rev.* **2002**, *102*(12), 4639–4750.
- (7) Baillie, T. A. Targeted Covalent Inhibitors for Drug Design. *Angewandte Chemie International Edition* **2016**, *55*(43), 13408–13421.
- (8) Roberts, G. C. K.; Dennis, E. A.; Meadows, D. H.; Cohen, J. S.; Jardetzky, O. The Mechanism of Action of Ribonuclease. *PNAS* **1969**, *62*(4), 1151–1158.
- (9) Loehrer, P. J.; Einhorn, L. H. Cisplatin. *Ann Intern Med* **1984**, *100*(5), 704–713.
- (10) Neault, J. F.; Novetta-Delen, A.; Tajmir-Riahi, H. A. Interaction of Cisplatin Drug with RNase A. *J Biomol Struct Dyn* **1999**, *17*(1), 101–109.
- (11) Marino, T.; Parise, A.; Russo, N. The Role of Arsenic in the Hydrolysis and DNA Metalation Processes in an Arsenous Acid–Platinum(II) Anticancer Complex. *Phys. Chem. Chem. Phys.* **2017**, *19*(2), 1328–1334.
- (12) Miodragović, Đ.; Merlino, A.; Swindell, E. P.; Bogachkov, A.; Ahn, R. W.; Abuhadba, S.; Ferraro, G.; Marzo, T.; Mazar, A. P.; Messori, L.; O'Halloran, T. V. Arsenoplatin-1 Is a Dual Pharmacophore Anticancer Agent. *J. Am. Chem. Soc.* **2019**, *141*(16), 6453–6457.
- (13) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*(2), 785–789.
- (14) Kato, G. J. Human Genetic Diseases of Proteolysis. *Human Mutation* **1999**, *13*(2), 87–98.
- (15) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* **2020**, *582*(7811), 289–293.
- (16) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.;Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry*, **2009**, *30*(16), 7.
- (17) Sofia, H. J. Radical SAM, a Novel Protein Superfamily Linking Unresolved Steps in Familiar Biosynthetic Pathways with Radical Mechanisms: Functional Characterization Using New Analysis and Information Visualization Methods. *Nucleic Acids Research* **2001**, *29*(5), 1097–1106.

- (18) Gizzi, A. S.; Grove, T. L.; Arnold, J. J.; Jose, J.; Jangra, R. K.; Garforth, S. J.; Du, Q.; Cahill, S. M.; Dulyaninova, N. G.; Love, J. D.; Chandran, K.; Bresnick, A. R.; Cameron, C. E.; Almo, S. C. A Naturally Occurring Antiviral Ribonucleotide Encoded by the Human Genome. *Nature* **2018**, *558*(7711), 610–614.
- (19) Hinson, E. R.; Cresswell, P. The Antiviral Protein, Viperin, Localizes to Lipid Droplets via Its N-Terminal Amphipathic  $\alpha$ -Helix. *Proceedings of the National Academy of Sciences* **2009**, *106* (48), 20452–20457.
- (20) Nicolet, Y.; Amara, P.; Mouesca, J.-M.; Fontecilla-Camps, J. C. Unexpected Electron Transfer Mechanism upon AdoMet Cleavage in Radical SAM Proteins. *Proc Natl Acad Sci U S A* **2009**, *106* (35), 14867–14871.
- (21) Zhou, S.; Wei, W.-J.; Liao, R.-Z. QM/MM Study of the Mechanism of the Noncanonical S-C $\gamma$  Bond Scission in S-Adenosylmethionine Catalyzed by the CmnDph2 Radical Enzyme. *Top Catal* **2021**, *1*, 11.
- (22) Blachly, P. G.; Sandala, G. M.; Giammona, D. A.; Bashford, D.; McCammon, J. A.; Noodleman, L. Broken-Symmetry DFT Computations for the Reaction Pathway of IspH, an Iron-Sulfur Enzyme in Pathogenic Bacteria. *Inorg. Chem.* **2015**, *54*(13), 6439–6461.
- (23) de la Lande, A.; Salahub, D. R. Derivation of Interpretative Models for Long Range Electron Transfer from Constrained Density Functional Theory. *Journal of Molecular Structure: THEOCHEM* **2010**, *943* (1), 115–120.
- (24) Blumberger, J. Free Energies for Biological Electron Transfer from QM/MM Calculation: Method, Application and Critical Assessment. *Phys. Chem. Chem. Phys.* **2008**, *10*(37), 5651–5667.
- (25) Jiang, Y.; Yin, W.; Xu, H. E. RNA-Dependent RNA Polymerase: Structure, Mechanism, and Drug Discovery for COVID-19. *Biochem Biophys Res Commun* **2021**, *538*, 47–53.
- (26) McDonald, S. M. RNA Synthetic Mechanisms Employed by Diverse Families of RNA Viruses. *WIREs RNA* **2013**, *4*(4), 351–367.
- (27) Wood, J. M.; Evans, G. B.; Grove, T. L.; Almo, S. C.; Cameron, S. A.; Furneaux, R. H.; Harris, L. D. Chemical Synthesis of the Antiviral Nucleotide Analogue DdhCTP. *J Org Chem* **2021**, *86*(13), 8843–8850.
- (28) Seifert, M.; Bera, S. C.; van Nies, P.; Kirchdoerfer, R. N.; Shannon, A.; Le, T.-T.-N.; Meng, X.; Xia, H.; Wood, J. M.; Harris, L. D.; Papini, F. S.; Arnold, J. J.; Almo, S.; Grove, T. L.; Shi, P.-Y.; Xiang, Y.; Canard, B.; Depken, M.; Cameron, C. E.; Dulin, D. Inhibition of SARS-CoV-2 Polymerase by Nucleotide Analogs from a Single-Molecule Perspective. *eLife*, **2021**, *10*, e70968.
- (29) Gordon, C. J.; Tchesnokov, E. P.; Woolner, E.; Perry, J. K.; Feng, J. Y.; Porter, D. P.; Götte, M. Remdesivir Is a Direct-Acting Antiviral That Inhibits RNA-Dependent RNA Polymerase from Severe Acute Respiratory Syndrome Coronavirus 2 with High Potency. *J Biol Chem* **2020**, *295*(20), 6785–6797.
- (30) Nayeem, S. M.; Sohail, E. M.; Sudhir, G. P.; Reddy, M. S. Computational and Theoretical Exploration for Clinical Suitability of Remdesivir Drug to SARS-CoV-2. *European Journal of Pharmacology* **2021**, *890*, 173642.
- (31) Chang, S. K.; Alasalvar, C.; Shahidi, F. Review of Dried Fruits: Phytochemicals, Antioxidant Efficacies, and Health Benefits. *Journal of Functional Foods* **2016**, *21*, 113–132. <https://doi.org/10.1016/j.jff.2015.11.034>.
- (32) Kashyap, P.; Shikha, D.; Thakur, M.; Aneja, A. Functionality of Apigenin as a Potent Antioxidant with Emphasis on Bioavailability, Metabolism, Action Mechanism and in Vitro and in Vivo Studies: A Review. *Journal of Food Biochemistry* **2021**, e13950.
- (33) Chaiyasit, W.; Elias, R. J.; McClements, D. J.; Decker, E. A. Role of Physical Structures in Bulk Oils on Lipid Oxidation. *Crit Rev Food Sci Nutr* **2007**, *47* (3), 299–317. <https://doi.org/10.1080/10408390600754248>.
- (34) Galano, A.; Martínez, A. Capsaicin, a Tasty Free Radical Scavenger: Mechanism of Action and Kinetics. *J Phys Chem B* **2012**, *116*(3), 1200–1208.
- (35) Li, X.; Wang, T.; Zhou, B.; Gao, W.; Cao, J.; Huang, L. Chemical Composition and Antioxidant and Anti-Inflammatory Potential of Peels and Flesh from 10 Different Pear Varieties (*Pyrus* Spp.). *Food Chemistry* **2014**, *152*, 531–538.



- (36) Castañeda-Arriaga, R.; Alvarez-Idaboy, J. R. Lipoic Acid and Dihydrolipoic Acid. A Comprehensive Theoretical Study of Their Antioxidant Activity Supported by Available Experimental Kinetic Data. *J. Chem. Inf. Model.* **2014**, *54* (6), 1642–1652.
- (37) Galano, A.; Mazzone, G.; Alvarez-Diduk, R.; Marino, T.; Alvarez-Idaboy, J. R.; Russo, N. Food Antioxidants: Chemical Insights at the Molecular Level. *Annual Review of Food Science and Technology* **2016**, *7*(1), 335–352.
- (38) Li, X.; Ren, Z.; Wu, Z.; Fu, Z.; Xie, H.; Deng, L.; Jiang, X.; Chen, D. Steric Effect of Antioxidant Diels-Alder-Type Adducts: A Comparison of Sanggenon C with Sanggenon D. *Molecules* **2018**, *23* (10), 2610.
- (39) Marcus, R. A. Chemical and Electrochemical Electron-Transfer Theory. *Annual Review of Physical Chemistry* **1964**, *15* (1), 155–196. <https://doi.org/10.1146/annurev.pc.15.100164.001103>.
- (40) Marcus, R. A. Electron Transfer Reactions in Chemistry. Theory and Experiment. *Rev. Mod. Phys.* **1993**,
- (41) Pryor, W. A. Why Is the Hydroxyl Radical the Only Radical That Commonly Adds to DNA? Hypothesis: It Has a Rare Combination of High Electrophilicity, High Thermochemical Reactivity, and a Mode of Production That Can Occur near DNA. *Free Radical Biology and Medicine* **1988**, *4* (4), 219–223.
- (42) Galano, A.; Alvarez-Idaboy, J. R. A Computational Methodology for Accurate Predictions of Rate Constants in Solution: Application to the Assessment of Primary Antioxidant Activity. *Journal of Computational Chemistry* **2013**, *34* (28), 2430–2445.



## Section A

### Paper I



The generation of the oxidant agent of a mononuclear nonheme Fe(II) biomimetic complex by oxidative decarboxylation. A DFT investigation





Article

# The Generation of the Oxidant Agent of a Mononuclear Nonheme Fe(II) Biomimetic Complex by Oxidative Decarboxylation. A DFT Investigation

Angela Parise, Maria Costanza Muraca, Nino Russo \* , Marirosa Toscano and Tiziana Marino \*

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Cubo 14C, Via P. Bucci, 87036 Arcavacata di Rende, CS, Italy; angela.parise@unical.it (A.P.); muraca.mc@gmail.com (M.C.M.); marirosa.toscano@unical.it (M.T.)

\* Correspondence: nrusso@unical.it (N.R.); tiziana.marino65@unical.it (T.M.); Tel.: +39-0984-492106 (N.R.); +39-0984-492085 (T.M.)

Academic Editor: Sérgio F. Sousa

Received: 10 December 2019; Accepted: 10 January 2020; Published: 14 January 2020



**Abstract:** The oxidative decarboxylation of the iron(II)  $\alpha$ -hydroxy acid (mandelic acid) complex model, biomimetic of Rieske dioxygenase, has been investigated at the density functional level. The explored mechanism sheds light on the role of the  $\alpha$ -hydroxyl group on the dioxygen activation. The potential energy surfaces have been explored in different electronic spin states. The rate-determining step of the process is the proton transfer. The oxidative decarboxylation preferentially takes place on the quintet state.

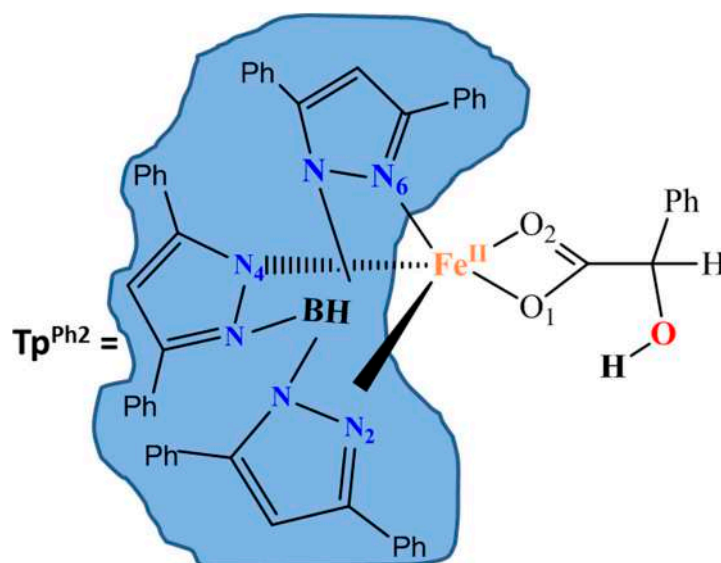
**Keywords:** nonheme biomimetic; potential energy surface; density functional theory; oxidative decarboxylation

## 1. Introduction

In recent years, the employment of biomimetic catalysts considerably increased not only for their related high selectivity, but because they are allowed to operate in very mild conditions and they adapt to different types of substrates [1,2].

Encouraged by the new reactions catalyzed by Rieske dioxygenases [1,3,4], sustainable mononuclear nonheme iron complexes have been proposed for bioinspired oxidation catalysis. Dioxygenases are involved in the degradation of a wide range of natural and synthetic compounds by incorporating both atoms of O<sub>2</sub> into substrates.

Bioinspired catalysis using peroxides or peracids for the oxidation of hydrocarbon substrates is more common and better characterized, than that based on dioxygen in catalytic oxidations, which has not been fully explored yet and the relative reaction mechanism is often unknown. This last reaction path requires both electron and proton sources [5–13] and cosubstrate-assisted O<sub>2</sub> molecular activation by transition metal containing complexes can be an alternative pathway [14–19] resorting to the sacrificial reductants represented a good solution [12,20]. Paine's and other groups have investigated a family of iron complexes supported by the monoanionic facial trinitrogen donor ligand, hydrotris(3,5-diphenylpyrazole-1-yl)borate (Tp<sup>Ph2</sup> ligand) that activates O<sub>2</sub> by 2 electrons reduction in the oxidative decarboxylation of the  $\alpha$ -hydroxy acids as sacrificial reductants [3,21–23]. These studies provide evidence that the the hydroxyl group in the dioxygen activation is important. The iron complexes mentioned above reacting with O<sub>2</sub> are able to carry out the stoichiometric cis-dihydroxylation of olefins producing the diol product [3], as in that with H<sub>2</sub>O<sub>2</sub> [4,24,25]. The mononuclear non-heme iron(II) complex shown in Scheme 1, the object of the present investigation, represents the first biomimetic complex reproducing the dioxygenase activity of  $\alpha$ -ketoglutarate-dependent enzymes.



**Scheme 1.** Structure of hydrotris(3,5-diphenylpyrazole-1-yl)borate FeII (mandelate).

Since, as occurred in similar systems [26] no intermediate was detected in the course of the reaction, the steps up to the oxidative decarboxylation of the keto acid play a key role in the control of the rate of the reaction.

To provide deeper insights on the reaction catalyzed by the biomimetic complex of Scheme 1, we have undertaken a theoretical investigation taking into account the spin state of the analyzed species and building up the potential energy surfaces for the reaction path leading to the formation of an iron–oxygen oxidant upon oxidative decarboxylation of  $\alpha$ -hydroxy acid.

## 2. Results and Discussion

As a first step of work, we have optimized the naked biomimetic iron(II)  $\alpha$ -hydroxy acid complex (represented by model **a** in Figure S1 of the Supplementary Information) in septet, quintet, triplet and singlet spin states.

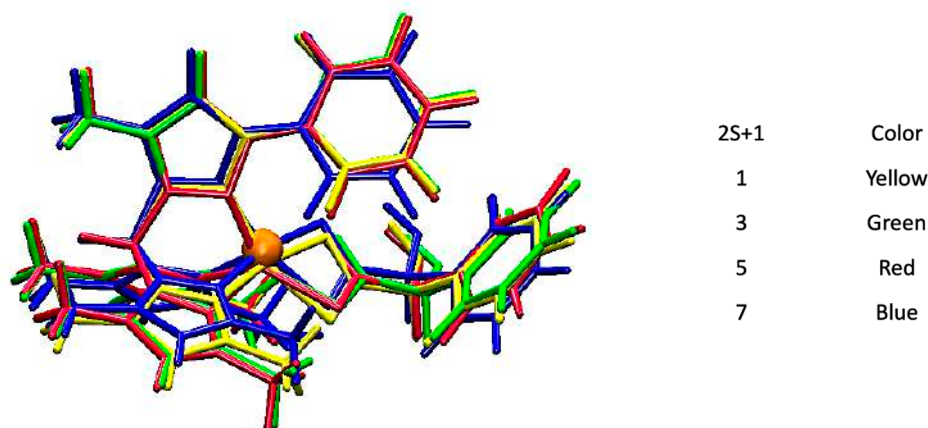
The coordination bond distances, which we have obtained together with X-ray counterpart [3] and relative electronic energies are collected in Table 1. Results show that the complex assumes the most stable configuration in the high-spin quintet state in well agreement with the available data for similar systems [3,27,28]. The singlet, triplet and septet spin states lie at 17.2, 13.1, and 65.1 kcal/mol above ground one, respectively.

**Table 1.** Main geometrical parameters for naked complex **a** at all the spin states. Relative energies provided at M06L level of theory are kcal/mol and compared with the experimental X-ray counterpart [3]. Distances are in Å. See Scheme 1 for labels.

Bond	2S + 1				X-Ray
	1	3	5	7	
Fe-O1	2.06 Å	2.04 Å	2.01 Å	2.06 Å	2.01 Å
Fe-O2	2.08 Å	2.03 Å	2.19 Å	2.29 Å	2.17 Å
Fe-N2	1.96 Å	2.18 Å	2.10 Å	2.09 Å	2.17 Å
Fe-N4	1.96 Å	1.95 Å	2.16 Å	2.05 Å	2.10 Å
Fe-N6	1.92 Å	2.00 Å	2.09 Å	1.97 Å	2.11 Å
DE	17.2	13.1	0.0	65.1	-

Comparison with the crystallographic structure (Table 1) shows that the geometrical parameters in the quintet state are reproduced fairly faithfully. Furthermore, the calculated Fe-N bond distances well match with other high-spin ( $\text{Tp}^{\text{Ph}_2}$ )  $\text{Fe}^{\text{II}}$  complexes [23,27].

The maximum deviation results to be less than 0.06 Å. From the superposition of the optimized structures at the considered spin states, it emerges that the mandelate ring shows major deviations whereas the coordination around the iron does not suffer drastic changes (Figure 1). This suggests that in the quintet, the carboxylic group of the mandelic acid acts as monodentate ligand.



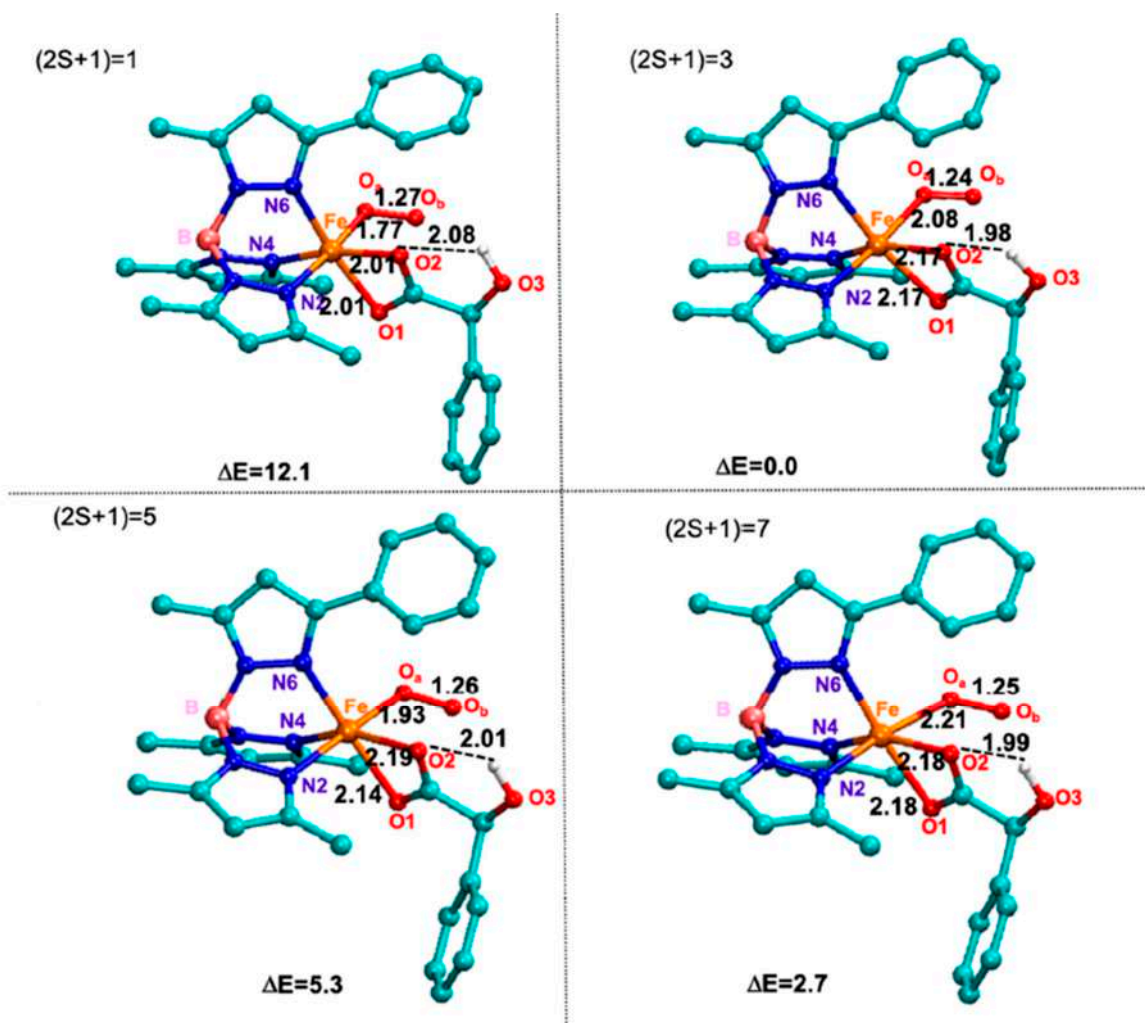
**Figure 1.** Superposition of the M06L optimized structures of model **a** at the different spin states considered.

For the  $\text{O}_2$  addition (**a**- $\text{O}_2$  complex), two different binding modes, namely end-on (eo) and side-on (so), have been taken into account. From computations, the eo binding mode appears to be the preferred one in all the considered spin states. Looking at the Figure 2 collecting the optimized structures of the eo obtained adducts, the triplet **a**- $\text{O}_2$  complex represents the lowest energy configuration followed by the septet, quintet and singlet. A support to this finding arises from the analysis of the triplet spin state HOMO frontier orbital (see Figure 3) that mainly involves the metal center and its surrounding with the mandelate ligand less interested in the occurring electron redistribution.

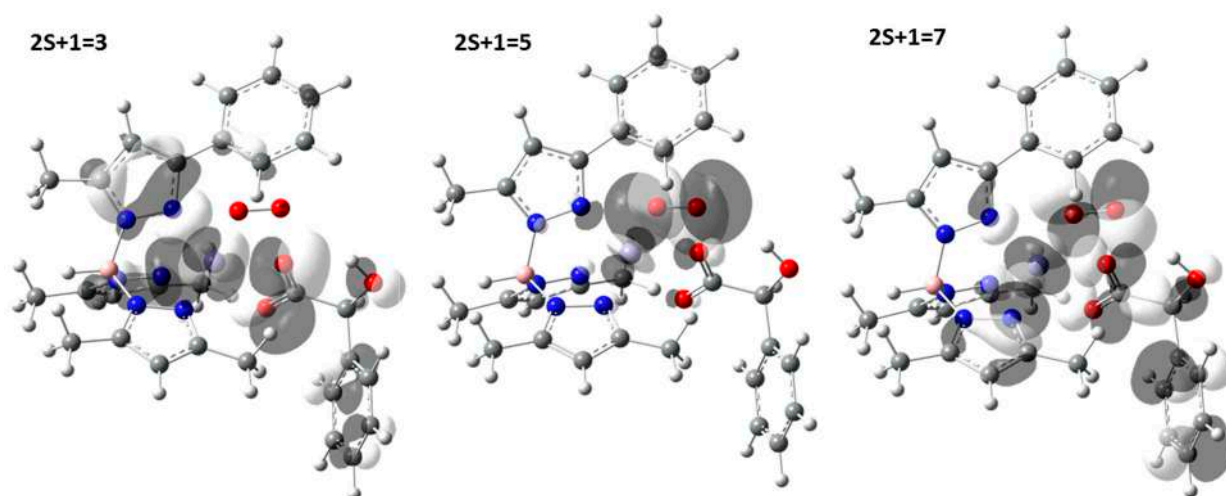
As concerns the geometrical structure, the entry of dioxygen induces a reorganization around the metal center that now appears to be hexacoordinated (octahedral geometry). The  $\text{O}_2$  occupies the apical position, while the carboxylate moiety (lying on the equatorial plane) acts as bidentate ligand with both the Fe-O bond lengths of 2.17 Å. These values well compare with those present in dioxygenase enzymes [20,26].

The Fe- $\text{O}_a$  bond assumes the value of 2.08, 1.93, and 2.21 Å in the triplet, quintet and septet electronic states, respectively. The O-O bond distance in the ground state (1.24 Å) reveals a significant deviation from the corresponding value in the enzymatic environment (1.31 Å). This behavior can be ascribable to the weaker electro-donating character of  $\text{Tp}^{\text{Ph}_2}$  ligand towards the iron(III), present in the **a**- $\text{O}_2$  complex, relative to the usual biological ligands in the enzymes [26].

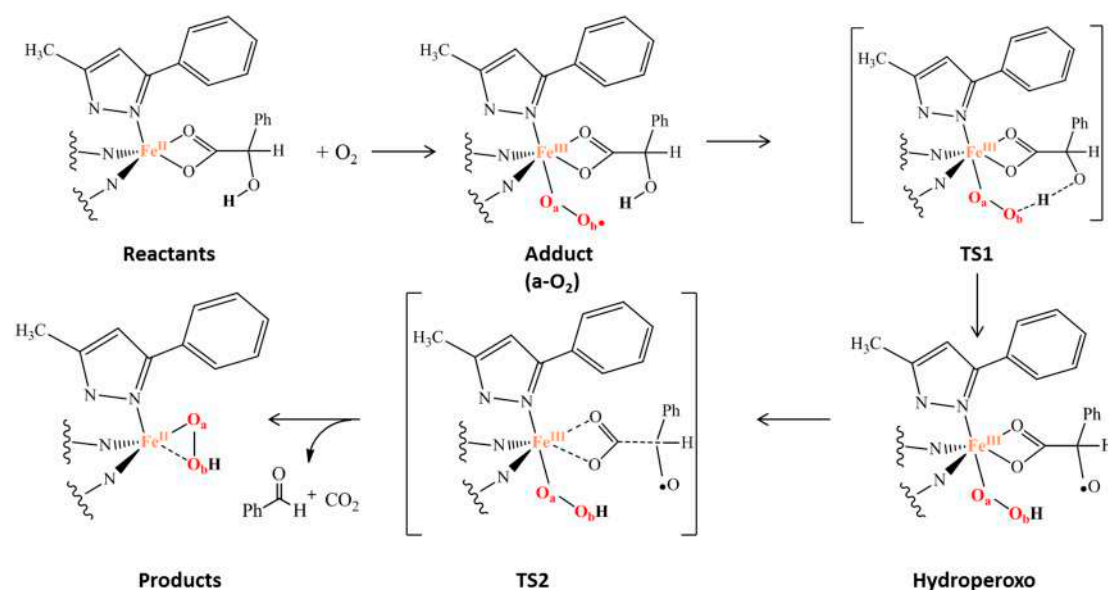
The considered reaction mechanism (Scheme 2) arises from experimental indications. In fact, biomimetic and enzymatic studies suggested that the iron(III) superoxide species, obtained upon addition of  $\text{O}_2$  at the iron(II) center, has the ability to abstract hydrogen atom from the O-H group of  $\alpha$ -hydroxy acid.



**Figure 2.** M06L optimized structures and main geometrical parameters of  $a\text{-O}_2$  adducts. Selected bonds are in Å. Relative electronic energies are in kcal/mol.

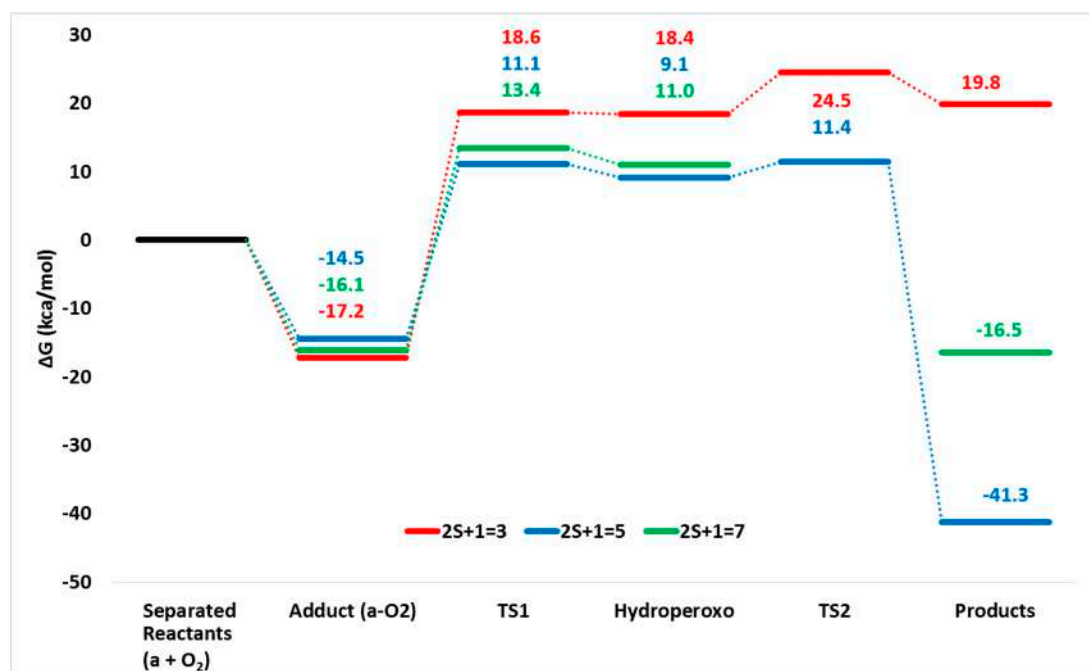


**Figure 3.** Highest occupied molecular orbital (HOMO) composition of  $a\text{-O}_2$  adduct, in the triplet, quintet and septet states.



**Scheme 2.** Proposed mechanism for the oxidative decarboxylation promoted by the non-heme biomimetic complex model (a).

Since the singlet electronic spin state of iron-dioxygen complexes lies at very high energy, only the septet, quintet and triplet potential energy surfaces (PESs) have been considered (see Figure 4).



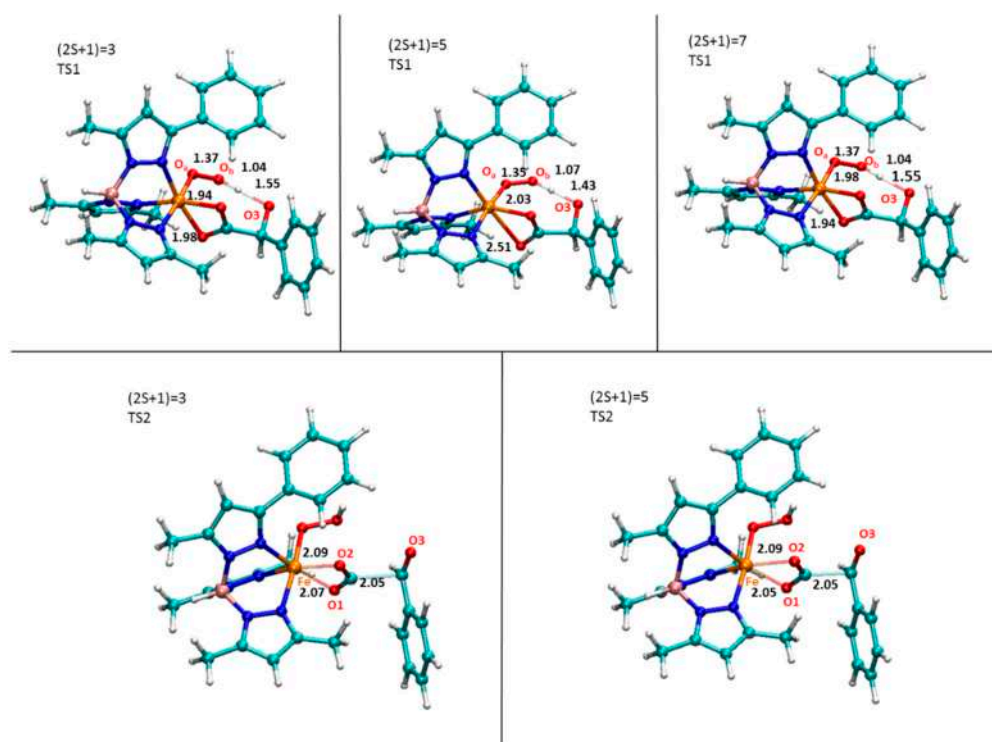
**Figure 4.** M06L potential energy surfaces (PESs) for oxidative decarboxylation operated by non-heme biomimetic model (a) at different spin states (triplet, quintet and septet).

As far as the PES for the septet is concerned, no transition state connecting the hydroperoxide with the final oxidant agent (TS2 of Scheme 2) has been intercepted along the surface. A similar finding has been previously obtained for two enzymatic systems [26,29]. Instead, on the quintet and triplet PESs, both TS1 and TS2 have been characterized confirming that the oxidative decarboxylation is a multistep process.



The addition of O<sub>2</sub> molecule starting to the complex (a) leads to the formation of adducts that in all the three considered spin states lie below the reactants' asymptote suggesting as their formation is an exergonic process (see Figure 4).

On the three computed potential energy surfaces, the first transition state (TS1) is located at 18.6 (triplet), 11.1 (quintet) and 13.4 (septet) kcal/mol and their corresponding optimized structures are shown in Figure 5.



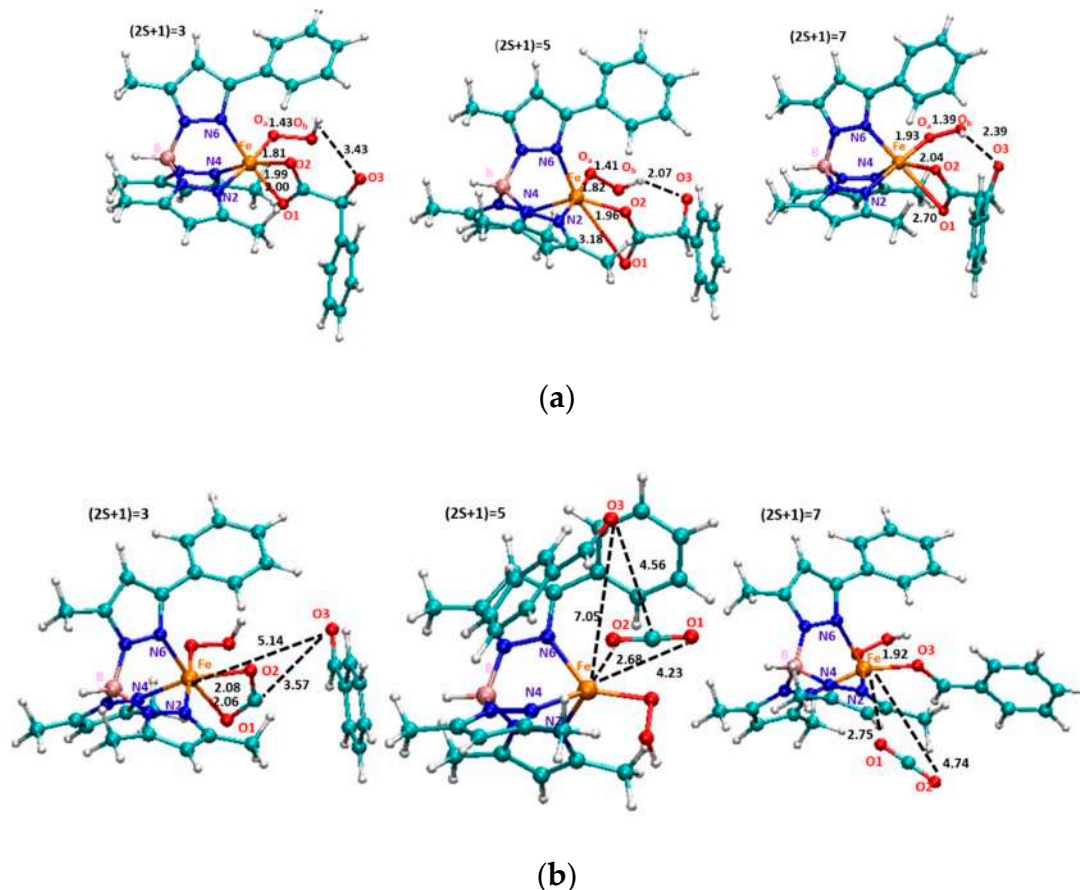
**Figure 5.** M06L optimized structures of TS1 and TS2 intercepted along the three explored PESs. Selected bonds are in Å.

For the TS1 in the triplet state, the barrier is 7.5 kcal/mol higher than that of quintet and a spin crossing from these two PESs occurs. This is due to the fact that the resulting iron(III) hydroperoxo species prefers to be a quintet (9.1 kcal/mol), while the corresponding in the triplet state is 9.3 higher in energy. The TS1 describes the proton transfer from the  $\alpha$ -OH group of the mandelate to the O<sub>b</sub> atom of the molecular oxygen, as confirmed by the imaginary frequency ( $638.1i\text{ cm}^{-1}$ ). This result enhances the role played by the  $\alpha$ -hydroxy acid ligand in this process as proposed by Paine et al. [3].

The next formed species, containing the Fe(III)-OOH moiety, is found at 9.1 kcal/mol and 11.0 kcal/mol in the quintet and septet state, respectively. In contrast, the iron(III) hydroperoxo in the triplet state lies at higher energy. This finding reflects the relative geometrical differences (see Figure 6a). In fact, in the quintet state, the carboxylate moiety is monocoordinated to iron (O<sub>1</sub>-Fe: 1.96 Å, O<sub>2</sub>-Fe: 3.18 Å) and the C-C bond results to be more elongated (1.71 Å) with respect to the previous steps. The benzaldehyde is prone to be released but it still retained by an H bond between its carbonyl (2.07 Å) and the OOH moiety. From the iron(III) hydroperoxo, the reaction proceeds with the definitive C-C bond breaking (2.05 Å) via TS2 that is characterized by an imaginary frequency related to the C-C stretching mode ( $211.8i\text{ cm}^{-1}$ ) as shown in Figure 4. In addition, this bond's elongation generates the loss of the H bond since the OH-O=C is now 3.35 Å.

The product in its quintet spin state, at 41.3 kcal/mol below the separated reactants, evidences as the decarboxylation is a highly exergonic process. In the product, the CO<sub>2</sub> is already formed (O-C-O angle is 180°) and it lies at 2.67 Å from the metallic center engaged in a sandwich-like topology between

two phenyl rings (ligand and benzaldehyde) (see Figure 6b). Some changes occur also in the metallic core due to the elongation of  $O_a-O_b$  bond (1.48 Å) and to the  $Fe-O_a$  distance of 1.91 Å. In this way, the iron(II) hydroperoxo oxidant agent is formed and ready to act as nucleophilic agent in the next oxidation reactions that are not object of the present study.



**Figure 6.** M06L optimized structures and main geometrical parameters of hyperperoxo (a) and prod species (b) in all the three electronic states considered. Selected bonds are in Å.

Paine et al. underline as “this iron(II) hydroperoxo species may also undergo O-O bond heterolysis to form an iron(IV)-oxo-hydroxo intermediate and both oxidant species can carry out the oxidation on different kind of substrates” [3]. To give a contribution to this aspect, geometry optimizations of the two forms have been performed and the resulting optimized geometries (shown in Figure S2) indicate that the quintet spin state proves to be the most stable one.

### 3. Conclusions

On the basis of our density functional theory investigation on the oxidative decarboxylation promoted by a biomimetic nonheme complex, the following conclusions can be drawn:

1. The favored reaction path occurs in the quintet spin state as in the corresponding Rieske oxygenase enzymes;
2. A spin cross takes place between the triplet and quintet electronic spin states after the adduct formation;
3. The  $\alpha$ -hydroxy group plays an important role in the oxidative decarboxylation;
4. The rate determining step is the proton transfer;
5. The overall reaction results to be more exergonic in the quintet PES.



#### 4. Computational Details

All density functional theory (DFT) based calculations have been performed by using Gaussian09 code [30]. Following previous suggestions available in literature about the reliable results of the M06L [31,32] exchange-correlation functional in describing electronic structure and properties of metal containing systems also in various spin states [33–36], this functional has been adopted for the geometry optimizations.

The crystal structure (X ray) of the biomimetic catalyst [3] has been adopted as starting structure for our computations. In order to reduce the computational time, five phenyl rings of the Tp<sup>Ph2</sup> ligand have been replaced by methyl groups (see a of Figure S1). The phenyl group of the Tp<sup>Ph2</sup> ligand closely oriented to the iron center has been retained to better evaluate its role on the activated dioxygen species during the catalytic process.

Geometry optimizations and vibrational frequencies have been performed employing differentiated basis set. In particular, for O, N and H atoms in the active center the 6-31+G(d,p) set has been used while for the remaining ones the 6-31G(d,p) one has been chosen. The LANL2DZ effective core potential and the related basis set has been employed for iron.

In order to improve the energetic values, single point calculations with the larger basis set 6-311+G(d,p) for all atoms except for metal ion, has been performed on all the previously optimized structures. The solvent corrections were evaluated using the polarizable continuum model [37,38] as implemented in Gaussian09 taking into account the dielectric constant ( $\epsilon = 24.93$ ) that better reproduces the polarity of the solvents mixture used in the experimental work [3].

Owing to the unrestricted formalism of all calculations, the spin contamination and its effect on the calculated energetics was monitored as done for similar systems [35,39] and the  $\langle S^2 \rangle$  values are collected in Supplementary Information section (Table S1).

**Supplementary Materials:** Supplementary materials can be found online.

**Author Contributions:** T.M. and N.R. conceived of the presented idea and supervised the project. A.P. and M.C.M. performed the computations. T.M. verified and analyzed the data. T.M. and N.R. wrote the manuscript with support from M.T., A.P. contributed to the final version of the manuscript, drafted the manuscript and designed the figures. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** The University of Calabria is gratefully acknowledged.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### Abbreviations

Tp <sup>Ph2</sup>	hydrotris (3,5-diphenylpyrazole-1-yl) borate
DFT	Density Functional Theory
PES	Potential Energy Surface
PCM	Polarizable Continuum Model
HOMO	High Occupied Molecular Orbital

#### References

1. Costas, M.; Mehn, M.P.; Jensen, M.P.; Que, L., Jr. Dioxygen activation at mononuclear nonheme iron active sites: Enzymes, models, and intermediates. *Chem. Rev.* **2004**, *104*, 939–986. [[CrossRef](#)] [[PubMed](#)]
2. Talsi, E.P.; Bryliakov, K.P. Chemo- and stereoselective CH oxidations and epoxidations/cis-dihydroxylations with H<sub>2</sub>O<sub>2</sub>, catalyzed by non-heme iron and manganese complexes. *Coord. Chem. Rev.* **2012**, *256*, 1418–1434. [[CrossRef](#)]
3. Paria, S.; Chatterjee, S.; Paine, T.K. Reactivity of an Iron–Oxygen Oxidant Generated upon Oxidative Decarboxylation of Biomimetic Iron(II)  $\alpha$ -Hydroxy Acid Complexes. *Inorg. Chem.* **2014**, *53*, 2810–2821. [[CrossRef](#)] [[PubMed](#)]

4. Chen, K.; Costas, M.; Kim, J.; Tipton, A.K.; Que, L., Jr. Olefin Cis-Dihydroxylation versus Epoxidation by Non-Heme Iron Catalysts: Two Faces of an Fe<sup>III</sup>–OOH Coin. *J. Am. Chem. Soc.* **2002**, *124*, 3026–3035. [[CrossRef](#)] [[PubMed](#)]
5. Thibon, A.; England, J.; Martinho, M.; Young, V.G., Jr.; Frisch, J.R.; Guillot, R.; Girerd, J.-J.; Münck, E.; Que, L., Jr.; Banse, F. Proton- and Reductant-Assisted Dioxygen Activation by a Nonheme Iron(II) Complex to Form an Oxoiron(IV) Intermediate. *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 7064–7067. [[CrossRef](#)] [[PubMed](#)]
6. Hong, S.; Lee, Y.-M.; Shin, W.; Fukuzumi, S.; Nam, W. Dioxygen Activation by Mononuclear Nonheme Iron(II) Complexes Generates Iron–Oxygen Intermediates in the Presence of an NADH Analogue and Proton. *J. Am. Chem. Soc.* **2009**, *131*, 13910–13911. [[CrossRef](#)]
7. Mukherjee, A.; Martinho, M.; Bominaar, E.L.; Münck, E.; Que, L., Jr. Shape-Selective Interception by Hydrocarbons of the O<sub>2</sub>- Derived Oxidant of a Biomimetic Nonheme Iron Complex. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 1780–1783. [[CrossRef](#)]
8. Martinho, M.; Blain, G.; Banse, F. Activation of dioxygen by a mononuclear non-heme iron complex: Characterization of a Fe(III)-(OOH) intermediate. *Dalton Trans.* **2010**, *39*, 1630–1634. [[CrossRef](#)]
9. Lee, Y.-M.; Hong, S.; Morimoto, Y.; Shin, W.; Fukuzumi, S.; Nam, W. Dioxygen Activation by a Non-Heme Iron(II) Complex: Formation of an Iron(IV)–Oxo Complex via C–H Activation by a Putative Iron(III)–Superoxo Species. *J. Am. Chem. Soc.* **2010**, *132*, 10668–10670. [[CrossRef](#)]
10. Mandon, D.; Jaafar, H.; Thibon, A. Exploring the Oxygen sensitivity of FeCl<sub>2</sub> complexes with tris(2-pyridylmethyl)amine-type ligands: O<sub>2</sub> coordination and a quest for superoxide. *N. J. Chem.* **2011**, *35*, 1986–2000. [[CrossRef](#)]
11. Badiei, Y.M.; Siegler, M.A.; Goldberg, D.P. O<sub>2</sub> Activation by Bis(imino)pyridine Iron(II)-Thiolate Complexes. *J. Am. Chem. Soc.* **2011**, *133*, 1274–1277. [[CrossRef](#)] [[PubMed](#)]
12. He, Y.; Goldsmith, R. Observation of a ferric hydroperoxide complex during the non-heme iron catalysed oxidation of alkenes and alkanes by O<sub>2</sub>. *Chem. Commun.* **2012**, *48*, 10532–10534. [[CrossRef](#)] [[PubMed](#)]
13. Li, F.; Van Heuvelen, K.M.; Meier, K.K.; Münck, E.; Que, L., Jr. Sc<sup>3+</sup>-Triggered Oxoiron(IV) Formation from O<sub>2</sub> and its Non-Heme Iron(II) Precursor via a Sc<sup>3+</sup>–Peroxo–Fe<sup>3+</sup> Intermediate. *J. Am. Chem. Soc.* **2013**, *135*, 10198–10201. [[CrossRef](#)] [[PubMed](#)]
14. Kim, S.O.; Sastri, C.V.; Seo, M.S.; Kim, J.; Nam, W. Dioxygen Activation and Catalytic Aerobic Oxidation by a Mononuclear Nonheme Iron(II) Complex. *J. Am. Chem. Soc.* **2005**, *127*, 4178–4179. [[CrossRef](#)]
15. Bruijninx, P.C.A.; van Koten, G.; Gebbink, R.J.M.K. Mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad: Recent developments in enzymology and modeling studies. *Chem. Soc. Rev.* **2008**, *37*, 2716–2744. [[CrossRef](#)]
16. Shook, R.L.; Peterson, S.M.; Greaves, J.; Moore, C.; Rheingold, A.L.; Borovik, A.S. Catalytic Reduction of Dioxygen to Water with a Monomeric Manganese Complex at Room Temperature. *J. Am. Chem. Soc.* **2011**, *133*, 5810–5817. [[CrossRef](#)]
17. Paine, T.K.; Que, L., Jr. Dioxygen Activation by Biomimetic Iron(II) Complexes of  $\alpha$ -Keto Acids and  $\alpha$ -Hydroxy Acids. *Struct. Bond.* **2014**, *160*, 39–56.
18. Sallmann, M.; Limberg, C. Utilizing the Trispyrazolyl Borate Ligand for the Mimicking of O<sub>2</sub>-Activating Mononuclear Nonheme Iron Enzymes. *Acc. Chem. Res.* **2015**, *48*, 2734–2743. [[CrossRef](#)]
19. Sahu, S.; Goldberg, D.P. Activation of Dioxygen by Iron and Manganese Complexes: A Heme and Nonheme Perspective. *J. Am. Chem. Soc.* **2016**, *138*, 11410–11428. [[CrossRef](#)]
20. Funabiki, T. Functional model oxygenations by nonheme iron complexes. In *Advances in Catalytic Activation of Dioxygen by Metal Complexes*; Simandi, L.I., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003; pp. 157–226.
21. Chatterjee, S.; Paine, T.K. Olefin cis-Dihydroxylation and Aliphatic C-H Bond Oxygenation by a Dioxygen-Derived Electrophilic Iron–Oxygen Oxidant. *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 9338–9342. [[CrossRef](#)]
22. Chatterjee, S.; Paine, T.K. Hydroxylation versus Halogenation of Aliphatic C-H Bonds by a Dioxygen-Derived Iron–Oxygen Oxidant: Functional Mimicking of Iron Halogenases. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 7717–7722. [[CrossRef](#)] [[PubMed](#)]
23. Paria, S.; Que, L., Jr.; Paine, T.K. Oxidative Decarboxylation of Benzilic Acid by a Biomimetic Iron(II) Complex: Evidence for an Iron(IV)–Oxo–Hydroxo Oxidant from O<sub>2</sub>. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 11129–11132. [[CrossRef](#)] [[PubMed](#)]

24. Costas, M.; Que, L., Jr. Ligand Topology Tuning of Iron-Catalyzed Hydrocarbon Oxidations. *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 2179–2181. [[CrossRef](#)]
25. Oldenburg, P.D.; Feng, Y.; Pryjomska-Ray, I.; Ness, D.; Que, L., Jr. Olefin cis-dihydroxylation with bio-inspired iron catalysts. evidence for an Fe(II)/Fe(IV) catalytic cycle. *J. Am. Chem. Soc.* **2010**, *132*, 17713–17723.
26. Borowski, T.; Bassan, A.; Siegbahn, P.E.M. A Hybrid Density Functional Study of O–O Bond Cleavage and Phenyl Ring Hydroxylation for a Biomimetic Non-Heme Iron Complex. *Inorg. Chem.* **2004**, *43*, 3277–3291. [[CrossRef](#)]
27. Mehn, M.P.; Fujisawa, K.; Hegg, E.L.; Que, L., Jr. Oxygen Activation by Nonheme Iron(II) Complexes:  $\alpha$ -Keto Carboxylate versus Carboxylate. *J. Am. Chem. Soc.* **2003**, *125*, 7828–7842. [[CrossRef](#)]
28. Siegbahn, P.E.M.; Borowski, T. Modeling Enzymatic Reactions Involving Transition Metals. *Acc. Chem. Res.* **2006**, *39*, 729–738. [[CrossRef](#)]
29. Borowski, T.; Bassan, A.; Siegbahn, P.E.M. Mechanism of Dioxygen Activation in 2-Oxoglutarate-Dependent Enzymes: A Hybrid DFT Study. *Chem. Eur. J.* **2004**, *10*, 1031–1041. [[CrossRef](#)]
30. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Zakrzewski, V.G.; Montgomery, J.A., Jr.; Stratmann, R.E.; Burant, J.C.; et al. *Gaussian 2009*; Gaussian, Inc.: Wallingford, CT, USA, 2009.
31. Zhao, Y.; Truhlar, D.G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006**, *125*, 194101. [[CrossRef](#)]
32. Valero, R.; Costa, R.; Moreira, I.D.P.R.; Truhlar, D.G.; Illas, F. Performance of the M06 family of exchange-correlation functionals for predicting magnetic coupling in organic and inorganic molecules. *J. Chem. Phys.* **2008**, *128*, 114103. [[CrossRef](#)]
33. Marzouk, A.; Bolvin, H.; Reinhardt, P.; Manceron, L.; Perchard, J.P.; Tremblay, B.; Alikhani, M.E. A Combined Experimental and Theoretical Study of the  $Ti^{2+} + N_2O$  Reaction. *J. Phys. Chem. A* **2014**, *118*, 561–572. [[CrossRef](#)]
34. Gusev, D.G. Assessing the Accuracy of M06-L Organometallic Thermochemistry. *Organometallics* **2013**, *32*, 4239–4243. [[CrossRef](#)]
35. Marino, T.; Fortino, M.G.; Russo, N.; Toscano, M.; Alberto, M.E. Computational Mechanistic Insights on the NO Oxidation Reaction Catalyzed by Non-Heme Biomimetic Cr-N-Tetramethylated Cyclam Complexes. *Int. J. Mol. Sci.* **2019**, *20*, 3955. [[CrossRef](#)] [[PubMed](#)]
36. Fortino, M.; Marino, T.; Russo, N. Theoretical Study of Silver-Ion-Mediated Base Pairs: The Case of C–Ag–C and C–Ag–A Systems. *J. Phys. Chem. A* **2015**, *119*, 5153–5157. [[CrossRef](#)] [[PubMed](#)]
37. Boone, A.J.; Chang, C.H.; Greene, S.N.; Herz, T.; Richards, N.G.J. Modeling the spin-dependent properties of open-shell Fe(III)-containing systems: Towards a computational description of nitrile hydratase. *Coord. Chem. Rev.* **2003**, *238–239*, 291–314. [[CrossRef](#)]
38. Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001. [[CrossRef](#)]
39. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24*, 669–681. [[CrossRef](#)] [[PubMed](#)]

**Sample Availability:** Samples of the compounds are not available from the authors.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).





**Paper II**  
The platination mechanism of RNase A by arsenoplatin: insight from the theoretical study



## RESEARCH ARTICLE



Cite this: *Inorg. Chem. Front.*, 2021, **8**, 1795

# The platination mechanism of RNase A by arsenoplatin: insight from the theoretical study†

A. Parise, <sup>a,b</sup> N. Russo <sup>a</sup> and T. Marino \*<sup>a</sup>

Herein, we present a detailed metalation process of the bovine pancreatic ribonuclease (RNase A) by a novel multitarget anti-cancer agent arsenoplatin-1, **AP1**, ( $[\text{Pt}(\mu\text{-NHC}(\text{CH}_3)\text{O})_2\text{ClAs}(\text{OH})_2]$ ), on the basis of quantum chemical investigation, employing the B3LYP and M062x functionals and a large model of the active site. The proposed mechanism is consistent with the structural data. The role of water molecules in the active site is also analyzed. These studies revealed that the N $\delta$  of His119 binds platinum(II), preserving the Pt–As bond. To better rationalize the different behavior of **AP1** (bound to His119) and the reference cisplatin (bound to Met29) towards the same target RNase A, also these processes have been considered. The final platinated complex structure agrees well with the crystallographic one. Our results evidence that the metalation process takes place more favorably in water than in the protein environment in agreement with the nature of the protein binding pocket residues.

Received 25th September 2020,  
Accepted 18th January 2021

DOI: 10.1039/d0qi01165g

rs.c.li/frontiers-inorganic

## Introduction

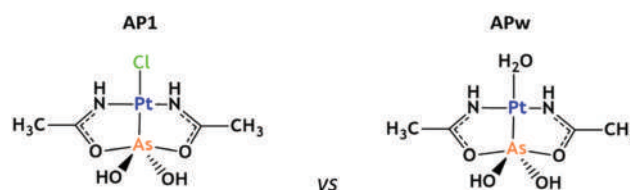
The O'Halloran group successfully developed a synthesis method to obtain good amounts of stable aqueous arsenous acid–platinum complex with biological activity that is distinct from cisplatin and arsenic trioxide individually. This newly synthesized compound, named arsenoplatin-1 (**AP1**),  $[\text{Pt}(\mu\text{-NHC}(\text{CH}_3)\text{O})_2\text{ClAs}(\text{OH})_2]$  (Scheme 1), presented anticancer biological activity in several cancer cell lines and is stable in an aqueous environment. In addition, it can be beneficial to load it into a liposome based drug delivery system. With the aim of improving the efficacy of these two common anti-cancer therapeutics, a method to co-encapsulate high levels of arsenous acids and aqua-cisPt into 100 nm liposomes, capable of releasing their drug in the low-pH endosome, has been developed.<sup>1</sup>

Arseoplatins are a good example of hybrid bifunctional compounds obtained from merging the pharmacophores of active molecules featuring different mechanisms of action. They are adducts of two chemically important anticancer drugs, cisplatin and arsenic trioxide, characterized by an arsenous acid moiety covalently linked to the platinum(II) center equipped with an unusual five coordinate As(III)-geometry. This can be interpreted, in the drug discovery field, as a

borderline choice between bio-inspired and rational design generating molecules with a dual mode of action for creating new efficient drugs that can act on different biological targets. Preliminary tests *in vitro* aimed to evaluate the cytotoxicity proposed that arsenoplatins may have a distinct mode of action in comparison with cisplatin and  $\text{As}_2\text{O}_3$  alone, other than the ability to overcome platinum resistance mechanisms.<sup>2</sup>

Arseoplatin-1 (**AP1**) resulted to be more potent than either cisplatin or arsenic trioxide in most of the cancer cell lines evaluated. This behavior was deemed attributable to a possible cooperative effect between platinum and arsenic pharmacophores detected in lungs,<sup>3</sup> oral squamous cell carcinoma,<sup>4</sup> and ovarian cancer cell lines.<sup>5</sup>

Despite the insights gained into the interaction of cisplatin and **AP1** with DNA at theoretical and experimental levels,<sup>2,6</sup> those concerning **AP1** and proteins are often lacking. Recently, with the aim of better learning its biological chemistry, a novel investigative strategy based on the combined use of electrospray ionization mass spectrometry (ESI MS) and X-ray crystallography<sup>7,8</sup> allowed the in-depth investigation of the



**Scheme 1** Representation of the **AP1** complex and its aquated form **APw**.

<sup>a</sup>Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy.

E-mail: tiziana.marino65@unical.it

<sup>b</sup>Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0qi01165g



metalation process of small model proteins. In particular, the structural information of the hen egg white lysozyme (HEWL) and bovine pancreatic ribonuclease (RNase A), interacting with cisplatin, is available.<sup>9</sup> Recent structural studies have shown that **AP1** bound to RNase A with platinum(II) that binds to the N-atoms of the solvent exposed His105 and His119 side chains on the surface of the protein preserving the Pt–As bond.<sup>10</sup> Since His119 residues are implicated in the catalytic activity of the RNase A enzyme, the binding of **AP1** on this site results in the inhibition of the enzyme function, as indicated by a catalytic activity assay. Unlike **AP1**, cisplatin showed the adduct formation with the methionine sulfur atom under the same conditions. In addition, analogous to what was observed with the tested proteins, the same authors observed **AP1** readily enters the cells and binds to DNA, maintaining the Pt–As bond (Pt:As ratio of 1). After longer incubation times, however, the Pt:As ratio in DNA samples increases, suggesting the occurrence of cleavage of the Pt–As bond and the release of the As(OH)<sub>2</sub> moiety. This finding suggested that arsenoplatin-1 has the potential to deliver both Pt and As species to a variety of hematological and solid cancers with superior effects to those observed in most parts of the NCI-60 human tumor cell lines when compared to the use of cisplatin or arsenic trioxide as single agents.

Starting from these interesting experimental outcomes<sup>2,10</sup> and stimulated by the reliable results of the previous theoretical investigation on the **AP1**–DNA interaction mechanisms,<sup>6</sup> it seemed right, but also intriguing, to explore the inhibition mechanism induced by the metalation process of RNase A at the expense of **AP1**. To do this, a rather large quantum chemical model including amino acid residues present in the catalytic task and suitable to describe electronic structures and related energetic profiles has been adopted. Furthermore, to find an explanation of the different behavior of **AP1** (bound to His119) and cisplatin (bound to Met29) towards the same target RNase A, the mechanism has been explored by using smaller models.

Theoretical calculations represent nowadays an effective way to separate and individually quantify the physical components that contribute to the activation free energy and, therefore, are an optimal tool to characterize all the species including intermediates and transition states, intercepted along a selected reaction pathway, often suggested by the experimental findings.

## Computational methods

All the calculations were performed employing the Gaussian 09.D01 software package.<sup>11</sup> The B3LYP<sup>12,13</sup> (20% HF exchange) and M062x<sup>14</sup> (54% HF exchange) functionals coupled to the 6-31+G(d,p) basis set for the C, N, O, and H atoms were used for the small models (Scheme S1†). The results for geometry optimization, electronic barrier height and electronic energy of reaction are reported in Fig. S1.† On the basis of these results, for larger systems, the M062x functional was employed.

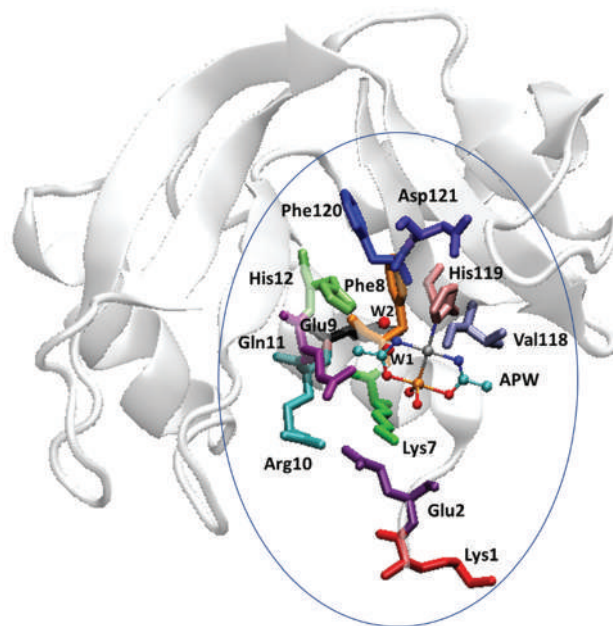
Furthermore, to evaluate the influence of the exchange–correlation functional also single point B3LYP computations on the M06-2x optimized geometries were performed.

For both Pt and As atoms, the effective core potential SDD<sup>15</sup> coupled with their related orbital basis set was selected. The D3 dispersion contribution was considered in all the computations.<sup>16</sup>

In order to avoid large artificial movements of the amino acids in the larger quantum cluster model, the coordinate-locking scheme was assumed in which the amino acid residues are usually truncated at C $\alpha$  atoms that are fixed during geometry optimizations. The stars in Fig. 1 indicate the atoms where the truncation is made. The residues have been modelled according to standard cluster model procedures.<sup>17–24</sup> This adopted procedure assumes that the overall structure is kept very close to the experimental one. The nature of minima (no negative vibrational frequencies) and transition states (one negative vibrational frequency) along potential energy surfaces was established by frequency calculation at the same level of theory. Due to the presence of constrained atoms, small vibrational frequencies <40i cm<sup>-1</sup> have been obtained in all stationary points that do not alter the thermochemical analysis and zero-point energy (ZPE) corrections and do not obscure the main negative TS-frequency. An analogous computational protocol was successfully applied in the mechanistic studies of many enzymes.<sup>23–29</sup>

The transition state has been checked by intrinsic reaction coordinate (IRC) analysis,<sup>30,31</sup> by assessing that the localized transition states correctly connect to the corresponding minima along the imaginary mode of vibration.

To model the protein environment effects, the CPCM<sup>32,33</sup> polarizable continuum model was used with a dielectric con-



**Fig. 1** The used cluster model is shown inside the circle and the arsenoplatin complex is shown as balls and sticks.



stant that simulates the protein ( $\epsilon = 4$ ) and water ( $\epsilon = 80$ ) environments. The last choice well reproduces the nature of the catalytic pocket that is widely solvent exposed.<sup>10</sup> To improve the electronic energies, the more extended 6-311+G(2d,2p)|SDD basis set was adopted in single point energy calculations done by using M062x and B3LYP-D3 on the M062x optimized geometries. In the potential energy surfaces (PESs), the reported energies include the ZPE corrections, excluding the contributions of frozen atoms in vibrational analysis<sup>34</sup> and the standard approximation for free energy. In order to properly take into account the entropic contributions, free energy corrections have been estimated adding a term equal to  $RT \ln(V_{\text{molar gas}}/V_{\text{molar solution}})$ , where  $R$  = gas constant,  $T$  = absolute temperature,  $V_{\text{molar gas}}$  is the volume occupied by one mole of ideal gas at the considered temperature and  $V_{\text{molar solution}}$  is the volume occupied by one mole of species in a standard solution of concentration  $1 \text{ mol L}^{-1}$ , as successfully adopted in other works.<sup>35,36</sup> To obtain internal entropies, we interpolate the rotational and harmonic vibrational approximations.  $S_V$  and  $S_R$  are combined by using a weighting function, following the Head-Gordon damping function, reported in eqn (1) and (2) as proposed by S. Grimme.<sup>18</sup>

$$S = W(\omega)S_V + [1 - W(\omega)]S_R \quad (1)$$

$$W(\omega) = 1/[1 + (\omega_0/\omega)^a]^{-1} \quad (2)$$

NBO<sup>37</sup> analysis was performed on all the stationary points intercepted on the PESs and the related results are given in Table S1.†

## Results and discussion

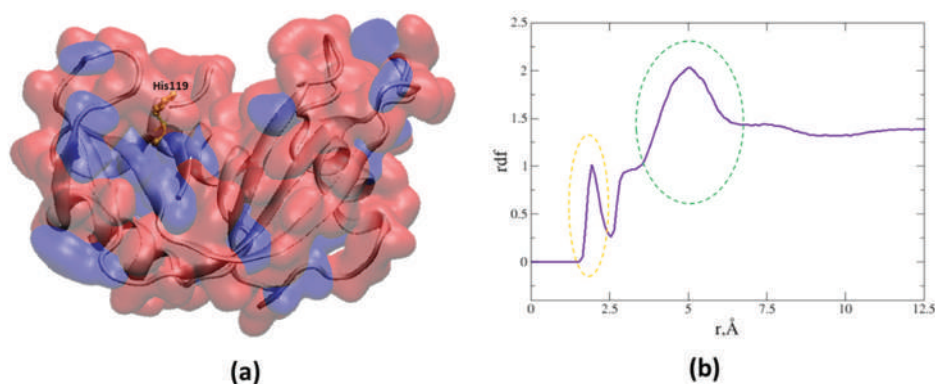
### Active site model

The starting molecular model used for the QM calculations was obtained from the X-ray structure of the adduct formed in the reaction between pancreatic ribonuclease isolated from *Bos taurus*<sup>10</sup> and arsenoplatin-1 at a resolution of  $2.15 \text{ \AA}$  (PDB id:

5NJ7). All protonation states were obtained using Propka 3.0 (see Table S2†).<sup>38,39</sup>

To construct the active site model in the initial enzyme-substrate complex form, the  $N\delta$  (His119)-Pt was broken and modified back to the native form with the  $\text{H}_2\text{O}$  molecule replacing the imidazole ring of the His119 residue. The resulting cluster (depicted in Fig. 1) consists of 255 atoms with an overall charge of zero and contains amino acid residues of the active site positioned in the solvent exposed surface of the protein of the catalytic groove, which are Lys1, Glu2, Lys 7, Phe8, Glu9, Arg10, Gln11, His12, Val118, Phe120, and Asp121. RNase A shows a V-shaped kidney structure, with two opposite domains forming an elongated substrate-binding groove between the two arms. The high content of positively charged (lysine and arginine) residues present in the active site cleft is linked to the reaction catalyzed by RNase A consisting of the hydrolysis of the phosphodiesteric bond at the 5'-ribose of the negatively charged substrate, single-stranded RNA.

Other polar residues ensure a dense network of local interactions. The distribution of the positively and negatively charged polar and nonpolar residues is shown in Fig. 2a, where it is possible to see that His119, present along the walls of the active site groove, is surrounded almost exclusively by polar residues. The dominant distribution of these polar amino acids at the protein surface attracts water molecules in the first hydration shell. A detailed analysis of the water molecule distribution in the binding cavity in the RNase A protein has been done by using Molecular Dynamics simulations (see the ESI†). The H-bonding interactions from the MD simulation of the apoprotein are also reported in Table S3.† Fig. 2b shows the water radial distribution function (RDF) obtained as a function of the distance between the water oxygen and the  $N\delta$  of the His119 residue. The observed trend of RDF, evidenced by the sharp peak circled in orange at a distance of  $<2.50 \text{ \AA}$ , indicates that a more ordered water structure is present in the proximity of the protein surface rather than in bulk water where broader peaks are present. This behavior is essentially due to the specific interactions between His119 and water.



**Fig. 2** (a) Structure of RNase A with the backbone evidencing the secondary structures. The surface is illustrated based on the different nature (negative, red and positive, blue) of the residues. (b) The radial distribution function of the bulk water inside  $12.5 \text{ \AA}$  from  $N\delta$  of the His119 residue.

This confirms the presence of a conspicuous number of water molecules in the crystal that encouraged us to retain two explicit water molecules (**W1** and **W2**) in the QM cluster (Fig. 1).

### Small models

Preliminary calculations were carried out on small models (single amino acids and platinum complex), as depicted in Scheme S1,† in order to evaluate the reactivity of **AP1** in its original form with the chloride ligand (a) or in its corresponding aquated form (b) towards histidine and methionine. The barriers related to the platination process on the small models are reported in Fig. S1.† The results confirmed that the more favorable energetics concern the aquated form of **AP1** ((b) of Scheme S1†) and the histidine resulted in the preferred target in agreement with the experimental observations.<sup>10</sup>

This behavior is similar to that occurring in platinum ligation to DNA purine bases investigated previously.<sup>6</sup> Furthermore, the energy trend obtained using the two functionals (Fig. S1†) is similar and matches well with a recent study performed on analog platinum(II) complexes.<sup>40</sup>

### Reaction mechanism

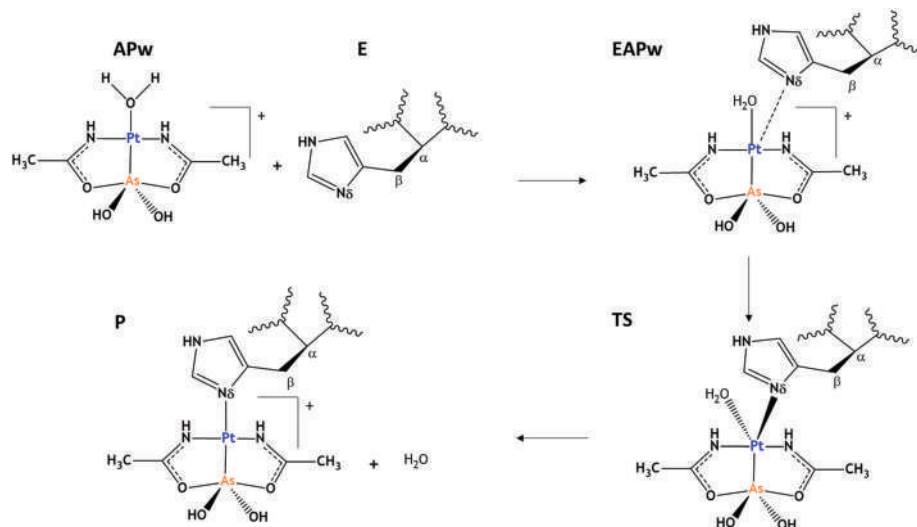
From the water exposure of the binding pocket of RNase A and the results from our calculations on small models (see Fig. S1†), we believed that the ligation to histidine of the RNase A takes place by the aquated form **APw** and not by **AP1** (Scheme 1).

The explored mechanism for the platination of RNase A (studied by employing the larger cluster), shown in Scheme 2, indicated as **E**, proposes a single step process in which, after the barrierless formation of the adduct (**EAPw**), the reaction proceeds towards the covalent product (**P**) through a transition state characterized by the Pt of **APw** linked to N $\delta$  of His119 and the water molecule.

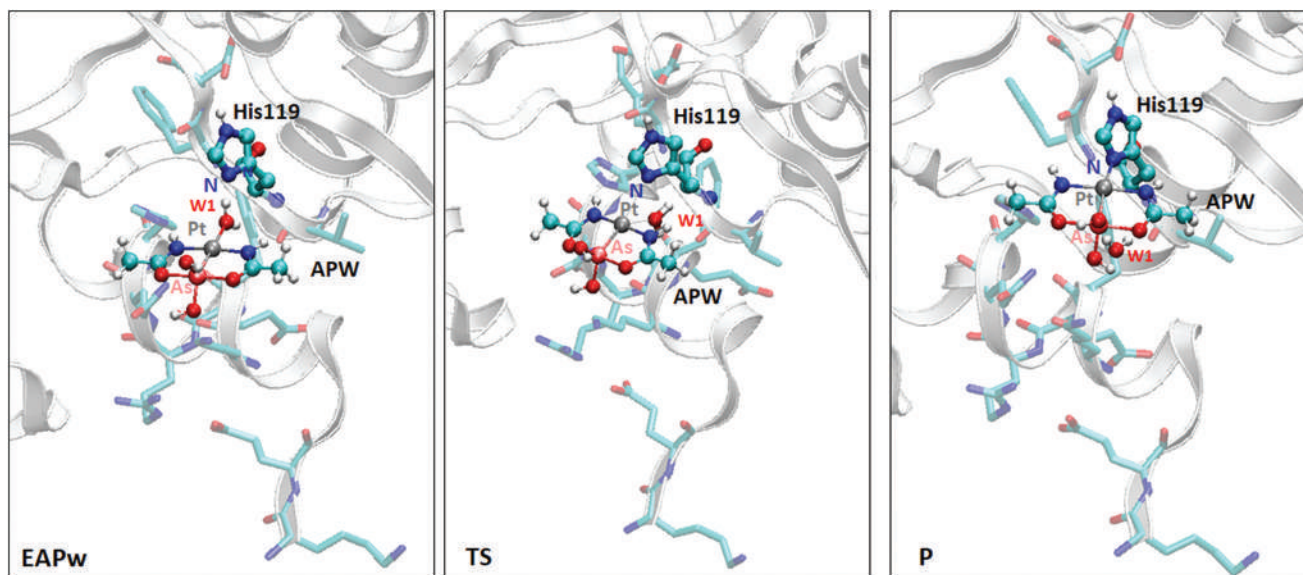
All the M062x/6-31+G(d)|SDD optimized structures for the obtained stationary points are depicted in Fig. 3 along with the main geometrical parameters, while Fig. 4 shows the M062x-D3 (a) and B3LYP-D3 (b) potential energy surfaces in protein and water environments.

In the **EAPw** complex (see Fig. 3), **APw** is found to bind such that the water coordinated to Pt points toward the N $\delta$  of His119 establishing a H-bond of 1.55 Å. An additional H-bond interaction involves the O<sub>3</sub>H moiety and the carbonyl group of the backbone of the Phe120 residue (1.61 Å). The distance between N $\delta$  His119 and Pt, corresponding to the reaction coordinate linking the starting point to the product of the platination process, is equal to 3.83 Å. This confirms that the neutral nucleophile His119 is not still included in the inner coordination shell of the platinum, thus ensuring the planarity around the Pt center since the dihedral angle (N2–N1–O1–O2) is  $-3.54^\circ$ . The **EAPw** complex is 4.30 (3.83) kcal mol<sup>-1</sup> at M062x-D3 (B3LYP-D3) lower than the separated reactants (**E** + **APw**) in the protein environment and the corresponding values in the water phase are 0.90 kcal mol<sup>-1</sup> (M062x-D3) and 2.37 kcal mol<sup>-1</sup> (B3LYP-D3) (Fig. 3).

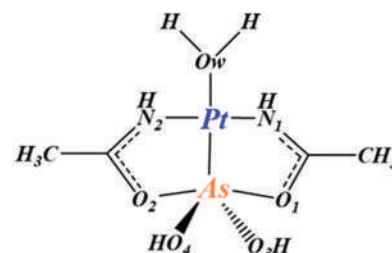
The reaction proceeds towards the product by overcoming the barrier dictated by the **TS** whose optimized structure is depicted in Fig. 3. The RNase A platination process takes place at the expense of His119 (N $\delta$  imidazole ring) that now is located at 2.83 Å, confirming the shortening of the distance by 1 Å with respect to the **EAPw** adduct. The water molecule is placed at 2.16 Å showing only slight lengthening relative to the previous species. This behaviour can be assimilated to an associative-like substitution mechanism in which the incoming ligand (His119) binds to the Pt center before the departing ligand (water). The Pt appears pentacoordinated and arranged in a trigonal bipyramidal geometry as also observed in the corresponding step of the **AP1** aquation process.<sup>6</sup> The imaginary frequency in this **TS** is about 59 cm<sup>-1</sup> and the analysis of



Scheme 2 The mechanism of the metalation process promoted by **APw**.



Distance (Å)	EAPw	TS	P	X-Ray <sup>(a)</sup>
Pt-N <sub>His</sub>	3.83	2.83	2.21	2.10
Pt-As	2.30	2.32	2.32	2.20
Pt-O <sub>w</sub>	2.16	2.30	3.37	4.02
Pt-N1	2.03	2.03	2.03	2.00
Pt-N2	2.04	2.04	2.07	2.00
As-O1	1.94	1.96	1.96	1.98
As-O2	1.99	1.99	1.97	2.02
As-O3	1.73	1.74	1.74	1.72
As-O4	1.74	1.72	1.73	1.72
H <sub>w</sub> -N <sub>His</sub>	1.55	1.75	2.23	4.25
<b>Angle (°)</b>				
N-Pt-O	43.78	61.97	62.69	84.40



**Fig. 3** Optimized geometries of the stationary points at the M062x/6-31G+(d,p)||SDD level of theory. <sup>a</sup> Main distances as defined in the structure are included in the connected table.<sup>10</sup>

the corresponding vibrational mode clearly indicates the stretching of the N $\delta$  His119–Pt bond formation associated with that of metal–water (Pt–O<sub>w</sub>).

The required energetic cost for this step is about 15 kcal mol<sup>-1</sup> using both M062x-D3 and B3LYP-D3 functionals in the protein environment. The same barrier is reduced to about 9 and 10 kcal mol<sup>-1</sup> in water (see Fig. 4). This can be considered the consequence of the “synergic” effect due to the destabilization of the **EAPw** adduct and the stabilization of the transition state observed with both the employed functionals. This evidence further highlights the role played by the first hydration shell of the protein surface around the Pt where the bond cleavage and formation occur.

As evidenced above in the small models, in the QM cluster larger model, both the used functionals provide electronic barrier heights and electronic energies very close in energy in the protein and water environments as in the case of other chemical reactions.<sup>41</sup> Furthermore, the inclusion of D3 dispersion does not influence the energetics of the process.

Looking at the NBO charges collected in Table S1,<sup>†</sup> it is possible to note that the charge on Pt in the **TS** becomes less

positive than that in the adduct due to the effect of the presence of different atoms of the ligand (O<sub>w</sub> in **EAPw** and N $\delta$  in the **TS**) on the metal center. In the final platinated complex (**P**) the imidazole of His119 lies in the trans position to the As(OH)<sub>2</sub> moiety at about 2.21 Å from Pt while the water leaving group is now at 3.37 Å and is in close contact with His119 by an H-bond (2.23 Å). During the platination process the geometrical parameters in the As(OH)<sub>2</sub> and diacetylamido moieties do not suffer significant variations and the Pt–As bond maintains the same distance (see Fig. 2). This behavior is in agreement with that of the experimental counterpart that proposes the Pt–As unit to remain intact upon RNase A binding. The additional water molecule (W2) retained in the QM model lies very close to the platination reaction site establishing an H-bond (2.00 Å) interaction with the oxygen atom of the Val118 residue, and with the backbone N atom of Phe8 with an N–H<sub>w</sub> distance of 2.30 Å.

The stability of the final covalent platinated product with the N $\delta$ –Pt formed bond is confirmed by the  $\Delta G$  value of the process that is a little bit more pronounced with B3LYP-D3 than that with M062x-D3 (see Fig. 4). The exothermicity of the

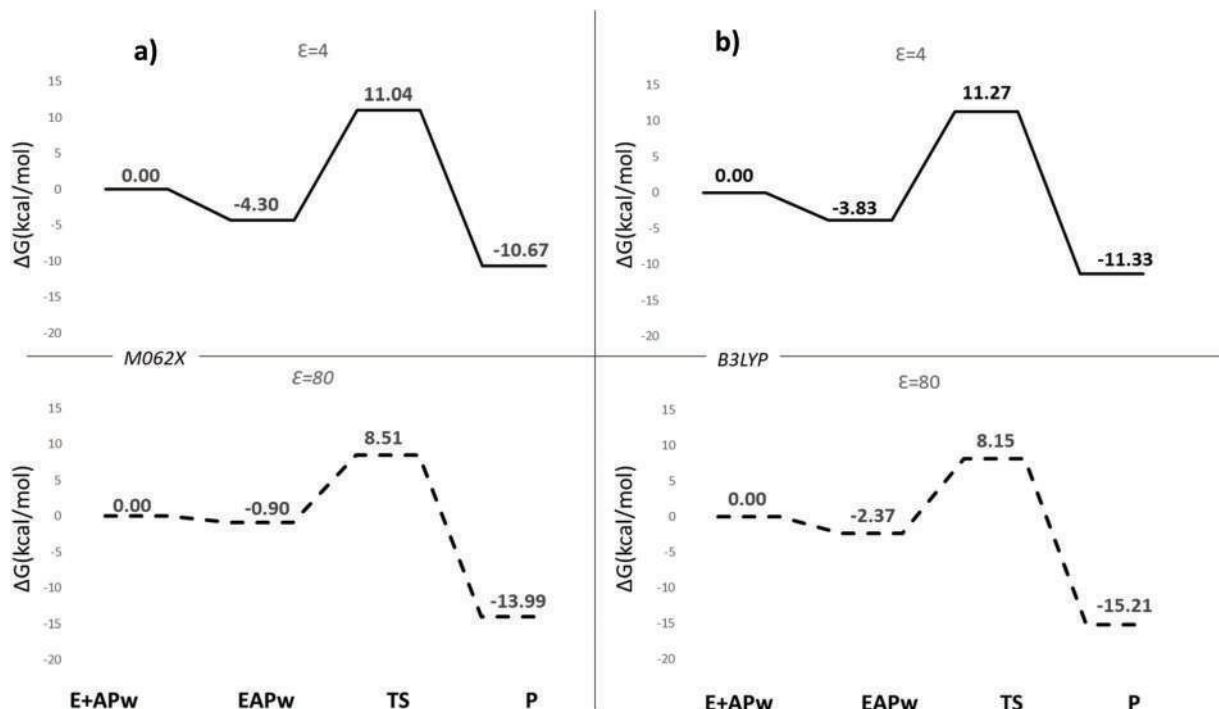


Fig. 4 M062x-D3/6-311+G(2d,2p)|SDD//M062x/6-31+G(d,p)|SDD (a) and B3LYP-D3/6-311+G(2d,2p)|SDD//M062x/6-31+G(d,p)|SDD (b) free energy profiles of the RNase A metalation process promoted by APw in protein ( $\epsilon = 4$ ) and water ( $\epsilon = 80$ ) environments.

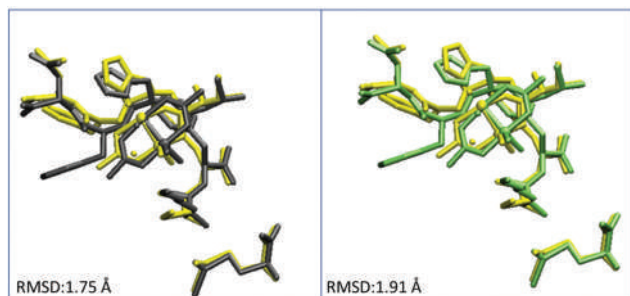


Fig. 5 Superposition of the M062x/6-31+G(d)|SDD (gray) and B3LYP-D3/6-31+G(d)|SDD (green) optimized structures of the final covalent platinated product (**P**) with the corresponding crystallographic structure (yellow).

process indicates that the reverse reaction is still possible but much slower for the higher barriers (21.71 and 22.50 kcal mol<sup>-1</sup> at the M062x level and 22.60 and 23.36 kcal mol<sup>-1</sup> at the B3LYP-D3 level in protein and water environments, respectively).

Comparison with the experimental results,<sup>10</sup> (see Fig. 5) possible only for the **P** species, shows a good agreement with the calculated geometry with both the employed functionals as indicated by the superposition of the optimized structure of **P**.

To better evaluate the nature of the interactions present inside the catalytic pocket during the process, in Fig. 6, we have reported the density of isosurfaces, arising from Non-Covalent Interactions (NCI) analysis, indicating the contributions of the residues retained in the model. The red isosur-

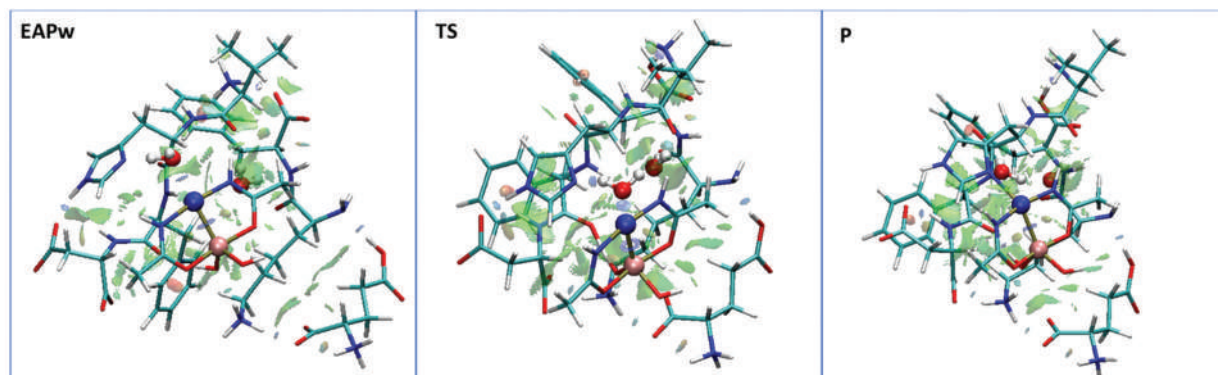


Fig. 6 Nonbonding interaction plot calculated for the stationary points at the M062x/6-311+G(2d,2p)|SDD//M062x/6-31+G(d,p)|SDD level.



faces account for the repulsive interactions related to the center of the  $\pi$  systems of Phe8 and Phe120 residues. Further information arises from the green regions indicative of the dispersion contributions (van der Waals forces) characterizing the binding cavity of the adduct that increases during the platination process (**EAPw**  $\rightarrow$  **P**). The blue region around Pt reveals significant attractive interactions (H-bond) that are absent in the **EAPw** but appear in the **TS** structure where **W1** engages N $\delta$  of His119 and the **W2** molecule forms a H-bond.

## Conclusions

Drug discovery remains a great challenge for scientific and economic reasons. In particular, the development of new anticancer agents is a long-term process, which involves multiple steps and can not always exclude the occurrence of side effects. The preparation of hybrid molecules with two or more structural domains acting as two distinct pharmacophores represents a more fast and efficient pathway. This is particularly the case of arsenoplatin-1 (**AP1**), [Pt( $\mu$ -NHC(CH<sub>3</sub>)O)<sub>2</sub>ClAs(OH)<sub>2</sub>], the object of the theoretical investigation presented herein. It is a new anticancer agent that is more active than the parent drugs As<sub>2</sub>O<sub>3</sub> and cisplatin in most of the cancer cell lines tested. The metalation process of RNase A by **AP1** has been elucidated by DFT calculations using a larger cluster model able to represent the environment of the binding site of **AP1** quite realistically. The proposed one-step mechanism of the RNase A inhibition by **AP1**, examined in both protein and water environments by the solvation continuum model, shows low barriers compatible with a very fast process. The occurrence in aqueous medium is more favorable from a kinetic point of view with both B3LYP and M062x functionals and compatible with the protein's active site present on the surface area accessible to solvents. The structure of the metalated protein product (**P**) matches very well with the corresponding crystallographic one, confirming that the binding site is well reproduced by the used cluster model.

## Conflicts of interest

The authors declare that the research was conducted in the absence of any potential conflict of interest.

## Acknowledgements

The Università degli Studi della Calabria-Dipartimento di Chimica e Tecnologie Chimiche (CTC) is gratefully acknowledged.

## References

- 1 H. Chen, S. Pazieni, N. Krett, R. Ahn, J. Penner-Hahn, S. Rosen and T. O'Halloran, Coencapsulation of Arsenic- and Platinum-Based Drugs for Targeted Cancer Treatment, *Angew. Chem.*, 2009, **121**, 9459–9463.
- 2 D. U. Miodragović, J. A. Quentzel, J. W. Kurutz, C. L. Stern, R. W. Ahn, I. Kandela, A. Mazar and T. V. O'Halloran, Robust Structure and Reactivity of Aqueous Arsenous Acid-Platinum(II) Anticancer Complexes, *Angew. Chem., Int. Ed.*, 2013, **52**, 10749–10752.
- 3 C. Zheng, S. Lam, Y. Li, B. M. Fong, J. C. Mak and J. C. Ho, Combination of Arsenic Trioxide and Chemotherapy in Small Cell Lung Cancer, *Lung Cancer*, 2013, **82**, 222–230.
- 4 T. Nakaoka, A. Ota, T. Ono, S. Karnan, H. Konishi, A. Furuhashi, Y. Ohmura, Y. Yamada, Y. Hosokawa and Y. Kazaoka, Combined Arsenic Trioxide-Cisplatin Treatment Enhances Apoptosis in Oral Squamous Cell Carcinoma Cells, *Cell. Oncol.*, 2014, **37**, 119–129.
- 5 N. Zhang, Z.-M. Wu, E. McGowan, J. Shi, Z.-B. Hong, C.-W. Ding, P. Xia and W. Di, Arsenic Trioxide and Cisplatin Synergism Increase Cytotoxicity in Human Ovarian Cancer Cells: Therapeutic Potential for Ovarian Cancer, *Cancer Sci.*, 2009, **100**, 2459–2464.
- 6 T. Marino, A. Parise and N. Russo, The Role of Arsenic in the Hydrolysis and DNA Metalation Processes in an Arsenous Acid-Platinum(II) Anticancer Complex, *Phys. Chem. Chem. Phys.*, 2017, **19**, 1328–1334.
- 7 A. Merlino, T. Marzo and L. Messori, Protein Metalation by Anticancer Metallodrugs: A Joint ESI MS and XRD Investigative Strategy, *Chem. – Eur. J.*, 2017, **23**, 6942–6947.
- 8 L. Messori and A. Merlino, Protein Metalation by Metal-Based Drugs: X-Ray Crystallography and Mass Spectrometry Studies, *Chem. Commun.*, 2017, **53**, 11622–11633.
- 9 A. Casini, G. Mastrobuoni, C. Temperini, C. Gabbiani, S. Francese, G. Moneti, C. T. Supuran, A. Scozzafava and L. Messori, ESI Mass Spectrometry and X-Ray Diffraction Studies of Adducts between Anticancer Platinum Drugs and Hen Egg White Lysozyme, *Chem. Commun.*, 2007, 156–158.
- 10 D. Miodragović, A. Merlino, E. P. Swindell, A. Bogachkov, R. W. Ahn, S. Abuhadba, G. Ferraro, T. Marzo, A. P. Mazar, L. Messori and T. V. O'Halloran, Arsenoplatin-1 Is a Dual Pharmacophore Anticancer Agent, *J. Am. Chem. Soc.*, 2019, **141**, 6453–6457.
- 11 M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. Hratchian, A. Izmaylov, J. Bloino, G. Zheng, J. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. Peralta, F. Ogliaro, M. Bearpark, J. Heyd, E. Brothers, K. Kudin, V. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. Burant, S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. Millam, M. Klene, J. Knox, J. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. Stratmann, O. Yazyev, A. Austin, R. Cammi, C. Pomelli, J. Ochterski, R. Martin, K. Morokuma, V. Zakrzewski, G. Voth, P. Salvador, J. Dannenberg, S. Dapprich, A. Daniels, Ö. Farkas,

- J. Foresman, J. Ortiz, J. Cioslowski and D. Fox, *Gaussian 09, Revision D.01*, Gaussian, Inc., Wallingford CT.
- 12 A. D. Becke, Density-functional Thermochemistry. III. The Role of Exact Exchange, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
  - 13 C. Lee, W. Yang and R. G. Parr, Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.
  - 14 Y. Zhao and D. G. Truhlar, The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
  - 15 D. Andrae, U. Hausermann, M. Dolg, H. Stoll and H. Preuss, Energy-Adjusted *ab Initio* Pseudopotentials for the Second and Third Row Transition Elements, *Theor. Chim. Acta*, 1990, **77**, 123–141.
  - 16 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu, *J. Chem. Phys.*, 2010, **132**, 154104.
  - 17 T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery, K. Morokuma and M. J. Frisch, Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM, *J. Chem. Theory Comput.*, 2006, **2**, 815–826.
  - 18 S. Grimme, Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory, *Chem. – Eur. J.*, 2012, **18**, 9955–9964.
  - 19 R. P. P. Neves, P. A. Fernandes and M. J. Ramos, Mechanistic Insights on the Reduction of Glutathione Disulfide by Protein Disulfide Isomerase, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, E4724–E4733.
  - 20 F. E. Medina, R. P. P. Neves, M. J. Ramos and P. A. Fernandes, QM/MM Study of the Reaction Mechanism of the Dehydratase Domain from Mammalian Fatty Acid Synthase, *ACS Catal.*, 2018, **8**, 10267–10278.
  - 21 F. E. Medina, M. J. Ramos and P. A. Fernandes, Complexities of the Reaction Mechanisms of CC Double Bond Reduction in Mammalian Fatty Acid Synthase Studied with Quantum Mechanics/Molecular Mechanics Calculations, *ACS Catal.*, 2019, **9**, 11404–11412.
  - 22 M. Prejanò, T. Marino and N. Russo, On the Inhibition Mechanism of Glutathione Transferase P1 by Piperlongumine. Insight From Theory, *Front. Chem.*, 2018, **6**, 606.
  - 23 M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao and P. E. M. Siegbahn, Quantum Chemical Studies of Mechanisms for Metalloenzymes, *Chem. Rev.*, 2014, **114**, 3601–3658.
  - 24 A. J. M. Ribeiro, M. E. Alberto, M. J. Ramos, P. A. Fernandes and N. Russo, The Catalytic Mechanism of Protein Phosphatase 5 Established by DFT Calculations, *Chem. – Eur. J.*, 2013, **19**, 14081–14089.
  - 25 P. Piazzetta, T. Marino and N. Russo, Insight into the Promiscuous Activity of Human Carbonic Anhydrase against the Cyanic Acid Substrate from a Combined QM and QM/MM Investigation, *Phys. Chem. Chem. Phys.*, 2014, **16**, 16671.
  - 26 A. J. M. Ribeiro, D. Santos-Martins, N. Russo, M. J. Ramos and P. A. Fernandes, Enzymatic Flexibility and Reaction Rate: A QM/MM Study of HIV-1 Protease, *ACS Catal.*, 2015, **5**, 5617–5626.
  - 27 H. Chen, S. Pazicni, N. Krett, R. Ahn, J. Penner-Hahn, S. Rosen and T. O'Halloran, Coencapsulation of Arsenic- and Platinum-Based Drugs for Targeted Cancer Treatment, *Angew. Chem.*, 2009, **121**, 9459–9463.
  - 28 R.-Z. Liao, F. Himo, J.-G. Yu and R.-Z. Liu, Dipeptide Hydrolysis by the Dinuclear Zinc Enzyme Human Renal Dipeptidase: Mechanistic Insights from DFT Calculations, *J. Inorg. Biochem.*, 2010, **104**, 37–46.
  - 29 C. Gonzalez and H. B. Schlegel, An Improved Algorithm for Reaction Path Following, *J. Chem. Phys.*, 1989, **90**, 2154–2161.
  - 30 C. Gonzalez and H. Bernhard, Schlegel, Reaction Path Following in Mass-Weighted Internal Coordinates, *J. Phys. Chem.*, 1990, **94**, 5523–5527.
  - 31 M. Cossi, N. Rega, G. Scalmani and V. Barone, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model, *J. Comput. Chem.*, 2003, **24**, 669–681.
  - 32 V. Barone and M. Cossi, Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
  - 33 J. Yan and S. Chen, How To Produce Methane Precursor in the Upper Ocean by An Untypical Non-Heme Fe-Dependent Methylphosphonate Synthase?, *ChemPhysChem*, 2020, **21**, 385–396.
  - 34 M. Prejanò, I. Romeo, L. Sgrizzi, N. Russo and T. Marino, Why Hydroxy-Proline Improves the Catalytic Power of the Peptidoglycan *N*-Deacetylase Enzyme: Insight from Theory, *Phys. Chem. Chem. Phys.*, 2019, **21**, 23338–23345.
  - 35 M. Prejanò, N. Russo and T. Marino, How Lanthanide Ions Affect the Addition–Elimination Step of Methanol Dehydrogenases, *Chem. – Eur. J.*, 2020, **26**, 11334–11339.
  - 36 E. D. Glendening, A. E. Reed, J. E. Carpenter and F. Weinhold, *NBO, Version 3.1*, Theoretical Chemistry Institute, University of Wisconsin, Madison, 1995.
  - 37 P. E. M. Siegbahn and F. Himo, Recent Developments of the Quantum Chemical Cluster Approach for Modeling Enzyme Reactions, *JBIC, J. Biol. Inorg. Chem.*, 2009, **14**, 643–651.
  - 38 V. Srivastava, In Situ Generation of Ru Nanoparticles to Catalyze CO<sub>2</sub> Hydrogenation to Formic Acid, *Catal. Lett.*, 2014, **144**, 1745–1750.
  - 39 F. J. Pérez-Alonso, M. L. Granados, M. Ojeda, T. Herranz, S. Rojas, P. Terreros, J. L. G. Fierro, M. Gracia and J. R. Gancedo, Relevance in the Fischer–Tropsch Synthesis of the Formation of Fe–O–Ce Interactions on Iron–Cerium

- Mixed Oxide Systems, *J. Phys. Chem. B*, 2006, **110**, 23870–23880.
- 40 I. Tolbatov, C. Coletti, A. Marrone and N. Re, Reactivity of arsenoplatin complex versus water and thiocyanate: a DFT benchmark study, *Theor. Chem. Acc.*, 2020, **139**, 184.
- 41 A. T. Pereira, A. J. M. Ribeiro, P. A. Fernandes and M. J. Ramos, Benchmarking of density functionals for the kinetics and thermodynamics of the hydrolysis of glycosidic bonds catalyzed by glycosidases, *Int. J. Quantum Chem.*, 2017, **117**, e25409.





**Paper III**  
The Se-S Bond Formation in the Covalent Inhibition Mechanism of SARS-CoV-2 Main  
Protease by Ebselen-like Inhibitors: A Computational Study





## Article

# The Se–S Bond Formation in the Covalent Inhibition Mechanism of SARS-CoV-2 Main Protease by Ebselen-like Inhibitors: A Computational Study

Angela Parise <sup>1,2</sup>, Isabella Romeo <sup>1</sup> , Nino Russo <sup>1</sup> and Tiziana Marino <sup>1,\*</sup>

<sup>1</sup> Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy; angela.parise@unical.it (A.P.); isabella.romeo@unical.it (I.R.); nino.russo@unical.it (N.R.)

<sup>2</sup> Institut de Chimie Physique UMR8000, Université Paris-Saclay, CNRS, 91405 Orsay, France

\* Correspondence: tiziana.marino65@unical.it

**Abstract:** The inhibition mechanism of the main protease ( $M^{Pro}$ ) of SARS-CoV-2 by ebselen (EBS) and its analog with a hydroxyl group at position 2 of the benzisoselenazol-3(2H)-one ring (EBS-OH) was studied by using a density functional level of theory. Preliminary molecular dynamics simulations on the apo form of  $M^{Pro}$  were performed taking into account both the hydrogen donor and acceptor natures of the  $N\delta$  and  $N\epsilon$  of His41, a member of the catalytic dyad. The potential energy surfaces for the formation of the Se–S covalent bond mediated by EBS and EBS-OH on  $M^{Pro}$  are discussed in detail. The EBS-OH shows a distinctive behavior with respect to EBS in the formation of the noncovalent complex. Due to the presence of canonical H-bonds and noncanonical ones involving less electronegative atoms, such as sulfur and selenium, the influence on the energy barriers and reaction energy of the Minnesota hybrid meta-GGA functionals M06, M06-2X and M08HX, and the more recent range-separated hybrid functional wB97X were also considered. The knowledge of the inhibition mechanism of  $M^{Pro}$  by the small protease inhibitors EBS or EBS-OH can enlarge the possibilities for designing more potent and selective inhibitor-based drugs to be used in combination with other antiviral therapies.

**Keywords:** SARS-CoV-2 main protease; DFT; inhibition mechanism; Se–S covalent bond; potential energy surface



**Citation:** Parise, A.; Romeo, I.; Russo, N.; Marino, T. The Se–S Bond Formation in the Covalent Inhibition Mechanism of SARS-CoV-2 Main Protease by Ebselen-like Inhibitors: A Computational Study. *Int. J. Mol. Sci.* **2021**, *22*, 9792. <https://doi.org/10.3390/ijms22189792>

Academic Editor: József Tózsér

Received: 12 August 2021

Accepted: 7 September 2021

Published: 10 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



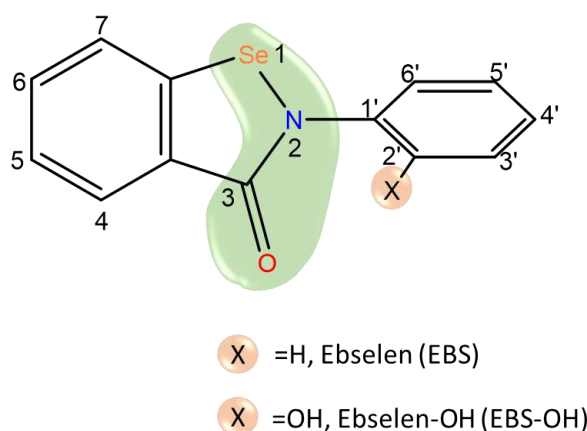
**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Since early 2020 the whole world has been trying to face the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1,2]. With a relatively long incubation period along with symptoms characterized by different levels of severity, the disease is still affecting billions of people and spreading in an unrelenting fashion causing widespread health, social, and economic disruption. Despite originating from the same *Coronaviridae* family as the Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) the shared sequence similarity is ~80% [2]. Viruses belonging to the genus Coronavirus are zoonotic and characterized by positive-stranded RNA encapsulated by a membrane envelope of 300–400 nm [3]. The morphology of virions is similar to a crown due to the mushroom-shaped proteins called spike glycoproteins (S proteins) responsible for the host tropism [4]. Human infections caused by the SARS coronavirus are distinguished by the interaction between the S-protein and the human angiotensin-converting enzyme 2 (ACE2); highly expressed on epithelial cells of the respiratory tract [5]. Following host cell entry, the viral single-stranded RNA is released for replication and translation of the virus polyproteins that are processed by two cysteine proteases (CPs), papain-like protease ( $PL^{Pro}$ ) and a main protease ( $M^{Pro}$ ) also known as 3-Chymotrypsin-like protease ( $3CL^{Pro}$ ) [6,7].  $M^{Pro}$  acts in the immune regulation and cleavage of the polyprotein at eleven

different sites generating nonstructural proteins, such as RNA polymerase, endo- and exoribonuclease, which are relevant in the replication process of the virus [8]. Considering that M<sup>Pro</sup> is only found in the virus rather than in the host cell, this protein represents an interesting target for the development of new promising anticoronavirus therapeutic agents [6,7]. Structurally, SARS-CoV-2 M<sup>Pro</sup> forms a homodimer (protomer A and B) and each protomer consists of three domains: domain I (residues 8–101), domain II (residues 102–184), and domain III (residues 201–303) which are connected by a loop region (residues 185–200) [6,7]. In M<sup>Pro</sup>, the catalytic dyad consisting of His41 and Cys145 is located in the cleft between domains I and II [9,10]. The main aim of this protein is cleaving the polyprotein pp1a and pp1ab translated from the viral RNA at 16 different positions to generate important structural proteins in addition to nanostructural proteins (NSPs) which are involved in arresting the process of viral assembly in the replication cycle [11]. The presence of a reactive sulphur in the active site of cysteine proteases provides a target for the design of many covalent and noncovalent inhibitors [11–15]. Among the in vitro investigated covalent inhibitors, it was observed that ebselen [(2-phenyl-1,2-benzisoxselenazol-3(2H)-one), (EBS) shows stronger inhibition against the SARS-CoV-2 virus than N3, the known Michael acceptor inhibitor [15–19]. EBS is a low-molecular-weight organoselenium drug having a pleiotropic mode of action and because of its very low toxicity has encountered no barriers for use in humans [19–21]. In fact, in the pre-COVID-19 era, EBS was known as a mimic of glutathione peroxidase and inhibits several enzymes involved in inflammatory processes, such as lipoxygenases, NO synthases and NADPH, providing it with remarkable anti-inflammatory, antiatherosclerotic and cytoprotective properties [22,23]. EBS is known to act as potent inhibitor of proteins implicated in the biosynthesis of the mycobacterial cell wall [20,24] by covalently modifying a noncatalytic cysteine through the reversible oxidation of the  $\gamma$ -sulfur of the amino acid by the EBS selenium, while in the case of M<sup>Pro</sup>, the cysteine target of ebselen is crucial for its protease activity [25]. In this system there is evidence arising from experimental study that the cysteine residue involved is present in the catalytic site (Cys145) [15]. Furthermore, EBS has also shown preclinical efficacy for cisplatin-induced ototoxicity [26,27] and has been considered for the treatment of bipolar disorders and hearing loss [28]. EBS inhibition properties have been evaluated in different clinical trials, exhibiting its safety in humans due to its extremely low cytotoxicity [29,30] and therefore it provides the promise of drug repurposing. On the basis of these preclinical studies the reaction of ebselen was reported with cysteine residues from completely unrelated proteins including the C-terminal domain of the HIV-1 capsid, Mycobacterium tuberculosis transpeptidase LdtMt2, glutamate dehydrogenase, Clostridium difficile toxins, Mycobacterium tuberculosis antigen 85C enzyme and many others [20,31–36]. Ebselen and some of its synthesized analogues were found to inhibit both SARS-CoV-2 M<sup>Pro</sup> and PL<sup>Pro</sup> [37,38]. Repurposing of known drugs can provide an accelerated path for approval and a likely option to address the current COVID-19 pandemic. In particular, EBS has displayed inhibition against the M<sup>Pro</sup> of SARS-CoV-2 virus with the concentration required to produce 50% of the maximum possible effect and indicating antiviral activity in cells, the EC<sub>50</sub>, equal to 4.67  $\mu$ M; superior to that of N3 (16.77  $\mu$ M) [15]. Furthermore, the improvement of the inhibitory potency of the ebselen derivative obtained by the hydroxyl group in the position ortho of the N-phenyl ring (EBS-OH of Scheme 1) on the other Achilles' heel SARS-CoV-2 PL<sup>Pro</sup>, induced us to investigate the M<sup>Pro</sup> inhibition mechanism by both EBS and EBS-OH inhibitors [21,39]. The behavior of the heavier chalcogens, S and Se, in biological systems has only recently gained more attention, in particular Se has been shown to form an Se–S intermediate in selenoprotein reductase and formate dehydrogenase. They play a central role in the enzyme's activation [40,41] and also in deiodinase biomimetics [42]. So far the SARS-CoV-2 M<sup>Pro</sup> inhibition mechanism has been studied by using Michael acceptor and peptidomimetic inhibitors at the QM-MM level of theory [43–47]. So, the mechanistic understanding of the Se–S covalent bond formation promoted by inhibitors, such as EBS and EBS-OH in the present study, could contribute to

enlarge the molecular inhibition mechanism of SARS-CoV-2  $M^{pro}$  and stimulate the design of other similar drugs.



**Scheme 1.** Sketch of ebselen, X group in position ortho of the phenyl ring can be H (EBS) or OH (EBS-OH). Ebselen warhead is evidenced in green.

Moreover, stimulated by the most recent literature regarding the inhibition mechanism of  $M^{pro}$  enzyme [43,45–47] we took into account the protonation state of the catalytic histidine (His41) [43,48,49]. Therefore, we have also performed computational simulations by using molecular docking and classical molecular dynamics (cMD) devoted to accurately model the structure and dynamics of  $M^{pro}$  in both the  $N\delta$  and  $N\epsilon$  protonated states of His41 to better evaluate their eventual influence on the covalent inhibition process by EBS. For clarity,  $M^{pro}$  in the  $N\delta$  form will in the whole manuscript be named  $M^{pro}$ -HID and  $M^{pro}$  in the  $N\epsilon$  form will be  $M^{pro}$ -HIE. To shed light into the inhibition mechanism of EBS and EBS-OH, the potential energy surfaces (PES) for the formation of the covalent complex between SARS-CoV-2  $M^{pro}$  and the two considered inhibitors have been calculated and analysed, considering the effects of the dielectric value,  $\epsilon = 80$  and  $\epsilon = 4$ , for the water and protein environments, respectively. In the case of EBS-OH only the inhibition mechanism by  $N\delta$ -His41 was performed. The calculated PES for the two inhibitors can be used to determine whether, and to what extent, a covalent inhibitor is reversible or not. This knowledge is important in order to provide in the near future, additional leads for covalent inhibitors obtained by incorporating small-molecules, such as EBS and EBS-OH, into hybrid molecules which also have few side effects. In addition, the ability of other DFT functionals ( $\omega$ B97X, M06, M06-2X and M08-HX) as single points on the B3LYP-D3 optimized geometries were tested to obtain more accurate data for the energy barriers and reaction energy of the inhibition mechanism driven by the sulfur–selenium covalent bond formation.

## 2. Results

In an effort to better understand the molecular mechanisms involved in the inhibition of  $M^{pro}$ , many works have focused on classical Molecular Dynamics simulations with [12,14,25,48–50] and without [11,51] inhibitors. In the present investigation cMD simulations performed on apo  $M^{pro}$ -HID and  $M^{pro}$ -HIE represent a preliminary step useful to obtain comparative structural and dynamic properties to deeply describe EBS and EBS-OH inhibition mechanisms.

### 2.1. MD Analysis of the Apo form of $M^{pro}$ -HID and $M^{pro}$ -HIE

SARS-CoV-2  $M^{pro}$  has been well characterized by crystallography. So far about 1000 entries of X-ray crystal structures for SARS-CoV-2  $M^{pro}$  are present on the Protein Data Bank in an apo form or a complexed one [15,51–55]. To better evaluate the effects on the structural properties in the catalytic pocket, the starting crystal structure with the PDB code

6W63 [56] was used for cMD simulations considering the apo form and the complexed one with EBS in both the  $N\delta$  and  $N\epsilon$  protonation states of the His41. The Cys145 residue was considered in the neutral form. As previously suggested [49], these investigations are helpful to monitor what happens during the simulation time in the His41-Cys145 catalytic dyad region, allowing us to better rationalize eventual changes in the shape of the inhibitor catalytic site and their possible influence on the explored inhibition mechanism of EBS [48,57].

To compare the dynamic behavior of  $M^{Pro}$ - $N\delta$  and  $M^{Pro}$ - $N\epsilon$ , several properties were taken into account: root mean square deviation (RMSD); RMSD-based clustering; root mean square fluctuation (RMSF), solvent accessible surface area (SASA), radial distribution function (RDF), volume of the binding pocket, salt bridges and the H-bond between residues. Results are reported in Figures S1–S9.

The analysis of the calculated RMSD for the residue pair Cys145-His41 in both the  $N\delta$  and  $N\epsilon$  forms (shown in Figure S3) allows us to observe that the spatial orientation of the dyad remains constant during the simulation time in the case of His41- $N\delta$ , while a dissimilar trend assumes the RMSD value of His41- $N\epsilon$ . The major fluctuations observed in RMSD for His41- $N\epsilon$  could suggest different conformations of the Cys145 side chain also previously found in the ortholog protease of SARS-CoV [46,58].

A different behavior may also be noticeable in the trend of RMSF (see Figure S4) where it is evident that mainly domain I (residues 1–100) and III (residues 200–300) suffer a major rearrangement revealing higher mobility.

From the superposition of the most representative structure derived from RMSD-based clustering (Figure S2 and Table S4) of the MD trajectory of SARS-CoV-2  $M^{Pro}$ -HID and HIE (Figure 1), it is possible to see no appreciable difference in the secondary structure except for the chain related to residue His41, that in  $M^{Pro}$ -HID is organized in a longer alpha-helix including Arg40-His41-Val42-Ile43-Cys44 versus Arg40-His41-Val42 present in  $M^{Pro}$ -HIE. In addition, the insets of Figure 1A emphasize the peculiar feature in the catalytic dyad: the altered rotational state of His41 along with the different conformation of the side chain of Cys145. In fact, the  $C=O-C_{\alpha}-C_{\beta}-S_{\gamma}$  torsional angle adopts a trans-like conformation ( $-149.15^{\circ}$ ) in  $M^{Pro}$ -HID and a gauche-like one ( $50.69^{\circ}$ ) in  $M^{Pro}$ -HIE.

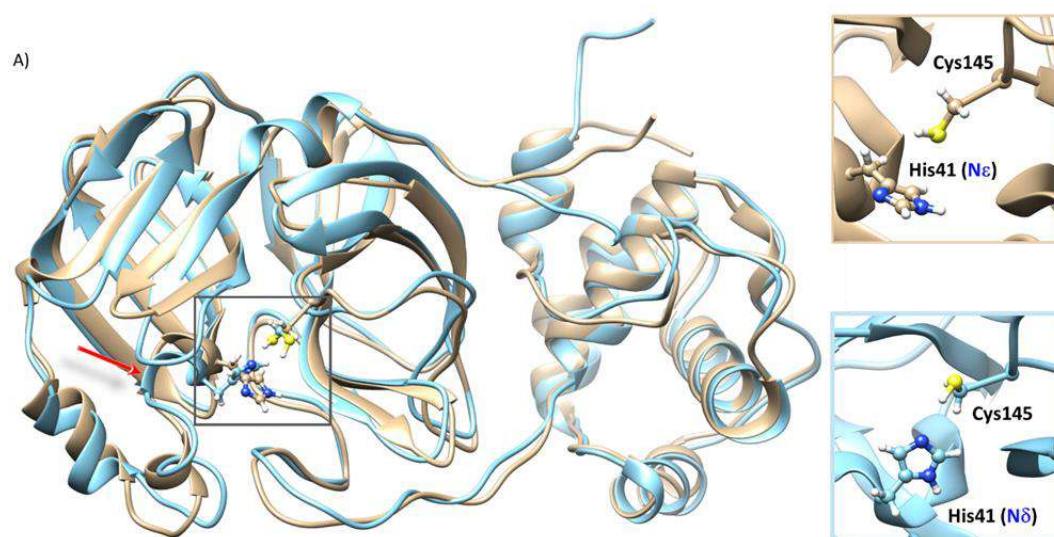
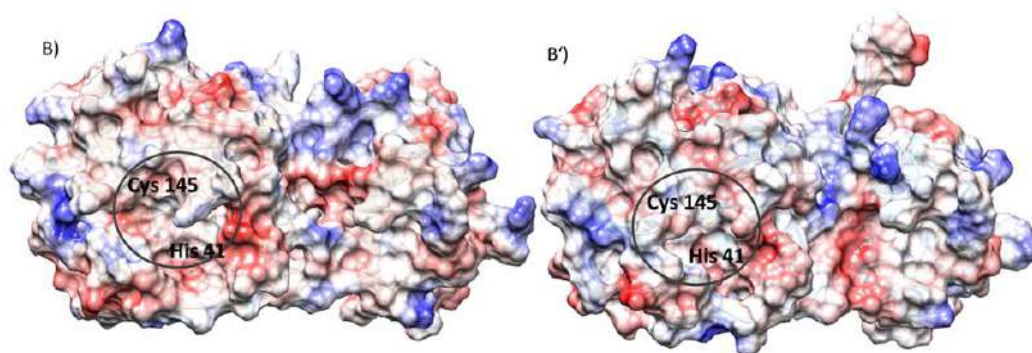


Figure 1. Cont.





**Figure 1.** (A) Alignment of the most representative structure of the SARS-CoV-2 M<sup>Pro</sup>-HID (cyan) and SARS-CoV-2 M<sup>Pro</sup>-HIE (gold), as obtained from a cluster analysis along the molecular dynamics trajectories. The insets evidence the structural organization inside the catalytic dyad (His41-Cys145). (B,B') Electrostatic potential surface (APBS) in the SARS-CoV-2 M<sup>Pro</sup>-HID (B) and SARS-CoV-2 M<sup>Pro</sup> HIE (B').

Furthermore, the comparison of the electrostatic potential, shown in Figure 1B,B', calculated for both protonation states of His41 by solving the Poisson–Boltzmann equation as implemented in the APBS code [59], reveals a more marked distribution of negative regions on the surface in proximity to the catalytic site of the SARS-CoV-2 M<sup>Pro</sup> HID. Moreover, the catalytic pocket volume changes dependently on M<sup>Pro</sup>-HID or M<sup>Pro</sup>-HIE (see Figure S5). Similar active site flexibility has been previously highlighted by Pavlova et al. [48].

## 2.2. Covalent Inhibition

Jin et al. [15], applying a screening strategy on over 10,000 inhibitors of the SARS-CoV-2 virus main protease identified EBS as a good candidate, revealing through tandem MS/MS analysis, its ability to covalently bind to the cysteine 145 residue of M<sup>Pro</sup>. Due to the observed stronger inhibition, they did not exclude the possibility that it can also act as a noncovalent inhibitor [15].

As reported in Scheme 2 the inhibition usually takes place in two steps. The compound must first bind noncovalently to the target protein, placing its moderately reactive electrophilic selenyl amide moiety, the warhead in ebselen (as circled in Scheme 1), close to the nucleophile (Cys145) of M<sup>Pro</sup> and giving rise to the enzyme–inhibitor complex (EI), in which the binding free energy ( $\Delta G_{\text{bind}}$ ) depends only on noncovalent interactions and is related to the inhibition constant  $K_i$ . In the second stage a chemical reaction transforms the EI complex generating the final covalent complex (E-I). In the case of EBS this mechanism is proposed by MS/MS study [15]. Irrespective of the exothermicity, if high reverse reaction barriers are observed the bond formation is effectively irreversible, so  $k_{-2}$  will be zero (see Scheme 2).



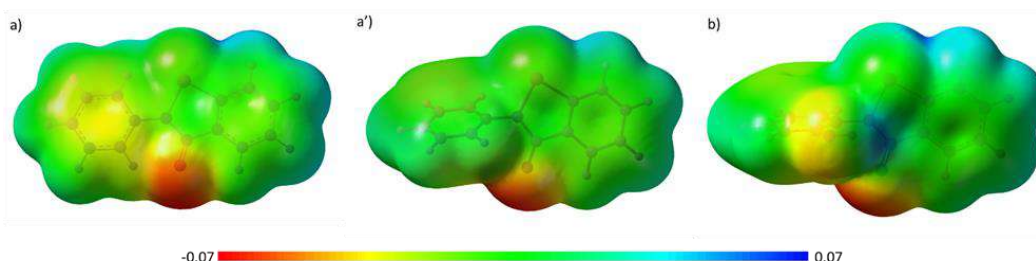
**Scheme 2.** General scheme of covalent interaction between a small molecule and its target.

### 2.2.1. Ebselen (EBS) and Its Derivative (EBS-OH)

The knowledge of the inhibition mechanism at an atomistic level represents a crucial step in the pathway of drug design since it allows us to characterize the reaction path, including characterization of the transition state species which are undetectable at an experimental level. The investigation also extended to EBS-OH which may prove helpful to better understand if this kind of ebselen derivative can improve the selectivity action, a feature actually not present in ebselen [15–19,21]. As above mentioned, the additional

hydroxyl group in the position ortho of the phenyl ring resulted in an increase of the inhibitory potency of ebselen by one order of magnitude [21].

The electrostatic potential maps (MEPs) of the two molecules, shown in Figure 2, enable us to visualize the charge distributions and charge related properties of the two molecules.



**Figure 2.** Molecular electrostatic potential (MEP) of: (a) planar EBS; (a') nonplanar EBS and (b) EBS-OH geometries optimized at B3LYP/6-31+G level of theory.

It is possible to remark that the presence of an OH moiety on the position ortho of the N-phenyl ring of EBS introduces a redistribution of the charge that can be fruitful in the course of the inhibition mechanism. The NBO charge values collected in Table S6 underline a major variation on the selenyl amide moiety. In particular, the nitrogen ( $N_2$ ) becomes more negative ( $-0.661$  e vs.  $-0.602$  e in EBS) and the oxygen ( $O_3$ ) decreases its charge ( $-0.597$  e vs.  $-0.618$  e in EBS). Furthermore, the charges on  $C3'$ ,  $C4'$  and  $C5'$  reflect the effect of OH as an activating group on  $C5'$  (ortho) and  $C3'$  (para), while as a deactivating one on  $C4'$  (meta) with respect to EBS.

Inside the catalytic pocket, EBS tends to assume a folded-like conformation with almost isoenergetic results ( $\sim 0.32$  kcal mol $^{-1}$ ) to the totally planar one (see Figure 2). It is possible to note that the charge distribution suffers a small modification if compared with the extended conformation, but is still different from that of EBS-OH, evidencing again the role played by its hydroxyl group on the N-phenyl ring.

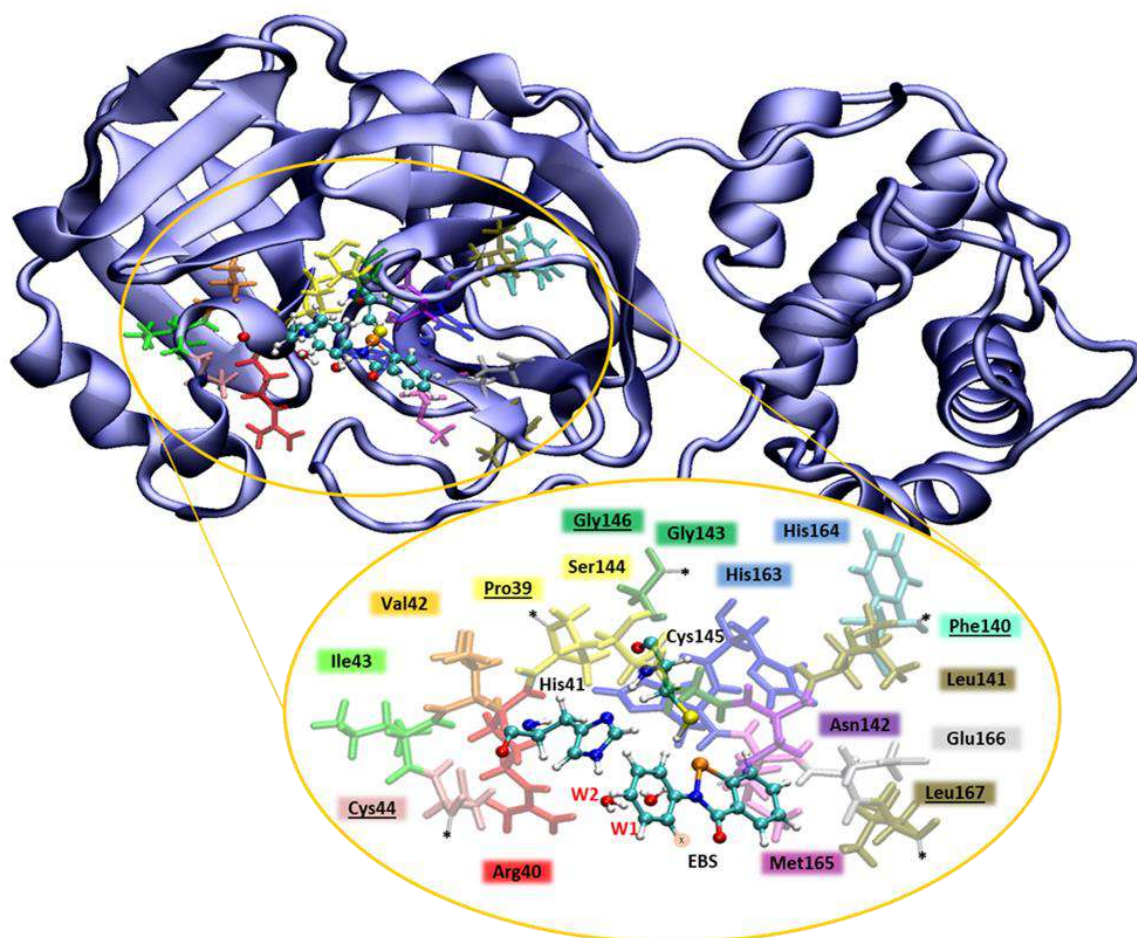
### 2.2.2. Determination of the Inhibitor-M<sup>Pro</sup> Complex

The most representative structure of the M<sup>Pro</sup>-HID and M<sup>Pro</sup>-HIE, obtained from the clustering MDs trajectory (Figure S2 and Table S4), was used in the next docking simulations with EBS and EBS-OH. Details of the procedure used are given in the Electronic Supplementary Materials.

The best docked poses M<sup>Pro</sup>-HID-EBS, M<sup>Pro</sup>-HIE-EBS and M<sup>Pro</sup>-HID-EBS-OH (shown in Figure S11) were employed, after solvation and MM minimization, as the starting point for building the active site model.

**Active site model:** The QM cluster model of the active site was derived following a well consolidated procedure [60–63]. The model includes: Pro39, Arg40, Val42, Ile43, Cys44, Phe140, Leu141, Asn142, Gly143, Ser144, His163, Met165, Glu166, Leu167, the catalytic dyad His41/Cys145, EBS, W1 and W2 (see Figure 3). Following the indications from the literature [46] and from our cMD results (Figure S8 and Tables S2 and S3), the Asp187 residue was not included in the model since its involvement in a salt bridge with Arg40 precluded its participation in the inhibition mechanism. As usually requested by the QM cluster model [64–66], the amino acid residues were truncated as shown in Figure 3, and the coordinates of the related selected atoms were fixed (indicated by stars) during geometry optimizations to prevent unrealistic movements of residues at the active site. In spite of this, the size of the model used in the present study was large enough to grant the flexibility required for the active site groups during the chemical events of the inhibition process. The model thus consists of 312 or 313 atoms including EBS or EBS-OH, respectively. In all the studied systems, the overall charge of the model is zero.





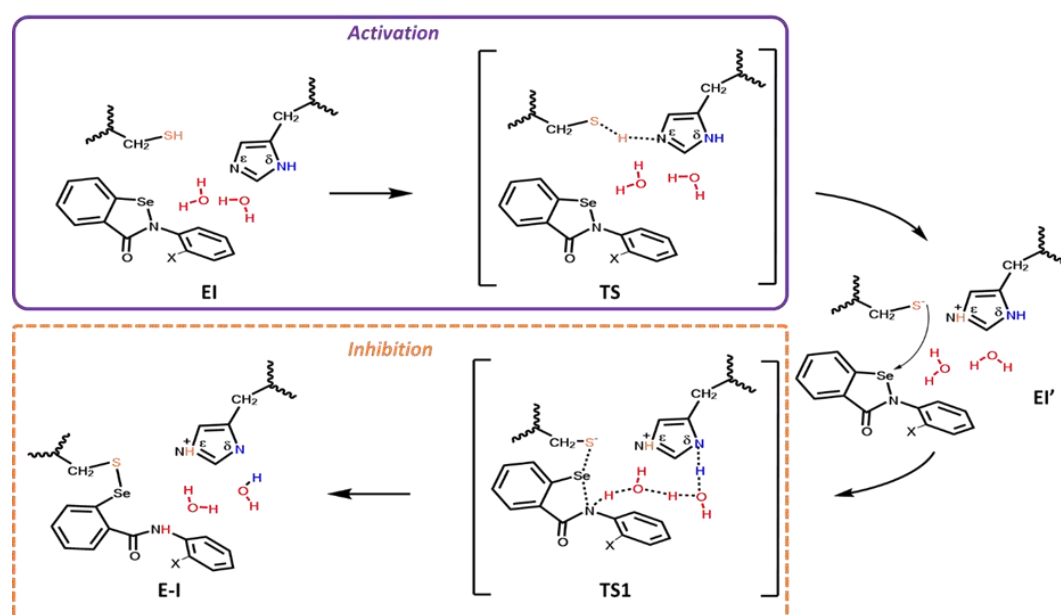
**Figure 3.**  $M^{Pro}$  cartoon representation. In the yellow circle the QM-cluster model of 312(3) atoms for EBS (EBS-OH). Amino acid residues are shown in sticks and colored by residue name, the catalytic dyad (Cys145, His41), EBS and water molecules W1 and W2 are in balls and sticks. Stars on residues Pro39, Cys44, Phe140, Gly146 and Leu167 point to frozen atoms and delimit the model truncations.

From the plot of solvent accessible surface area at the catalytic pocket (see Figure S6), we can note a smaller area in the case of  $M^{Pro}$ -HID.

From the water radial distribution function (RDF) analysis, obtained as a function of the distance between the water oxygen and side chains of Cys145 and His41 (black and purple line in Figure S9), a different water distribution can be noted in both  $M^{Pro}$ -HID and -HIE, in the first solvation shell of His41, while a similar distribution takes place in that of Cys145. Furthermore, the two water molecules (W1 and W2) are engaged in an extensive network of H-bonds between them and with the His41 imidazole ring. Such behavior confirms that due to the presence of a catalytic dyad (Cys145 and His41) in  $M^{Pro}$  [8,67,68], and differently from other cysteine and serine proteases where an Asp/Glu acts as the third catalytic residue, a buried water molecule plays a catalytic role.

### 2.3. Inhibition Mechanism

The explored inhibition mechanism of  $M^{Pro}$ -HID (Scheme 3) starts with the activation step that takes place from a proton shift between the thiol of Cys145 and the  $N\epsilon$  of the imidazole of His41, affording the Cys-His salt bridge present in the intermediate EI' adduct.



**Scheme 3.** Proposed mechanism for the covalent inhibition of M<sup>Pro</sup> SARS-CoV-2 for HID by EBS in two phases: activation (in violet) and inhibition through covalent bond formation (in dashed orange). X indicates H for EBS and OH for EBS-OH.

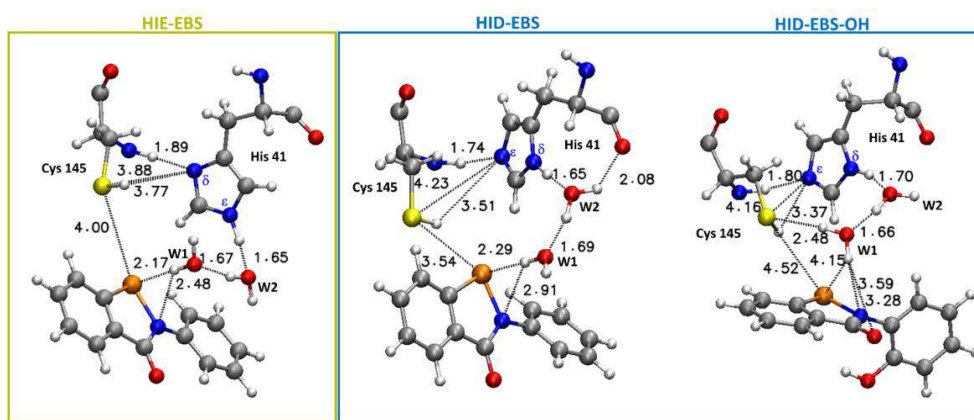
In this way the highly nucleophilic thiolate anion is formed. This step is very crucial for catalysis performed by CPs [69]. The formed efficient nucleophile  $-\text{CH}_2\text{S}^{(-)}$  thiolate, is now able to perform the nucleophile attack on the selenium (TS1) with the formation of a new selenylsulfide bond and the concomitant ring opening, also as result of the water-mediated proton delivery on behalf of His41, finalizing the covalent inhibition (E-I).

Contrary to what occurred in previous theoretical studies on proteolysis and on the inhibition of other CPs [70–72] and in agreement with a more recent work [45], both Cys145 and His41 residues of the catalytic dyad are in the neutral form.

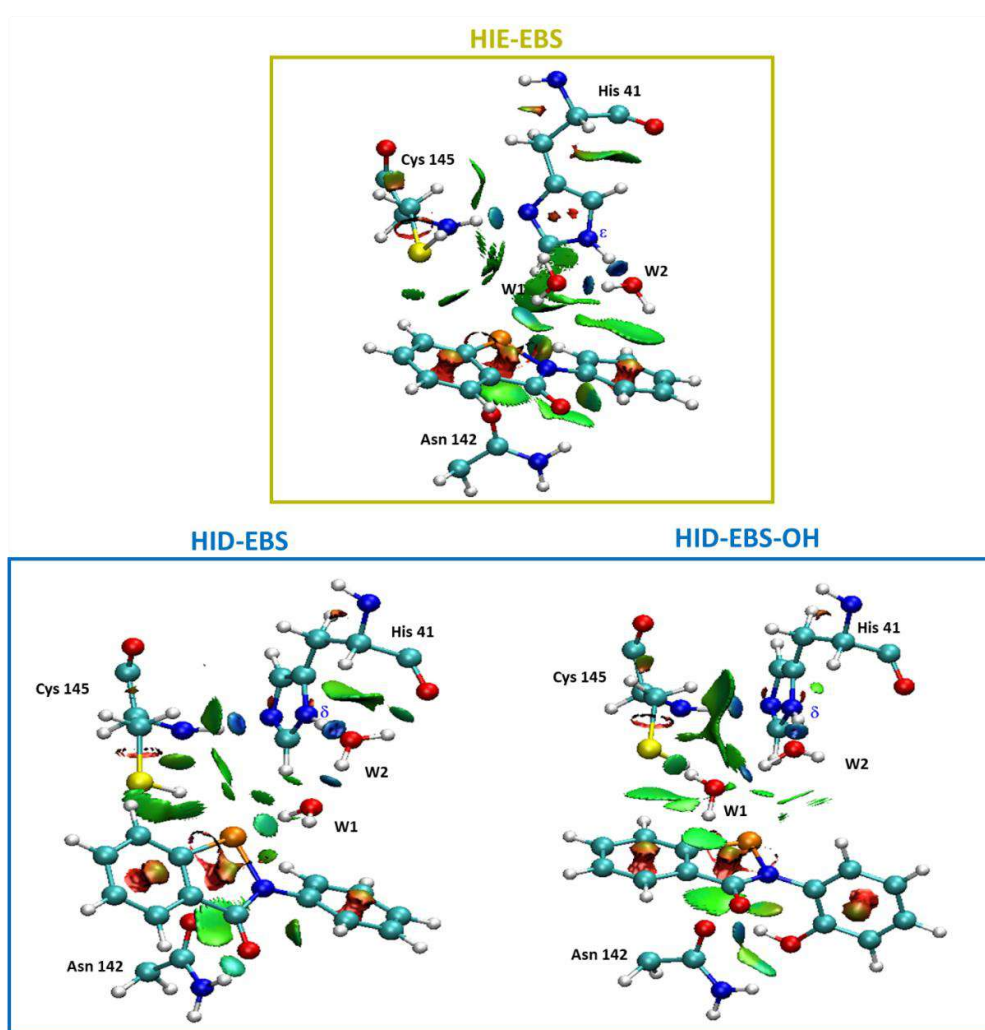
Further indications on the protonation states of the M<sup>Pro</sup> residues comes from pKa values predicted using H<sup>++</sup> computations [73] (Table S1). In fact, the pKa of the Cys145 buried in the active site appears to be lowered with respect to its intrinsic pKa  $\approx 8.6$  at pH = 7 [74], due to interaction with the nearby histidine residue, as has often been found in other CPs for catalytic cysteines [75,76]. Such an alteration of the intrinsic pKa value in the enzyme active site is usually a consequence of the local electronic environment generated by interactions involving the functional groups in the side chains of amino acids as well as the substrate, as documented in many serine protease members [77,78]. Such a scenario is well depicted in the EI optimized complexes referred to for EBS for both the HID and HIE forms (Figure 4), exhibiting the  $-\text{SH}$  group of Cys145 fruitfully oriented towards the imidazole of His41 and engaged in other different interactions.

In fact, the related noncovalent interaction (NCI) plots (Figure 5) evidence that the region including the catalytic dyad is characterized by the presence of strong and attractive H-bonds, while the van der Waals contributions appeared more pronounced in the case of the HIE form for EBS due to the planar arrangement assumed by EBS, in this case with respect to that in the HID one. In Figure 5 the NCI analysis of EI in the M<sup>Pro</sup>-HIE reveals a set of complex interactions between EBS and the amino residues of the catalytic pocket, which arise from a combination of specific hydrogen bonds (atom–atom interactions) with broader surfaces indicative of stabilizing vdW interactions, as a consequence of the planar conformation. In addition, we underline the presence of an interaction between selenium with the thiol group acting as donor of electron density typical of the chalcogen bond. This was more pronounced in the HID form (3.54 Å) than in the HIE one (4.00 Å) as also evidenced by the more extended isosurfaces of the NCI (Figure 5). These findings confirm that noncovalent interactions are of pivotal importance in controlling molecular recognition phenomena. In both M<sup>Pro</sup>-HID and HIE the water molecules trapped inside the active site

span the distance from the  $N_{\epsilon}$  ( $\delta$ ) of His41 and nitrogen of the selenyl amide moiety with typical H-bond values as shown in Figure 4.



**Figure 4.** Structure and main distances (in Å) between atoms of the  $M^{PTro}$  active site and EBS in the EI complex for both HIE (in yellow box) and HID forms and between atoms of the  $M^{PTro}$  active site HID and EBS-OH (in blue box) optimized at the B3LYP-D3/6-31+G (2d,2p) level. For clarity, only a small part of the model is shown. For the full model, see Figure 3.

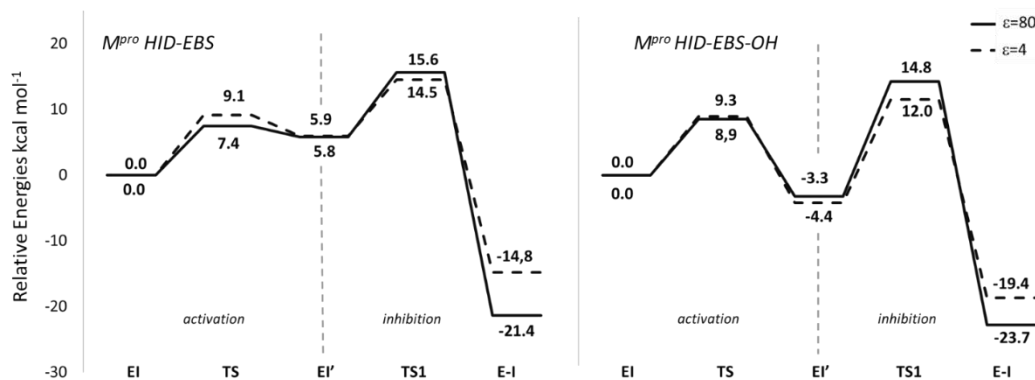


**Figure 5.** Noncovalent interaction plot calculated for the stationary point EI of the  $M^{PTro}$ -HIE inhibition process promoted by EBS (in yellow box).  $M^{PTro}$ -HID inhibition process promoted by EBS and EBS-OH (in blue box). For clarity, only relevant residues were kept.

When EBS-OH is considered, the torsion angle  $\Phi$  between two planes twists the biphenyl-like system to be perpendicular ( $\Phi$  84.5°), driven also by the formation of an H-bond between the OH substituent with the carbonyl side chain of the Asn142 present in the catalytic pocket (see Figure 5). NCI analysis clearly shows how the twisted conformation of the EBS-OH molecule results in a better fitting of the geometry of the active site. In the catalytic pocket of the M<sup>PRO</sup>-HID, the EBS assumes a less accentuated twisted conformation ( $\Phi$  -42.7°) than that in EBS-OH, but more evident than that in EBS in M<sup>PRO</sup>-HIE ( $\Phi$  -16.8°).

### 2.3.1. Inhibition Promoted by EBS

Only the inhibition mechanism related to the M<sup>PRO</sup>-HID will be deeply discussed, while details referring to M<sup>PRO</sup>-HIE are given in the Supplementary Materials (Figure S14 and S15, Table S8). Starting from the EI structure, reported in Figure 4, the first step of the reaction concerns the activation of the neutral cysteine thiol group to generate the thiolate one. This takes place by the formation of the transition state, TS (Figure 6). This transition state clearly depicts the activation of the -SH nucleophile center where the His41 acts as a proton acceptor (H—N $\epsilon$ : 1.68 Å in M<sup>PRO</sup>-HID) from the acid counterpart (thiol group of Cys145), S—H: 1.47 Å in HID. Analysis of the obtained imaginary frequency (Figure 7 and Figure S14) confirms this process. The distance between the sulfur and selenium of EBS is reduced (3.20 Å and 3.05 Å) with respect to the reactant EI, where it assumes the value of 4.00 Å and 3.54 Å in M<sup>PRO</sup>-HIE and M<sup>PRO</sup>-HID, respectively. The related activation energy is found to be 9.1 kcal mol<sup>-1</sup> in the protein environment ( $\epsilon$  = 4) and 7.4 kcal mol<sup>-1</sup> in water ( $\epsilon$  = 80) for the M<sup>PRO</sup>-HID (see Figure 6) [79]. The different barriers found in the activation step for M<sup>PRO</sup>-HIE (17.9 kcal mol<sup>-1</sup> in water and 18.5 kcal mol<sup>-1</sup> in protein, see Figure S15) clearly give a reason for the preferred protonation state (HID) in agreement with previous works [45–47].



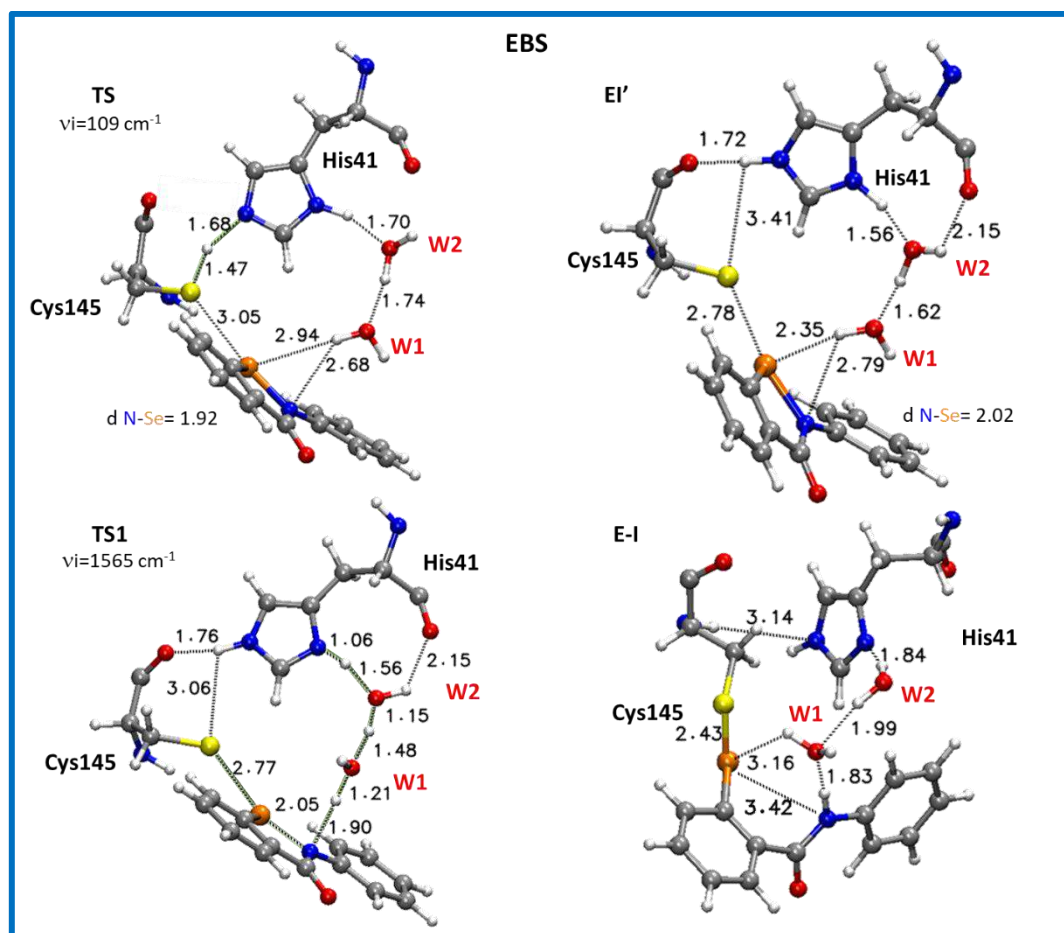
**Figure 6.** B3LYP-D3/6-31+G (2d,2p) free energy profiles of M<sup>PRO</sup> SARS-CoV-2 covalent inhibition mechanism by EBS and EBS-OH in two phases: activation and inhibition through covalent bond formation, expressed in kcal mol<sup>-1</sup>. Solid line indicates calculations in water ( $\epsilon$  = 80) and dotted line in the protein environment ( $\epsilon$  = 4).

Therefore, only the inhibition process promoted by the M<sup>PRO</sup>-HID has been taken into account in the case of EBS-OH (see Scheme 3).

At the end of the activation step, the EI' structure, is characterized by a CysS<sup>-</sup>/HisH<sup>+</sup> zwitterionic pair. Irrespective of an HIE or HID form, our calculations for ES therefore support the finding that Cys145 of the M<sup>PRO</sup> of SARS-CoV-2 works in the neutral form and represents the resting state of the enzyme as previously noted in similar systems [71,80].

Looking at Figure 6 and Figure S15 (HIE), the EI' species representing the noncovalent complex and simultaneously the starting point of the inhibition step and having the negatively charged sulphur, is the key player of the covalent bond formation, and now lies at 2.99 (2.78) Å from selenium and at 2.15 (3.41) Å from N $\delta$ H<sup>+</sup> (N $\epsilon$ H<sup>+</sup>) (see Figure S11). The shorter S–Se distance in the M<sup>PRO</sup>-HID accounts well for the greater stability (by about 4 kcal mol<sup>-1</sup>) than the corresponding species in M<sup>PRO</sup>-HIE in both considered environments (see Figure S15).





**Figure 7.** Optimized structures of the stationary points involved in the M<sup>Pro</sup>-HID inhibition process promoted by EBS. For clarity, only residues of the model implicated in the chemical event are shown. Main distances are in Å. Imaginary frequencies are reported in  $\text{cm}^{-1}$  for all transition states. Cartesian coordinates of the optimized structures with the full model for these transition states as well as for the intermediates along the reaction pathway are given in the Supplementary Materials.

The Se–N “scissile” bond suffers a sensible elongation (2.02 Å for both HIE and HID) making it more prone for its next cleavage.

Differently from the protease mechanisms by the Michael and peptidomimetic inhibitors investigated in recent works [43,45,46], where the acylation step follows the cysteine activation, in the present case the TS1 describes the covalent bond formation triggered by a series of chemical events occurring simultaneously. These are promoted by proton transfer from the protonated His41 mediated by W1 and W2 water molecules that cause the EBS ring opening described by the elongation of the Se–N bond in favor of the Se–S bond formation, also driven by the already present chalcogen bond (Se–S, 2.78 Å) (Figure 7). The TS1 vibrational mode mainly displays a linear stretching of these proton shifts involving the N<sub>ε</sub>, OW2 and OW1 atoms. This hydrogen traffic, promoted by water molecules, remarks their catalytic “help”, occupying the place of a third component of the triad in M<sup>Pro</sup> [43,44,46]. Furthermore, the occurring ring opening in TS1 gives rise to a more negatively charged oxygen of the warhead amide moiety as evidenced by the NBO charges trend (Tables S7 and S8), although it is not directly implicated in the nucleophilic attack by the sulfur atom of Cys145 as in the proteolytic event. Due to the more negative charge assumed by the EBS oxygen (Table S6), the carbonyl moiety engages in an H-bond with the NH of the side chain of Asn142 (Figure S12) remembering the “oxyanion hole” effect in proteases. The barrier to be overcome by EI' for affording the covalent complex E-I, is

9.8 kcal mol<sup>-1</sup> and 8.6 kcal mol<sup>-1</sup> for  $\epsilon=80$  and  $\epsilon = 4$ , respectively (Figure 6) and compares well with the analogous S–Se bond formation previously calculated [42,81–83].

The obtained barriers however do not change the fact that the kinetics of the inhibition by EBS in M<sup>PRO</sup>-HIE is governed by the high barrier of the activation step (see Figure S15), contrary to what is observed for the Michael and peptidomimetic-like inhibitors [43,45].

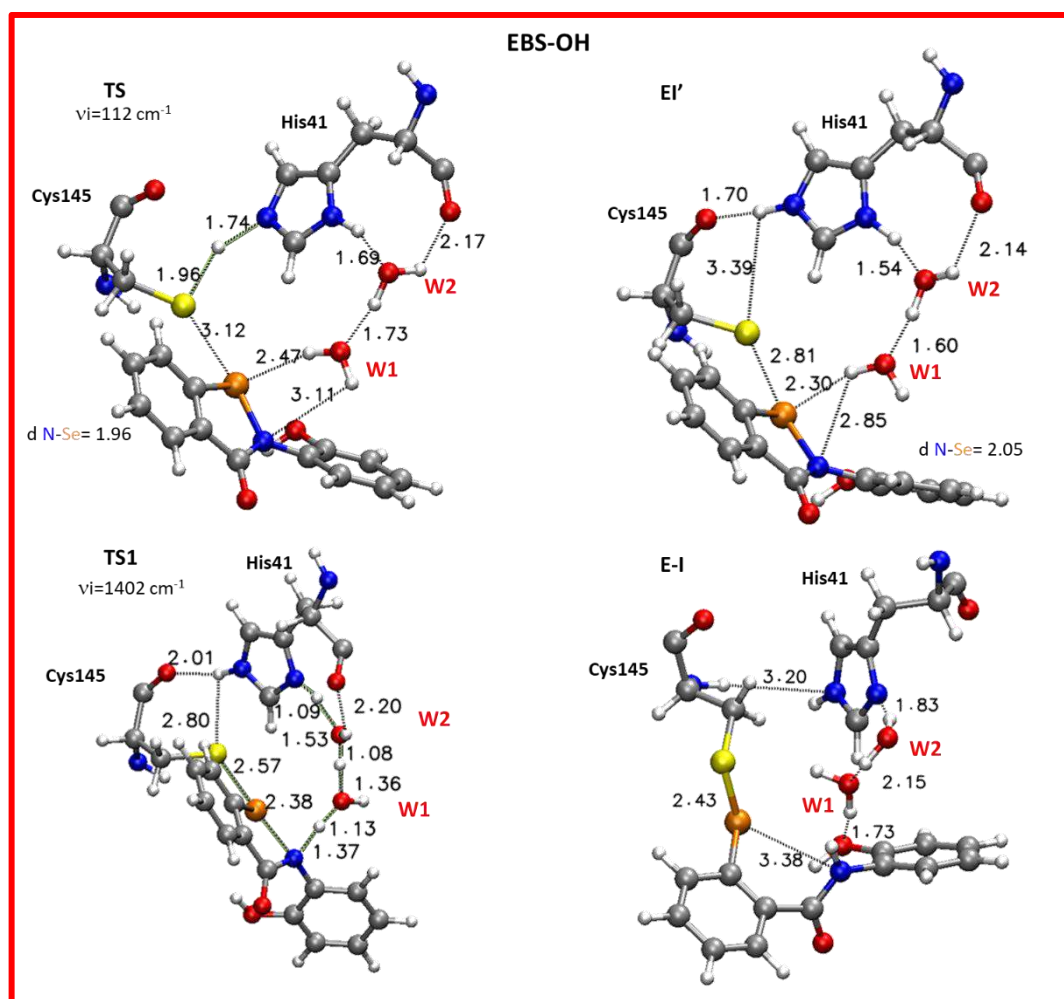
The final covalent product (E-I) accounts for the arrangements observed in TS1, where the S–Se bond is formed (2.43 Å), the Se–N is now completely broken (3.38 Å) and the neutral His41 is restored (Figure 7). The “selenosulfide intermediate nature” of the covalent product in this case, represents the stable end point, although the S–Se bond length is a little bit longer than that present in the literature for biomimetics [40,84,85]. This finding could be a consequence of a series of features created by the protein microenvironment (i.e., the weak interactions present in the catalytic pocket).

### 2.3.2. Inhibition Promoted by EBS-OH

On the basis of the obtained PES values reported in Figure 6, the activation step for EBS-OH requires an energetic amount of 8.9 kcal mol<sup>-1</sup> and 9.3 kcal mol<sup>-1</sup> in  $\epsilon = 80$  and  $\epsilon = 4$ , respectively that is comparable with those obtained in the case of EBS (Figure 6). As previously evidenced in the starting structure EI, the TS also exhibits a distinctive interaction of EBS-OH with the carbonyl of Asn142 (1.67 Å) using the OH moiety as a “hook” for anchoring to the side chain of the amino acid with the twisted conformation assumed by EBS-OH (Figure S13).

As a result of the activation step, the EI' noncovalent complex is formed showing the sulphur anion of the CysS<sup>-</sup>/HisH<sup>+</sup> zwitterionic pair at 2.81 Å from its parent atom (Se) accompanied by a consequent elongation of the selenium–sulfide bond (2.05 Å) (Figure 8). The tight interaction with Asn142 mentioned above represents an important factor for the stabilization of this complex with respect to the corresponding one for EBS. In fact, for EBS-OH, the results for this species results are more stable by 3.3 kcal mol<sup>-1</sup> (4.4 kcal mol<sup>-1</sup>) in a water (protein) environment; related to the starting complex with the neutral Cys145 and His41 (Figure 6) and, as already revealed the formation of the noncovalent complex is an exothermic process. This result represents a distinguishing aspect of EBS-OH with respect to EBS, suggesting the zwitterionic form of the dyad as the most preferred one in the catalytic site of M<sup>PRO</sup>-HID. The exothermicity of the EI' formation evidences a better affinity of EBS-OH towards the catalytic pocket of M<sup>PRO</sup>, strictly related to the inhibition constant as above mentioned for Scheme 2, and this can also help to explain the greater efficiency of EBS-OH than EBS [21]. Observing the related EI' optimized structures for EBS (Figure 7) and EBS-OH (Figure 8) the only difference can be ascribed to the presence of the OH moiety that is involved in an H-bond with the carbonyl of Asn142 during the whole examined mechanism (Figure S13). This aspect reduces the mobility of EBS-OH versus EBS in the catalytic pocket. So, the greater stability of the EI' for EBS-OH (see Figure 6) can justify the improved inhibitory potency observed [21].

Similarly to what occurred in the M<sup>PRO</sup>-HID inhibition promoted by EBS, the process continues with the nucleophilic attack of the S<sub>Y</sub> atoms on the selenium one and the proton transfer from the catalytic histidine to the nitrogen atom is also mediated by water molecules as depicted in the TS1 (see Figure 8). The optimized geometry offers a structure close to the covalent complex (E-I) owing to the formation of a S–Se covalent bond (2.57 Å) in comparison with the analogous complex in EBS (2.77 Å) (Figures 7 and 8). In agreement with other studies devoted to the inhibition process of M<sup>PRO</sup> by species different from our organoselenium compounds, therein investigated [43,45,47], TS1 is the rate limiting step since the barrier results, calculated to be 18.1 kcal mol<sup>-1</sup> (16.4 kcal mol<sup>-1</sup>) using  $\epsilon = 80$  ( $\epsilon = 4$ ), are related to the previous noncovalent complex EI' (see Figure 6).

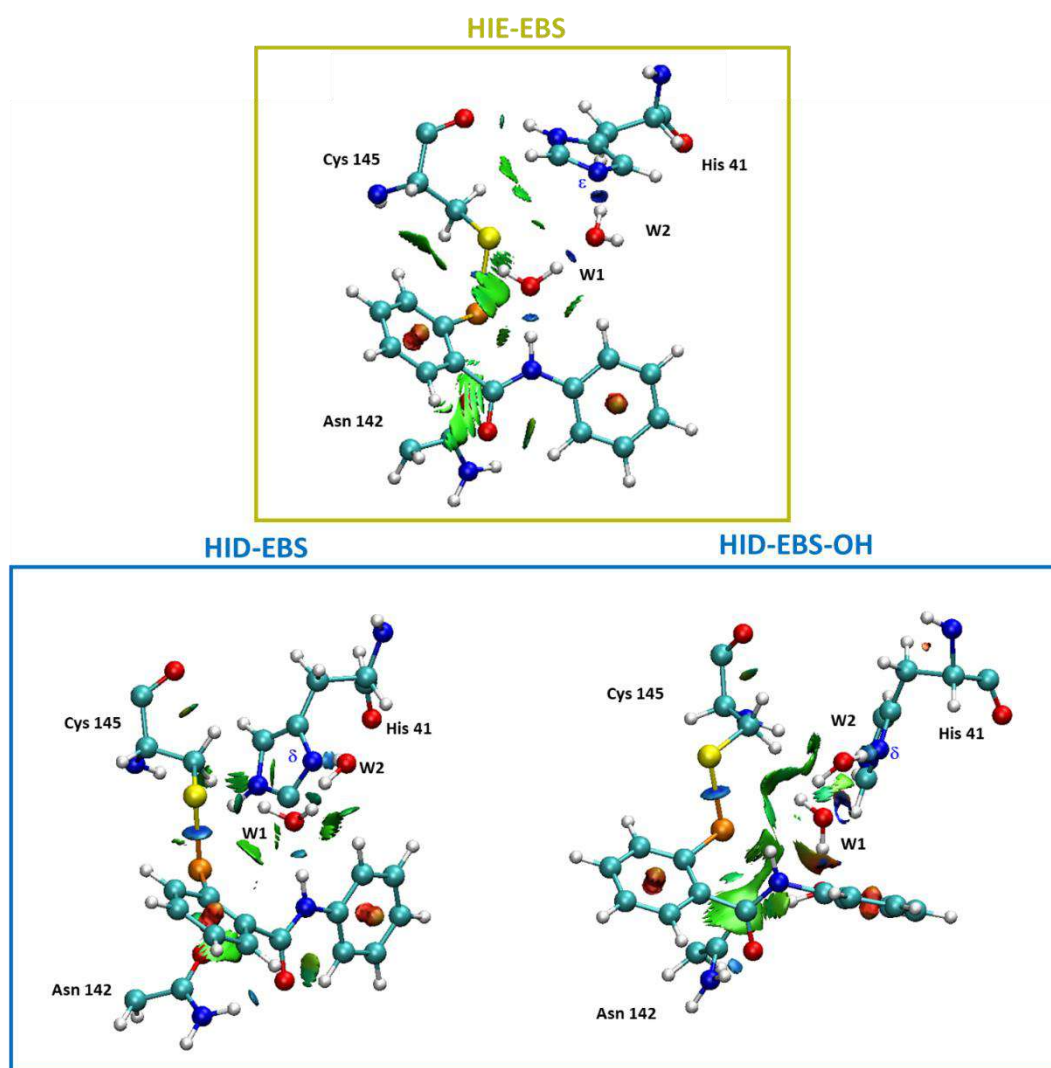


**Figure 8.** Optimized structures of the stationary points involved in the MPTO-HID inhibition process promoted by EBS-OH. For clarity, only residues of the model implicated in the chemical event are shown. Main distances are in Å. Imaginary frequencies are reported in cm<sup>-1</sup> for all transition states. Cartesian coordinates of the optimized structures with the full model for these transition states as well as for the intermediates along the reaction pathway are given in the Supplementary Materials.

In the E-I (Figure 8) the S–Se covalent bond is 2.43 Å while the distance of selenium from the nitrogen is now 3.38 Å. The W2 and W1 molecules are reorganized to establish the H-bond interactions with the inhibition product. From the NCI plots (Figure 9) it is also possible to evidence more extended isosurfaces attributable to vdW contributions as a consequence of the presence of the polarizing OH group with respect to EBS.

For both the considered inhibitors at the end of the inhibition process, the His41 presents a different tautomeric form from the starting complex (Scheme 2 and Schemes S1, Figures 6–8 and Figure S15). The strong stability of the E-I final complex for the covalent S–Se bond formation accounts well for the lack of observed reactivation of enzymes treated by ebselen [21] so, though the conversion between His tautomeric states is not a slow process, in the present studied inhibition it is not enough to recover the enzyme.





**Figure 9.** Noncovalent interaction plot calculated for the final covalent complex (E-I) for HIE (in yellow box) and  $M^{Pro}$  active site HID with EBS and EBS-OH (in blue box). For clarity, only relevant residues were kept.

The inhibition rate constant of SARS-CoV-2 protease by these compounds has not been experimentally determined so it is not possible to compare the calculated barrier with any observed data. From the inhibitory potency investigated for EBS-OH with  $PL^{Pro}$  SARS-CoV-2, an  $IC_{50}$  constant value of 236 nM was found versus 2  $\mu$ M for ebselen. We suppose that a similar behavior could take place in the case of  $M^{Pro}$  of SARS-CoV-2 with a half-maximal inhibitory concentration of EBS-OH better than that of EBS (0.67  $\mu$ M) [15]. On the basis of our results, we would expect a smaller value of  $EC_{50}$  for EBS-OH and therefore a more potent drug.

As can be seen from PESs (Figure 6), and the energy profiles for both EBS and EBS-OH do not suffer drastic variations based upon the effects of the two considered environments (protein and water). This is not surprising if a larger portion of the enzyme treated quantum mechanically (314 atoms) is considered. Indeed, careful systematic studies have widely demonstrated that the solvation effects of the surrounding enzyme environment decrease rapidly and even almost vanish in models with a size of around 200 atoms [61,62,64]. On the basis of our experience [61,62,82,83], and because of the comparative nature of the present study, the QM/MM methodology is not expected to introduce a big perturbation on the mechanism but mainly an improved energetic behavior.

That said, some deviation occurred in the case of  $\epsilon = 4$  that proposes for both inhibitors a minor exothermicity of the covalent product (see Figure 6). Probably the dielectric constant value of the water tends to evidence the effects of the –OH moiety.

Since the reversibility of the reaction depends on the barrier of the reverse reaction, and with the inhibition process promoted by EBS and EBS-OH showing comparable activation barriers, the process with more exothermic results will also be more irreversible. On the basis of the described behavior of EBS and EBS-OH in  $M^{PTO}$ -HID so far, our analysis leads us to propose [21] EBS-OH to be an irreversible inhibitor stronger than EBS. The assumption that the reactions of irreversible inhibitors are characterized by exothermicity higher than  $-22 \text{ kcal mol}^{-1}$  [86] emphasizes our findings.

Starting from the principle that no universal rule exists for the choice of the “best” functionals, the B3LYP results for the inhibition process were benchmarked against the latest-generation density functionals, such as M06-2X,  $\omega$ B97X, M06, M06-2X and M08HX in order to point out an adequate level of theory to use in future investigations on the inhibition promoted by ebselen derivatives. Results are collected in Table 1. Single-point calculations were carried out on the B3LYP optimized geometries for the  $M^{PTO}$ -EBS system and the whole benchmark was limited to the ebselen inhibitor. In fact, in the case of EBS-OH, the single point calculations with the above mentioned functionals were performed only on the EI and EI' complexes due to the marked behavior of the non-covalent complex (EI') with respect to that of EBS, with its implications for the future design of possible inhibitors with improved kinetics.

**Table 1.** Gibbs free energies for  $M^{PTO}$ -HID inhibition promoted by EBS calculated as a single point on B3LYP-D3/6-31+G (2d,2p) optimized geometries in water (protein) environments. In bold, values are reported only in water, the corresponding ones for the EBS-OH inhibitor are related to the EI' complex for the  $M^{PTO}$ -HID-EBS-OH system. All values are in  $\text{kcal mol}^{-1}$ .

	B3LYP-D3	M06	M062X	M08HX	$\omega$ B97X
EI	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
TS	7.4 (9.1)	7.5 (7.9)	6.1 (9.6)	0.7 (2.1)	0.9 (3.9)
EI'	5.8 (5.9) <b>−4.4</b>	0.2 (3.1) <b>−6.2</b>	0.9 (2.3) <b>−2.7</b>	2.3 (1.4) <b>−0.33</b>	−0.8 (−0.9) <b>−3.6</b>
TS1	15.6 (14.5)	11.0 (13.9)	9.2 (12.8)	8.2 (11.9)	10.5 (11.5)
E-I	−21.4 (−14.8)	−15.4 (−12.1)	−15.8 (−11.7)	−16.0 (−12.3)	−23.6 (−16.2)

In general, it is noteworthy that the first step of the reaction leading to the noncovalent complex is the most sensitive to the used functionals, in particular for M08HX and  $\omega$ B97X, in terms of barrier heights which become low enough to suggest that the activation for obtaining the zwitterionic catalytic dyad takes place more easily and regardless of the dielectric constant value (see Table 1). All the functionals, indeed, propose a very similar stability of the EI' to that of the EI starting complex, suggesting that both the forms could be responsible for the next inhibition step.

The same does not apply for the second step of the reaction, since from the results reported in Table 1 it emerges that the B3LYP functional provides a good description of the energetics of this kind of process. This finding also gives confirmation that the effect of the dielectric constant value on the DFT cluster model calculations saturates well with increasing the system size [87], thus the cluster used appears to be adequate for this system, providing information that should be exploited to generate good candidate SARS-CoV-2 inhibitors.

The  $\omega$ B97X functional (Table 1) results show it to be the best performing one as known in the literature [63,84], but this does not significantly affect our findings.

In the case of EBS-OH, the exothermicity of the formation of the EI' complex is confirmed by the used functionals (see Table 1) corroborating the outcomes arising from the B3LYP investigation.

### 3. Computational Methods

As far as the molecular dynamics simulations are concerned, a detailed description is given in the Supplementary Materials. Details on docking protocol procedure including figures (Figures S10 and S11) and table (Table S5) are given in the Supplementary Materials. The best docked pose shown in Figure S11 has been used as the starting structure for the inhibition mechanism at the QM level, and geometry optimizations were performed with a B3LYP/D3 functional [88–91] and 6-31+G basis set for all the atoms as implemented in the Gaussian 09 package [92]. To quantify the ZPE corrections, frequencies were calculated at the same level of theory, excluding the contributions of frozen atoms in the vibrational analysis [93]. To evaluate the environment effects, single-point calculations B3LYP-D3/6-31+G(2d,2p) in the framework of the SMD [94] model were performed on the optimized geometries by using the dielectric constants for protein ( $\epsilon = 4$ ) and water ( $\epsilon = 80$ ). The final energies reported are solvation energies arising from single point calculations with the larger basis set corrected for ZPE. In order to test if the 6-31+G basis set used was adequate for describing the inhibition phase, we conducted further calculations using 6-311+G(2d,p) and AUG-cc-pVTZ basis sets and the results are reported in Table S10. These basis sets confirm the results with the smallest one, since the obtained energy barriers of the inhibition process remain essentially unchanged.

Other latest-generation functionals explicitly developed for systems with long-range electron correlations were employed on the M<sup>Pro</sup>-HID inhibition process promoted by EBS for benchmarking, such as M06 [95], M06-2X [95],  $\omega$ B97X [96], and M08HX [97], which have been recently shown to be among the best-performing density functionals in the calculations on many systems [84,85,98–100]. The M06, M06-2x and M08HX functionals usually offer good results for main-group chemistry, including thermochemistry, excitation energies, barrier heights, and noncovalent interaction energies with a better performance of M08-HX for the latter ones [101].  $\omega$ B97X belongs to the long-range-corrected functionals [96] and was revealed to be useful to better discriminate the chemistry at stake in the two phases of the inhibition mechanism. It is known that the cluster model presents some limitations, such as that a single-conformation QM cluster approach does not explicitly take into consideration the potential influence of different enzyme conformations on the reaction energetics. There is therefore, a consequent reliance on the starting structure and the use of an implicit solvent model to simulate the cluster's environment, as discussed very deeply in previous excellent papers [7,10,64,102]. However, it was revealed to be appropriate to study enzymatic reactions, also including their inhibition processes [60–62,65,81]. NBO [103] analysis was carried out on all the stationary points intercepted on the PESs and the related results are collected in Tables S6–S9. All the results related to the QM investigation on M<sup>Pro</sup>-HIE are reported in the Supplementary Materials.

### 4. Conclusions

In the present work, the covalent inhibition of the main protease from SARS CoV-2 by repurposing organoselenium compounds, such as ebselen and its derivative on the position 2 of benzoselenazol-3(2H)-one ring, was investigated at the QM level in the framework of density functional theory. By using a large cluster of 314 atoms our quantum chemical computations have clarified, at an atomistic level, the inhibition mechanism of two ebselen-like non-Michael acceptor covalent inhibitors of M<sup>Pro</sup>. This enzyme plays a central role in the SARS-CoV-2 viral life cycle and for this reason became an attractive drug target against the COVID-19 disease. The inhibition process occurs in two steps: Cys145 is activated by His41, forming the ion pair Cys-His. The formed nucleophile species (deprotonated Cys145) performs the nucleophile attack on the selenium with the formation of the selenylsulfide bond, and takes place in concert with the ring opening promoted by the water-mediated proton delivery by His41, leading to the stable covalent enzyme-inhibitor complex. The S–Se covalent bond formation, the focus of covalent inhibition and a common step of the two examined inhibitors, implies a series of the chemical events that are caught well in the stationary points intercepted through potential energy surfaces. For

both inhibitors, the process result is irreversible even if it is more accentuated in the case of EBS-OH. Our outcomes allow us to suggest that, as already found in PL<sup>pro</sup> SARS-CoV-2, in M<sup>pro</sup> SARS-CoV-2, EBS-OH may result in being a more potent drug. Our calculations have provided deeper insights on the observed inhibitory potency of EBS-OH with respect to EBS as presented in the greater stability of the EI' complex relative to that of EBS. The cluster used appears to be adequate for the examined inhibition, providing information that should be exploited to generate good candidate SARS-CoV-2 inhibitors. Our results could contribute to increase the current knowledge of small-molecule covalent inhibitors and stimulate the design of drugs obtained by their incorporation into new hybrid non-Michael acceptor inhibitors. Although the winning strategy such as vaccines is available against the virus, the resource of antiviral drugs represents a helpful approach to reduce the signs of the disease caused by SARS-CoV-2 or by its variants.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ijms22189792/s1>.

**Author Contributions:** T.M. conceived of the presented idea and supervised the project. A.P. and I.R. performed the computations. A.P. and T.M. verified and analyzed the results. T.M. wrote the manuscript. T.M. and N.R. contributed to the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors thank the Dipartimento di Chimica e Tecnologie Chimiche of Università della Calabria for their financial support. The computing resources (in particular Visual Molecular Dynamics VMD 1.9.2) and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [69]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programs (see <http://www.cresco.enea.it/english> accessed on February 2021. for information).

**Conflicts of Interest:** The authors declare no conflict of interest. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; Qiu, Y.; Wang, J.; Liu, Y.; Wei, Y.; et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *Lancet* **2020**, *395*, 507–513. [[CrossRef](#)]
2. Hui, D.S.; Zumla, P.S.A. Severe Acute Respiratory Syndrome: Historical, Epidemiologic, and Clinical Features. *Infect. Dis. Clin. N. Am.* **2019**, *33*, 869–889. [[CrossRef](#)] [[PubMed](#)]
3. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.-L.; Abiona, O.; Graham, B.S.; McLellan, J.S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263. [[CrossRef](#)] [[PubMed](#)]
4. Woo, P.C.Y.; Huang, Y.; Lau, S.K.P.; Yuen, K.-Y. Coronavirus Genomics and Bioinformatics Analysis. *Viruses* **2010**, *2*, 1804–1820. [[CrossRef](#)]
5. Walls, A.C.; Park, Y.-J.; Tortorici, M.A.; Wall, A.; McGuire, A.T.; Velesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281–292.e6. [[CrossRef](#)]
6. Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S.-H. An Overview of Severe Acute Respiratory Syndrome–Coronavirus (SARS-CoV) 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy. *J. Med. Chem.* **2016**, *59*, 6595–6628. [[CrossRef](#)]
7. Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J.; et al. Design of Wide-Spectrum Inhibitors Targeting Coronavirus Main Proteases. *PLoS Biol.* **2005**, *3*, e324. [[CrossRef](#)]
8. Anand, K.; Ziebuhr, J.; Wadhvani, P.; Mesters, J.R.; Hilgenfeld, R. Coronavirus Main Proteinase (3CL<sup>pro</sup>) Structure: Basis for Design of Anti-SARS Drugs. *Science* **2003**, *300*, 1763–1767. [[CrossRef](#)] [[PubMed](#)]
9. Ren, Z.; Yan, L.; Zhang, N.; Guo, Y.; Yang, C.; Lou, Z.; Rao, Z. The newly emerged SARS-Like coronavirus HCoV-EMC also has an “Achilles’ heel”: Current effective inhibitor targeting a 3C-like protease. *Protein Cell* **2013**, *4*, 248–250. [[CrossRef](#)]
10. Anand, K.; Palm, G.J.; Mesters, J.R.; Siddell, S.G.; Ziebuhr, J.; Hilgenfeld, R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J.* **2002**, *21*, 3213–3224. [[CrossRef](#)]
11. Ullrich, S.; Nitsche, C. The SARS-CoV-2 Main Protease as Drug Target. *Bioorganic Med. Chem. Lett.* **2020**, *30*, 127377. [[CrossRef](#)]



12. Verma, N.; Henderson, J.A.; Shen, J. Proton-Coupled Conformational Activation of SARS Coronavirus Main Proteases and Opportunity for Designing Small-Molecule Broad-Spectrum Targeted Covalent Inhibitors. *J. Am. Chem. Soc.* **2020**, *142*, 21883–21890. [[CrossRef](#)] [[PubMed](#)]
13. Citarella, A.; Scala, A.; Piperno, A.; Micale, N. SARS-CoV-2 Mpro: A Potential Target for Peptidomimetics and Small-Molecule Inhibitors. *Biomolecules* **2021**, *11*, 607. [[CrossRef](#)]
14. Mengist, H.M.; Dilnessa, T.; Jin, T. Structural Basis of Potential Inhibitors Targeting SARS-CoV-2 Main Protease. *Front. Chem.* **2021**, *9*, 622898. [[CrossRef](#)]
15. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289–293. [[CrossRef](#)]
16. Haritha, C.; Sharun, K.; Jose, B. Ebselen, a new candidate therapeutic against SARS-CoV-2. *Int. J. Surg.* **2020**, *84*, 53–56. [[CrossRef](#)] [[PubMed](#)]
17. Menéndez, C.A.; Byléhn, F.; Perez-Lemus, G.R.; Alvarado, W.; De Pablo, J.J. Molecular characterization of ebselen binding activity to SARS-CoV-2 main protease. *Sci. Adv.* **2020**, *6*, eabd0345. [[CrossRef](#)]
18. Sies, H.; Parnham, M.J. Potential therapeutic use of ebselen for COVID-19 and other respiratory viral infections. *Free. Radic. Biol. Med.* **2020**, *156*, 107–112. [[CrossRef](#)]
19. Banerjee, R.; Perera, L.; Tillekeratne, L.V. Potential SARS-CoV-2 main protease inhibitors. *Drug Discov. Today* **2020**, *26*, 804–816. [[CrossRef](#)] [[PubMed](#)]
20. Goins, C.M.; Dajnowicz, S.; Thanna, S.; Sucheck, S.J.; Parks, J.M.; Ronning, D.R. Exploring Covalent Allosteric Inhibition of Antigen 85C from Mycobacterium tuberculosis by Ebselen Derivatives. *ACS Infect. Dis.* **2017**, *3*, 378–387. [[CrossRef](#)] [[PubMed](#)]
21. Weglarz-Tomczak, E.; Tomczak, J.M.; Talma, M.; Burda-Grabowska, M.; Giurg, M.; Brul, S. Identification of ebselen and its analogues as potent covalent inhibitors of papain-like protease from SARS-CoV-2. *Sci. Rep.* **2021**, *11*, 1–10. [[CrossRef](#)]
22. Parnham, M.; Sies, H. Ebselen: Prospective therapy for cerebral ischaemia. *Expert Opin. Investig. Drugs* **2000**, *9*, 607–619. [[CrossRef](#)]
23. Schewe, T. Molecular actions of Ebselen—An antiinflammatory antioxidant. *Gen. Pharmacol. Vasc. Syst.* **1995**, *26*, 1153–1169. [[CrossRef](#)]
24. Favrot, L.; Grzegorzewicz, A.E.; Lajiness, D.H.; Marvin, R.K.; Boucau, J.; Isailovic, D.; Jackson, M.; Ronning, D.R. Mechanism of inhibition of Mycobacterium tuberculosis antigen 85 by ebselen. *Nat. Commun.* **2013**, *4*, 1–10. [[CrossRef](#)]
25. Bzówka, M.; Mitusińska, K.; Raczyńska, A.; Samol, A.; Tuszyński, J.A.; Góra, A. Structural and Evolutionary Analysis Indicate That the SARS-CoV-2 Mpro Is a Challenging Target for Small-Molecule Inhibitor Design. *Int. J. Mol. Sci.* **2020**, *21*, 3099. [[CrossRef](#)]
26. Singh, N.; Halliday, A.C.; Thomas, J.M.; Kuznetsova, O.V.; Baldwin, R.; Woon, E.; Aley, P.K.; Antoniadou, I.; Sharp, T.; Vasudevan, S.R.; et al. A safe lithium mimetic for bipolar disorder. *Nat. Commun.* **2013**, *4*, 1–7. [[CrossRef](#)]
27. Kil, J.; Lobarinas, E.; Spankovich, C.; Griffiths, S.K.; Antonelli, P.J.; Lynch, E.D.; Le Prell, C. Safety and efficacy of ebselen for the prevention of noise-induced hearing loss: A randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* **2017**, *390*, 969–979. [[CrossRef](#)]
28. Singh, N.; Sharpley, A.L.; Emir, U.; Masaki, C.; Herzallah, M.; Gluck, M.A.; Sharp, T.; Harmer, C.; Vasudevan, S.R.; Cowen, P.; et al. Effect of the Putative Lithium Mimetic Ebselen on Brain Myo-Inositol, Sleep, and Emotional Processing in Humans. *Neuropsychopharmacology* **2015**, *41*, 1768–1778. [[CrossRef](#)] [[PubMed](#)]
29. Kil, J.; Pierce, C.; Tran, H.; Gu, R.; Lynch, E.D. Ebselen treatment reduces noise induced hearing loss via the mimicry and induction of glutathione peroxidase. *Hear. Res.* **2007**, *226*, 44–51. [[CrossRef](#)] [[PubMed](#)]
30. Zheng, W.; He, R.; Boada, R.; Subirana, M.A.; Ginman, T.; Ottosson, H.; Valiente, M.; Zhao, Y.; Hassan, M. A general covalent binding model between cytotoxic selenocompounds and albumin revealed by mass spectrometry and X-ray absorption spectroscopy. *Sci. Rep.* **2020**, *10*, 1–10. [[CrossRef](#)]
31. Thenin-Houssier, S.; de Vera, I.M.S.; Pedro-Rosa, L.; Brady, A.; Richard, A.; Konnick, B.; Opp, S.; Buffone, C.; Fuhrmann, J.; Kota, S.; et al. Ebselen, a Small-Molecule Capsid Inhibitor of HIV-1 Replication. *Antimicrob. Agents Chemother.* **2016**, *60*, 2195–2208. [[CrossRef](#)] [[PubMed](#)]
32. De Munnik, M.; Lohans, C.T.; Lang, P.A.; Langley, G.; Malla, T.R.; Tumber, A.; Schofield, C.J.; Brem, J. Targeting the Mycobacterium tuberculosis transpeptidase LdtMt2 with cysteine-reactive inhibitors including ebselen. *Chem. Commun.* **2019**, *55*, 10214–10217. [[CrossRef](#)]
33. Bender, K.O.; Garland, M.; Ferreyra, J.A.; Hryckowian, A.J.; Child, M.; Puri, A.W.; Solow-Cordero, D.; Higginbottom, S.K.; Segal, E.; Banaei, N.; et al. A small-molecule antivirulence agent for treating Clostridium difficile infection. *Sci. Transl. Med.* **2015**, *7*, 306ra148. [[CrossRef](#)]
34. Mukherjee, S.; Weiner, W.S.; Schroeder, C.E.; Simpson, D.S.; Hanson, A.M.; Sweeney, N.L.; Marvin, R.K.; Ndjomou, J.; Kolli, R.; Isailovic, D.; et al. Ebselen Inhibits Hepatitis C Virus NS3 Helicase Binding to Nucleic Acid and Prevents Viral Replication. *ACS Chem. Biol.* **2014**, *9*, 2393–2403. [[CrossRef](#)] [[PubMed](#)]
35. Nikawa, T.; Schuch, G.; Wagner, G.; Sies, H. Interaction of ebselen with glutathione S-transferase and papain in vitro. *Biochem. Pharmacol.* **1994**, *47*, 1007–1012. [[CrossRef](#)]
36. Azad, G.; Singh, V.; Mandal, P.; Singh, P.; Golla, U.R.; Baranwal, S.; Chauhan, S.; Tomar, R.S. Ebselen induces reactive oxygen species (ROS)-mediated cytotoxicity in Saccharomyces cerevisiae with inhibition of glutamate dehydrogenase being a target. *FEBS Open Bio* **2014**, *4*, 77–89. [[CrossRef](#)] [[PubMed](#)]

37. Zmudzinski, M.; Rut, W.; Olech, K.; Granda, J.; Giurg, M.; Burda-Grabowska, M.; Zhang, L.; Sun, X.; Lv, Z.; Nayak, D.; et al. Ebselen Derivatives Are Very Potent Dual Inhibitors of SARS-CoV-2 Proteases—PLpro and Mpro in in Vitro Studies. *bioRxiv* **2020**. [CrossRef]
38. Ma, C.; Hu, Y.; Townsend, J.A.; Lagarias, P.I.; Marty, M.T.; Kolocouris, A.; Wang, J. Ebselen, Disulfiram, Carmofur, PX-12, Tideglusib, and Shikonin Are Nonspecific Promiscuous SARS-CoV-2 Main Protease Inhibitors. *ACS Pharmacol. Transl. Sci.* **2020**, *3*, 1265–1277. [CrossRef]
39. Węglarz-Tomczak, E.; Tomczak, J.M.; Giurg, M.; Burda-Grabowska, M.; Brul, S. Discovery of Potent Inhibitors of PLproCoV2 by Screening a Library of Selenium-Containing Compounds. *bioRxiv* **2020**. [CrossRef]
40. Cerqueira, N.M.F.S.A.; Fernandes, P.A.; Gonzalez, P.J.; Moura, J.J.G.; Ramos, M.J. The Sulfur Shift: An Activation Mechanism for Periplasmic Nitrate Reductase and Formate Dehydrogenase. *Inorg. Chem.* **2013**, *52*, 10766–10772. [CrossRef]
41. Cheng, Q.; Sandalova, T.; Lindqvist, Y.; Arnér, E. Crystal Structure and Catalysis of the Selenoprotein Thioredoxin Reductase 1. *J. Biol. Chem.* **2009**, *284*, 3998–4008. [CrossRef]
42. Fortino, M.; Marino, T.; Russo, N.; Sicilia, E. Mechanism of Thyroxine Deiodination by Naphthyl-Based Iodothyronine Deiodinase Mimics and the Halogen Bonding Role: A DFT Investigation. *Chem.—Eur. J.* **2015**, *21*, 8554–8560. [CrossRef]
43. Mondal, D.; Warshel, A. Exploring the Mechanism of Covalent Inhibition: Simulating the Binding Free Energy of  $\alpha$ -Ketoamide Inhibitors of the Main Protease of SARS-CoV-2. *Biochemistry* **2020**, *59*, 4601–4608. [CrossRef]
44. 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. [CrossRef]
45. Arafet, K.; Serrano-Aparicio, N.; Lodola, A.; Mulholland, A.J.; González, F.V.; Świderek, K.; Moliner, V. Mechanism of inhibition of SARS-CoV-2 Mpro by N3 peptidyl Michael acceptor explained by QM/MM simulations and design of new derivatives with tunable chemical reactivity. *Chem. Sci.* **2021**, *12*, 1433–1444. [CrossRef]
46. Ramos-Guzmán, C.A.; Ruiz-Pernía, J.J.; Tuñón, I. Multiscale Simulations of SARS-CoV-2 3CL Protease Inhibition with Aldehyde Derivatives. Role of Protein and Inhibitor Conformational Changes in the Reaction Mechanism. *ACS Catal.* **2021**, *11*, 4157–4168. [CrossRef] [PubMed]
47. Ramos-Guzmán, C.A.; Ruiz-Pernía, J.J.; Tuñón, I. A microscopic description of SARS-CoV-2 main protease inhibition with Michael acceptors. Strategies for improving inhibitor design. *Chem. Sci.* **2021**, *12*, 3489–3496. [CrossRef] [PubMed]
48. Pavlova, A.; Lynch, D.L.; Daidone, I.; Zanetti-Polzi, L.; Smith, M.D.; Chipot, C.; Kneller, D.W.; Kovalevsky, A.; Coates, L.; Golosov, A.A.; et al. Inhibitor binding influences the protonation states of histidines in SARS-CoV-2 main protease. *Chem. Sci.* **2021**, *12*, 1513–1527. [CrossRef]
49. Zanetti-Polzi, L.; Smith, M.D.; Chipot, C.; Gumbart, J.C.; Lynch, D.L.; Pavlova, A.; Smith, J.C.; Daidone, I. Tuning Proton Transfer Thermodynamics in SARS-Cov-2 Main Protease: Implications for Catalysis and Inhibitor Design. *ChemRxiv Prepr. Serv. Chem.* **2020**, *12*, 1513–1527. [CrossRef]
50. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved  $\alpha$ -Ketoamide Inhibitors. *Science* **2020**, *368*, 409–412. [CrossRef] [PubMed]
51. Kneller, D.W.; Phillips, G.; O’Neill, H.M.; Jedrzejczak, R.; Stols, L.; Langan, P.; Joachimiak, A.; Coates, L.; Kovalevsky, A. Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat. Commun.* **2020**, *11*, 1–6. [CrossRef]
52. Hung, H.C.; Ke, Y.Y.; Huang, S.Y.; Huang, P.N.; Kung, Y.A.; Chang, T.Y.; Yen, K.J.; Peng, T.T.; Chang, S.E.; Huang, C.T.; et al. Discovery of M Protease Inhibitors Encoded by SARS-CoV-2. *Antimicrob. Agents Chemother.* **2020**, *64*, e00872-20. [CrossRef]
53. Rathnayake, A.D.; Zheng, J.; Kim, Y.; Perera, K.D.; Mackin, S.; Meyerholz, D.K.; Kashipathy, M.M.; Battaile, K.P.; Lovell, S.; Perlman, S.; et al. 3C-like protease inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV-infected mice. *Sci. Transl. Med.* **2020**, *12*, eabc5332. [CrossRef] [PubMed]
54. Ma, C.; Sacco, M.D.; Hurst, B.; Townsend, J.A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M.; Chen, Y.; et al. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res.* **2020**, *30*, 678–692. [CrossRef]
55. Jin, Z.; Zhao, Y.; Sun, Y.; Zhang, B.; Wang, H.; Wu, Y.; Zhu, Y.; Zhu, C.; Hu, T.; Du, X.; et al. Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nat. Struct. Mol. Biol.* **2020**, *27*, 529–532. [CrossRef] [PubMed]
56. Structure of SARS-CoV-2 Main Protease Bound to Potent Broad-Spectrum Non-Covalent Inhibitor X77. 2020. Available online: <https://www.rcsb.org/structure/6W63> (accessed on 25 March 2020).
57. Paasche, A.; Zipper, A.; Schäfer, S.; Ziebuhr, J.; Schirmeister, T.; Engels, B. Evidence for Substrate Binding-Induced Zwitterion Formation in the Catalytic Cys-His Dyad of the SARS-CoV Main Protease. *Biochemistry* **2014**, *53*, 5930–5946. [CrossRef]
58. Paasche, A.; Schirmeister, T.; Engels, B. Benchmark Study for the Cysteine–Histidine Proton Transfer Reaction in a Protein Environment: Gas Phase, COSMO, QM/MM Approaches. *J. Chem. Theory Comput.* **2013**, *9*, 1765–1777. [CrossRef]
59. Baker, N.A.; Sept, D.; Joseph, S.; Holst, M.J.; McCammon, J.A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10037–10041. [CrossRef]
60. Prejanò, M.; Medina, F.E.; Fernandes, P.; Russo, N.; Ramos, M.J.; Marino, T. The Catalytic Mechanism of Human Transketolase. *ChemPhysChem* **2019**, *20*, 2881–2886. [CrossRef] [PubMed]



61. Prejanò, M.; Romeo, I.; Sgrizzi, L.; Russo, N.; Marino, T. Why hydroxy-proline improves the catalytic power of the peptidoglycan N-deacetylase enzyme: Insight from theory. *Phys. Chem. Chem. Phys.* **2019**, *21*, 23338–23345. [[CrossRef](#)]
62. Parise, A.; Russo, N.; Marino, T. The RNase A Platination Process by Arsenoplatin: Insights from Theoretical Study. *Inorg. Chem. Front.* **2021**, *8*, 1795–1803. [[CrossRef](#)]
63. Neves, R.P.P.; Fernandes, P.; Ramos, M.J. Mechanistic insights on the reduction of glutathione disulfide by protein disulfide isomerase. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E4724–E4733. [[CrossRef](#)] [[PubMed](#)]
64. Blomberg, M.R.A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P.E.M. Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601–3658. [[CrossRef](#)] [[PubMed](#)]
65. Himo, F. Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780–6786. [[CrossRef](#)] [[PubMed](#)]
66. Siegbahn, P.E.M.; Himo, F. Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. *JBIC J. Biol. Inorg. Chem.* **2009**, *14*, 643–651. [[CrossRef](#)]
67. Ziebuhr, J. Molecular biology of severe acute respiratory syndrome coronavirus. *Curr. Opin. Microbiol.* **2004**, *7*, 412–419. [[CrossRef](#)]
68. Ho, B.-L.; Cheng, S.-C.; Shi, L.; Wang, T.-Y.; Ho, K.-I.; Chou, C.-Y. Critical Assessment of the Important Residues Involved in the Dimerization and Catalysis of MERS Coronavirus Main Protease. *PLoS ONE* **2015**, *10*, e0144865. [[CrossRef](#)] [[PubMed](#)]
69. Awoonor-Williams, E.; Rowley, C.N. Evaluation of Methods for the Calculation of the pKa of Cysteine Residues in Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 4662–4673. [[CrossRef](#)]
70. Arafet, K.; González, F.V.; Moliner, V. Quantum Mechanics/Molecular Mechanics Studies of the Mechanism of Cysteine Proteases Inhibition by Dipeptidyl Nitroalkenes. *Chem.—Eur. J.* **2020**, *26*, 2002–2012. [[CrossRef](#)]
71. Silva, J.R.A.; Cianni, L.; Araujo, D.; Batista, P.H.J.; de Vita, D.; Rosini, F.; Leitao, A.; Lameira, J.; Montanari, C.A. Assessment of the Cruzain Cysteine Protease Reversible and Irreversible Covalent Inhibition Mechanism. *J. Chem. Inf. Model.* **2020**, *60*, 1666–1677. [[CrossRef](#)]
72. Arafet, K.; Ferrer, S.; Moliner, V. Computational Study of the Catalytic Mechanism of the Cruzain Cysteine Protease. *ACS Catal.* **2017**, *7*, 1207–1215. [[CrossRef](#)]
73. Anandakrishnan, R.; Aguilar, B.; Onufriev, A.V. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012**, *40*, W537–W541. [[CrossRef](#)]
74. Roos, G.; Foloppe, N.; Messens, J. Understanding the pKa of Redox Cysteines: The Key Role of Hydrogen Bonding. *Antioxid. Redox Signal.* **2013**, *18*, 94–127. [[CrossRef](#)]
75. Cianni, L.; Feldmann, C.W.; 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. [[CrossRef](#)] [[PubMed](#)]
76. Quesne, M.G.; Ward, R.A.; De Visser, S.P. Cysteine protease inhibition by nitrile-based inhibitors: A computational study. *Front. Chem.* **2013**, *1*, 39. [[CrossRef](#)]
77. Batra, J.; Szabo, A.; Caulfield, T.R.; Soares, A.S.; Sahin-Tóth, M.; Radisky, E.S. Long-range Electrostatic Complementarity Governs Substrate Recognition by Human Chymotrypsin C, a Key Regulator of Digestive Enzyme Activation. *J. Biol. Chem.* **2013**, *288*, 9848–9859. [[CrossRef](#)]
78. Kumar, P.; Agarwal, P.K.; Waddell, M.B.; Mittag, T.; Serpersu, E.H.; Cuneo, M.J. Low-Barrier and Canonical Hydrogen Bonds Modulate Activity and Specificity of a Catalytic Triad. *Angew. Chem. Int. Ed.* **2019**, *58*, 16260–16266. [[CrossRef](#)] [[PubMed](#)]
79. Rut, W.; Groborz, K.; Zhang, L.; Sun, X.; Zmudzinski, M.; Pawlik, B.; Młynarski, W.; Hilgenfeld, R.; Drag, M. Substrate Specificity Profiling of SARS-CoV-2 Main Protease Enables Design of Activity-Based Probes for Patient-Sample Imaging. *bioRxiv* **2020**. [[CrossRef](#)]
80. Prejanò, M.; Romeo, I.; Russo, N.; Marino, T. On the Catalytic Activity of the Engineered Coiled-Coil Heptamer Mimicking the Hydrolase Enzymes: Insights from a Computational Study. *Int. J. Mol. Sci.* **2020**, *21*, 4551. [[CrossRef](#)] [[PubMed](#)]
81. Prejanò, M.; Marino, T.; Russo, N. On the Inhibition Mechanism of Glutathione Transferase P1 by Piperlongumine. Insight From Theory. *Front. Chem.* **2018**, *6*, 606. [[CrossRef](#)]
82. Prejanò, M.; Medina, F.E.; Ramos, M.J.; Russo, N.; Fernandes, P.; Marino, T. How the Destabilization of a Reaction Intermediate Affects Enzymatic Efficiency: The Case of Human Transketolase. *ACS Catal.* **2020**, *10*, 2872–2881. [[CrossRef](#)]
83. Piazzetta, P.; Marino, T.; Russo, N. Promiscuous Ability of Human Carbonic Anhydrase: QM and QM/MM Investigation of Carbon Dioxide and Carbodiimide Hydration. *Inorg. Chem.* **2014**, *53*, 3488–3493. [[CrossRef](#)]
84. Tolbatov, I.; Coletti, C.; Marrone, A.; Re, N. Insight into the Substitution Mechanism of Antitumor Au(I) N-Heterocyclic Carbene Complexes by Cysteine and Selenocysteine. *Inorg. Chem.* **2020**, *59*, 3312–3320. [[CrossRef](#)] [[PubMed](#)]
85. Nikoo, S.; Meister, P.J.; Hayward, J.J.; Gauld, J.W. An Assessment of Computational Methods for Calculating Accurate Structures and Energies of Bio-Relevant Polysulfur/Selenium-Containing Compounds. *Molecules* **2018**, *23*, 3323. [[CrossRef](#)]
86. Mladenovic, M.; Ansorg, K.; Fink, R.F.; Thiel, W.; Schirmeister, T.; Engels, B. Atomistic Insights into the Inhibition of Cysteine Proteases: First QM/MM Calculations Clarifying the Stereoselectivity of Epoxide-Based Inhibitors. *J. Phys. Chem. B* **2008**, *112*, 11798–11808. [[CrossRef](#)] [[PubMed](#)]
87. Liao, R.-Z.; Yu, J.-G.; Himo, F. Quantum Chemical Modeling of Enzymatic Reactions: The Case of Decarboxylation. *J. Chem. Theory Comput.* **2011**, *7*, 1494–1501. [[CrossRef](#)] [[PubMed](#)]
88. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789. [[CrossRef](#)]

89. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [[CrossRef](#)]
90. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [[CrossRef](#)]
91. Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chem.—Eur. J.* **2012**, *18*, 9955–9964. [[CrossRef](#)]
92. 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.; et al. *Gaussian 09*; Revision D.01; Gaussian, Inc.: Wallingford, CT, USA, 2013.
93. Yan, J.; Chen, S. How to Produce Methane Precursor in the Upper Ocean by an Untypical Non-Heme Fe-Dependent Methylphosphonate Synthase? *ChemPhysChem* **2020**, *21*, 385–396. [[CrossRef](#)]
94. Marenich, A.V.; Cramer, C.; Truhlar, D. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [[CrossRef](#)] [[PubMed](#)]
95. Zhao, Y.; Truhlar, D.G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Accounts* **2008**, *120*, 215–241. [[CrossRef](#)]
96. Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620. [[CrossRef](#)]
97. Zhao, Y.; Truhlar, D. Exploring the Limit of Accuracy of the Global Hybrid Meta Density Functional for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2008**, *4*, 1849–1868. [[CrossRef](#)] [[PubMed](#)]
98. Minenkov, Y.; Singstad, Å.; Occhipinti, G.; Jensen, V.R. The accuracy of DFT-optimized geometries of functional transition metal compounds: A validation study of catalysts for olefin metathesis and other reactions in the homogeneous phase. *Dalton Trans.* **2012**, *41*, 5526–5541. [[CrossRef](#)]
99. Qi, S.-C.; Hayashi, J.-I.; Zhang, L. Recent application of calculations of metal complexes based on density functional theory. *RSC Adv.* **2016**, *6*, 77375–77395. [[CrossRef](#)]
100. Mardirossian, N.; Head-Gordon, M. How Accurate Are the Minnesota Density Functionals for Noncovalent Interactions, Isomerization Energies, Thermochemistry, and Barrier Heights Involving Molecules Composed of Main-Group Elements? *J. Chem. Theory Comput.* **2016**, *12*, 4303–4325. [[CrossRef](#)]
101. Wang, Y.; Verma, P.; Jin, X.; Truhlar, D.G.; He, X. Revised M06 density functional for main-group and transition-metal chemistry. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 10257–10262. [[CrossRef](#)]
102. Huggins, D.J.; Biggin, P.C.; Dämgen, M.A.; Essex, J.W.; Harris, S.A.; Henchman, R.H.; Khalid, S.; Kuzmanic, A.; Laughton, C.A.; Michel, J.; et al. Biomolecular simulations: From dynamics and mechanisms to computational assays of biological activity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2019**, *9*, e1393. [[CrossRef](#)]
103. Glendenning, E.D.; Reed, A.E.; Carpenter, J.E.; Weinhold, F. *NBO*, version 3.1; Theoretical Chemistry Institute, University of Wisconsin: Madison, WI, USA, 1995.



**Paper IV**  
New Insights on the redox Properties of Viperin from Advanced Numerical Simulations



Article

# New Insights on the redox Properties of Viperin from Advanced Numerical Simulations

 Angela Parise <sup>1,2</sup>, Nino Russo <sup>2</sup>, Tiziana Marino<sup>2</sup>, Aurélien de la Lande<sup>1,\*</sup>

1 Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France  
 2 Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy  
 \* Correspondence: aurelien.de-la-lande@universite-paris-saclay.fr;

**Abstract:** The family of radical SAM enzymes plays a fascinating role in radical chemistry, and most of them catalyze the cleavage of the S-C bond of S-adenosylmethionine to initiate catalysis based on radical cascade processes. In the viperin enzyme, activation of these reactions is by electron transfer (ET) on 4Fe-4S cluster present in the active site. In this work we report a dynamics structural investigation of viperin through Molecular Dynamic simulations. We performed QM/MM calculations using constrained-DFT to clarify the favored spin configuration of the Fe-S cluster. All six possible antiferromagnetically coupled spin states ( $\alpha\alpha\beta\beta$ ,  $\alpha\beta\alpha\beta$ ,  $\alpha\beta\alpha$ ,  $\beta\beta\alpha\alpha$ ,  $\beta\alpha\beta\alpha$ ,  $\beta\alpha\alpha\beta$ ) were considered. The decoupled hypothesis QM + MM simplified scheme in the Marcus Theory framework was used to evaluate the ionization potential of the 4Fe-4S cofactor.

**Keywords:** Radical SAM enzyme; [4Fe-4S]; Spin configuration; QM/MM<sup>pol</sup>; Marcus Theory

## 1. Introduction

**Citation:** Lastname, F.; Lastname, F.; Lastname, F. Title. *Biomolecules* **2022**, *12*, x. <https://doi.org/10.3390/xxxxx>

Academic Editor: Firstname Lastname

Received: date

Accepted: date

Published: date

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

In cells, oxidation processes are catalysed by a number of enzymes and coenzymes that successively transfer reducing equivalents, either hydrogen atoms or electrons to redox partners. Some of the most studied redox metalloenzymes, involve in their catalytic cycle the formation of one or more of radical organic species. [1] The main enzyme superfamily related to radical reactions rely on two cofactors widely spread in all tree Kingdoms of life: iron-sulfur clusters and SAM (S-Adenosylmethionine). [2]

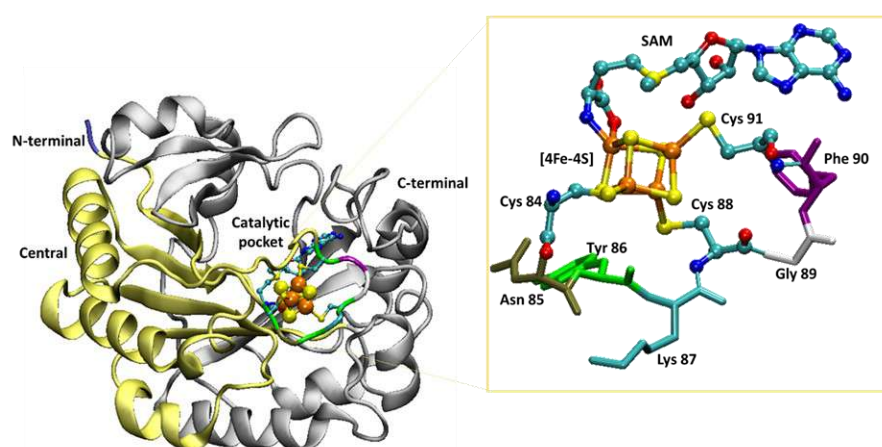
Viperin (standing for *Virus inhibitory protein, endoplasmic reticulum-associated, interferon-inducible*, also called RSAD2), is an enzyme found in humans and more generally mammals. It is a member of radical SAM superfamily of enzymes. [3] The first homologue of RSAD2, and the gene (*vig-1*) that encodes the protein, was identified in rainbow trout. [4] Then, came to light that viral infection also induce overexpression of *vig-1* gene in mouse. [5] These findings laid the foundation to get in 2001 that expression of RSAD2 in human cell lines, induced by interferons, inhibits the replication of cytomegalovirus. [6] Since then, viperin has been shown an effective antiviral activity, against RNA and DNA viruses, including HIV-1, West Nile virus, hepatitis C, dengue virus type 2, influenza A virus, and tick-borne encephalitis virus. [7–13]

Although viperin was discovered about 20 years ago, its 3D structure and its catalytic activity have been determined only recently. A recent study [14] revealed that viperin catalyses conversion reaction of cytidine triphosphate (CTP) to 3'-deoxy-3',4'-Didehydro-CTP (ddh-CTP) as illustrated in scheme S1. It is interesting to note that the catalytic product of viperin structurally resembles to antiviral drugs used in clinical treatment, which functions as novel antiviral nucleotide when mis-incorporated by viral RNA-dependent

RNA polymerases. [15,16] This aspect has attracted significant attention since viral infections continue to represent major challenges to public health.

Human viperin is a single polypeptide of 361 amino acids with a molecular weight of 42 kDa. [17] The N-terminal 42 amino acids forms amphipathic alpha-helix, which is needed for its localization into the endoplasmic reticulum membranes to inhibit protein secretion [18] and into lipid droplet to inhibit hepatitis C virus. [19] Overall, the N-terminal part consists of the first 70 amino acids.

The C-terminal 214–361 amino acids are highly conserved in different species and essential for pathogens inhibition. [20] The central domain contains the catalytic pocket that consists in a CxxxCxxC motif, characteristic of radical S-adenosylmethionine enzymes. [3] Indeed, the canonical CxxxCxxC motif (residues 83–90) permits to chelate three iron cations of the [4Fe-4S] cluster. The fourth iron cation within the cluster is liganded to the oxygen and nitrogen atoms of the SAM methionine fragment (Figure 1). [21]



**Figure 1.** Cartoon diagram of viperin where N-terminal domain is in light blue, central domain in yellow, C-terminal domain in grey. [4Fe-4S] cluster shown as orange and yellow spheres, S-Adenosylmethionine and cysteines bond to cluster are in ball and sticks, residues X of CxxxCxxC are in tube colored by residue. In the square zoom on viperin binding site. Hydrogen atoms are not shown for clarity. (PDB code 6Q2P) [22]

Multiple sequence alignment of active site show that vertebrate viperins are highly conserved and that fungi, bacteria, and archaebacteria express viperin-like enzymes. [16,21] The structural similarity among the radical SAM enzymes allows to put forward mechanistic hypotheses on catalytic reaction involving a single-electron reductive cleavage of SAM by the iron-sulphur cluster and generation of a highly reactive radical intermediate, 5'-deoxyadenosyl radical (dAdo). [14,23] In the presence of substrate, the radical abstracts a hydrogen atom from the substrate (CTP) and forms 5'-deoxyadenosine (dAdoH) as a by-product (see scheme S1). Many details of this reaction mechanism are still lacking though. Recent biochemical studies have shown that viperin binds to different triphosphate ribonucleotides [14], however, CTP shows the highest catalytic transformation efficiency. [22]

In this work, we focus on the structural characteristics controlling CTP binding with the active site. We report molecular dynamics simulations in its apo and CTP-bound forms, at physiological pH and temperature. We then focus on the redox properties of the Fe-S cluster. To this end, while the broken-symmetry density functional theory method is generally used to describe the various possible spin states of [4Fe-4S] cluster. [24,25] we propose an alternative approach based on constrained Density Functional Theory (cDFT) [26] to describe the antiferromagnetically-coupled spin configuration of [4Fe-4S] cluster in the



diabatic approximation. We then report QM + MM<sup>Pol</sup> calculation [27] to evaluate the redox potential of the cluster, taking into account the influence of protein environment.

## 2. Computational Details

Enzymes are biological catalysts whose dynamic behaviour is closely related to their reactivity. In Molecular Dynamics (MD) simulations, an initial model of the system is prepared from crystallographic structure, and different explicit solvation models is used to simulate these natural time-dependent conformational changing. MD utilizes classical mechanics approach based on Newton's laws to simulate atomic motion, where atoms are described as charged sphere interacting between them truth through covalent bond as springs and coulomb interaction. [28] This aspect allowed to determine the properties of big complex systems analytically in a reasonable computational time, but need a painstaking work in parameters setting, especially when non-standard residues and metals are present.

### 2.1 MD simulations set-up

We started our work with the recently reported crystal structure of mouse viperin bound to cytidine triphosphate and S-adenosylhomocysteine (SAH) cofactor with a resolution of 1.45 Å (PDB code 6Q2P). [22] To prepare a structure of the apo form, we removed the substrate and all the water molecules from the crystallographic structure. The chemically inactive SAH molecule was modified in silico by replacing the H atom bond to sulfur by a methyl group. We determined the protonation states on the amino acid residues at physiological pH using H++ server, [29] (see table S1), with a salt concentration of 0.15 M.

We have used the Amber16 software. [30] Force field parameters for the iron-sulfur cluster were taken from Ref. [31] The three cysteines are covalently linked to Fe-S cluster by using harmonic restrain in Amber to impose a Fe-S distance at  $2.2 \pm 0.2$  Å. The fourth iron cation interacts through a Lennard-Jones potential with the nitrogen and oxygen atoms of the amino and carboxylic group of SAM. FF parameters for S-adenosylmethionine were taken from. [32] Two chlorine counter ions were added to ensure electro neutrality.

We followed the same procedure for the enzyme-CTP complex except, of course, for cytidine triphosphate that was retained in the PDB file. Preparation of cytidine triphosphate substrate parameters was obtained by combining AMBER 94/99 force field data of monophosphate cytidine and parameters for the two terminal phosphate group of ATP available in AMBER parameter database. Two sodium counter ions were added to neutralize the enzyme-CTP system.

We inserted the protein within an orthorhombic box of 12 Å of TIP3P water molecules. The solvated structure was first minimized with harmonic restraints applied on all atoms of enzyme ( $50 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ ) using 5,000 steps of steepest descent algorithm, followed by 5,000 steps of conjugate gradient algorithm (CG). In the second minimization step, we released restraints on hydrogen atoms, with third and fourth minimizations being conducted with and without protein backbone atoms restraint, respectively. We carried out a progressive heating phase from 0 to 310 K for 20 ps using the Langevin thermostat in NVT ensemble, with a time step of 0.002 ps. The production phase complied to the following conditions: integration step of 2 fs coupling SHAKE algorithm, NPT ensemble at 1 bar pressure using the Berendsen barostat. The Particle Mesh Ewald summation method was employed for the electrostatic potential and the long-range electrostatic interactions were calculated with 12 Å cut-off distance. We carried out MD simulations for 150 ns for each the apo-enzyme (viperin) and the enzyme-substrate complex (CTP-viperin) The apo and complex system trajectories produced were saved every 0.2 ps and analyzed through the PTRAJ module. [33]

### 2.2 MD analysis



MD trajectories so obtained can then be used to assess the magnitude of structural changes (root-mean-square deviation [RMSD]), the propensity for a given residue or region to move (root-mean-square fluctuation [RMSF]), and the evolution of hydrogen bonding (HB) networks, among other things. The secondary structures were assigned with the DSSP (Dictionary of Secondary Structure of Proteins) algorithm. [34]

We carried out clustering analysis extract representative structures of the protein. RMSD-based clustering of the whole trajectories were carried out according to the relaxed complex scheme (RCS) protocol. After removing overall rotations and translations by RMS-fitting the  $C\alpha$  atoms' positions of the trajectory, the average linkage clustering algorithm is applied, implemented in cpptraj, identifying 10 representative conformations groups of the protein. Structures with a percentage of population > 15% have been structurally analysed.

In this study, the Protein-Ligand Interaction Profiler (PLIP) tool [35] was used to determine the non covalent interactions between CTP and viperin in each of the 3 most populated MD-clusters.

### 2.3 QM/MM calculations

We carried out hybrid QM/MM (Quantum Mechanics/Molecular Mechanics) geometry optimizations with the computational chemistry framework Cuby4 [36]. Cuby4 interfaces between Amber and deMon2k [37]. Cuby4 implements a subtractive QM/MM scheme with electrostatic embedding. [36] Hydrogen link-atom were added to handle QM/MM boundary when there covalent bonds are present over the QM and MM regions. For the QM methos, we have chosen the hybrid exchange correlation functional B3LYP [38] with the DZVP basis set [39] and GEN-A2\* [40,41] auxiliary basis set. The AMBER classic force field was used for the MM region.

After QM/MM geometry optimization, single-point calculations with PBE, [42] PBE0 [43,44] and B3LYP [38] functionals have been carried out to test different spin multiplicity states. To access the spin configuration within the cluster, we used constrained DFT.[45,46] In cDFT, *ad hoc* diabatic states in which it is possible to constraint the SCF procedure to converge toward an electron density that fulfills a spin charge constraint.[47,48] In this work the cDFT module implemented in deMon2k software (developer version 6.0.2) was used to predict the energetically favored spin configuration of the  $[4Fe-4S]^+$  cluster at the QM/MM level.[49] Files for calculations have been generated using QIB. [49] The QM regions used for this purpose are those shown in Figure S6 (a,b,d and f). Spin charge configurations shown in Table S5 were tested. The Onsager field [50] is applied to represent as a continuum the water solvent around the explicit MM water molecules.

### 2.4 Simplified QM+MM<sup>pol</sup> scheme

A simplified QM+MM scheme was used to calculate the redox potential of the Fe-S cluster witin [27] The approach assumes that the reorganization of the internal sphere upon redox change is mainly caused by local electronic and nuclear relaxation and is decoupled from the reorganization of the environment. This hypothesis has been confirmed in several redox enzymes.[51–53] The mechanical and electrostatic coupling between the redox cofactors and its environment is neglected. Inner-sphere and outer-sphere contributions are splitted and treated separately. The QM region consists of 4Fe-4S clusters, three coordinates cysteinates cut on the  $C\alpha$ - $C\beta$  bond, SAM carbonyl and amine groups connected by a  $-CH_2$  (as shown in Figure S6 b). The MM region includes the remaining protein moiety solvated by explicit water molecules.

Inner sphere contribution is calculated by DFT optimization of the QM region for the two redox states of the system, at B3LYP/DZVP basis set [39] and GEN-A2\* auxiliary basis stets level of theory. Out of sphere contribution is calculated at MM<sup>pol</sup> level by using FF02 [54] induced dipole model force field. 150 ns MD trajectory is sampled in 750 frames, for each frame the vertical  $\Delta E$  for the two redox states is evaluated by appropriately removing

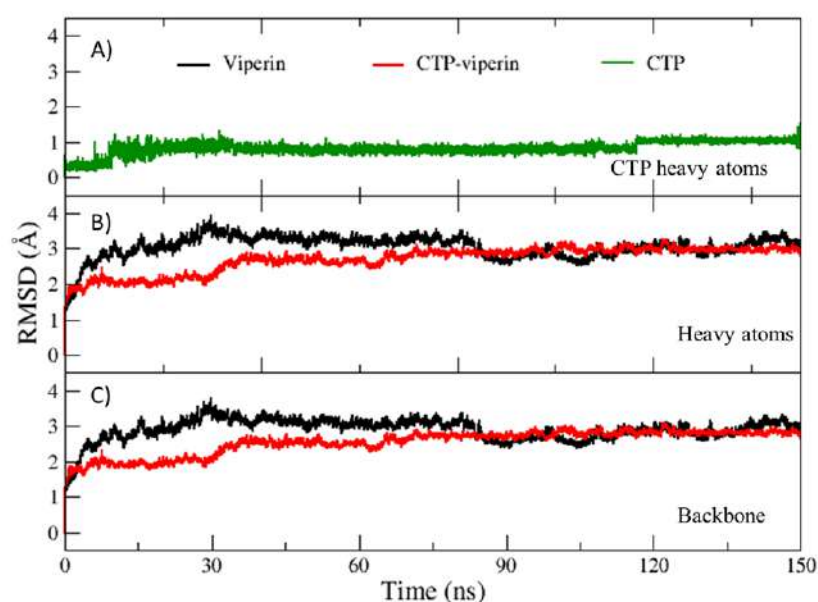
the MM energy contribution arising from the QM region. Parameters for the two redox states of the Fe-S cofactor come from DFT study.[31]

### 3. Results and Discussions

A representative collection of energetically accessible structures at physiological temperature and pH provides structural information of the enzyme and a statistical sample suitable to calculate the Ionization Potential of the 4Fe-4S cluster. In this effort classical Molecular Dynamics analyses have been carried out.

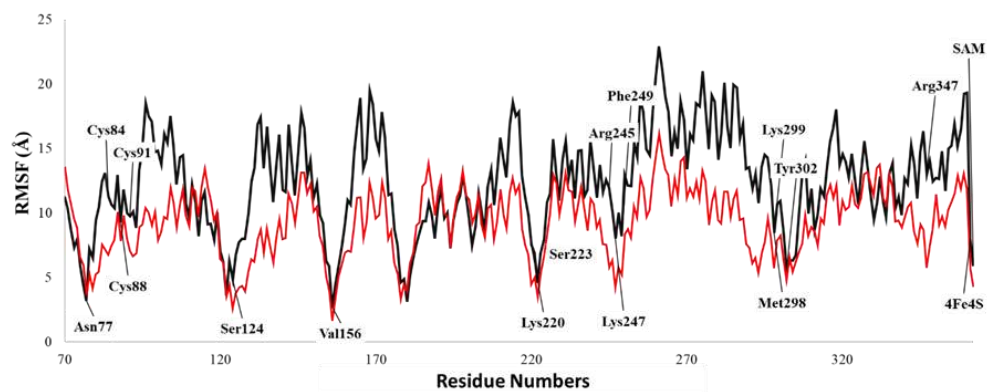
#### 3.1. MD analysis

We investigated the structural properties of the apo and ligand-bound system. The trajectory convergence during MD simulation was checked through RMSD values averaged over 150 ns MD simulation for backbone viperin and CTP-viperin, for viperin and CTP-viperin heavy atoms and ligand main atoms (figure 2).



**Figure 2.** 150 ns simulation RMSD values of CTP (in green) heavy atoms (panel A). RMSD values averaged over 150 ns runs of viperin (in black), CTP-viperin (in red) for the receptor heavy atoms (panel B) and backbone atoms (panel C).

The RMSD of backbone and heavy atoms are similar for the apo than for the complex system after a relaxation phase of 90 ns. In both systems, a stability at about 40ns after the beginning and remained stable throughout the simulation time for both viperin and CTP-viperin. To investigate the fluctuations of amino acid residues more in depth, we have calculated the root mean square fluctuations of the protein backbone, of [4Fe-4S] cluster, of SAM and of the substrate atoms during MD simulation (Figure 3).



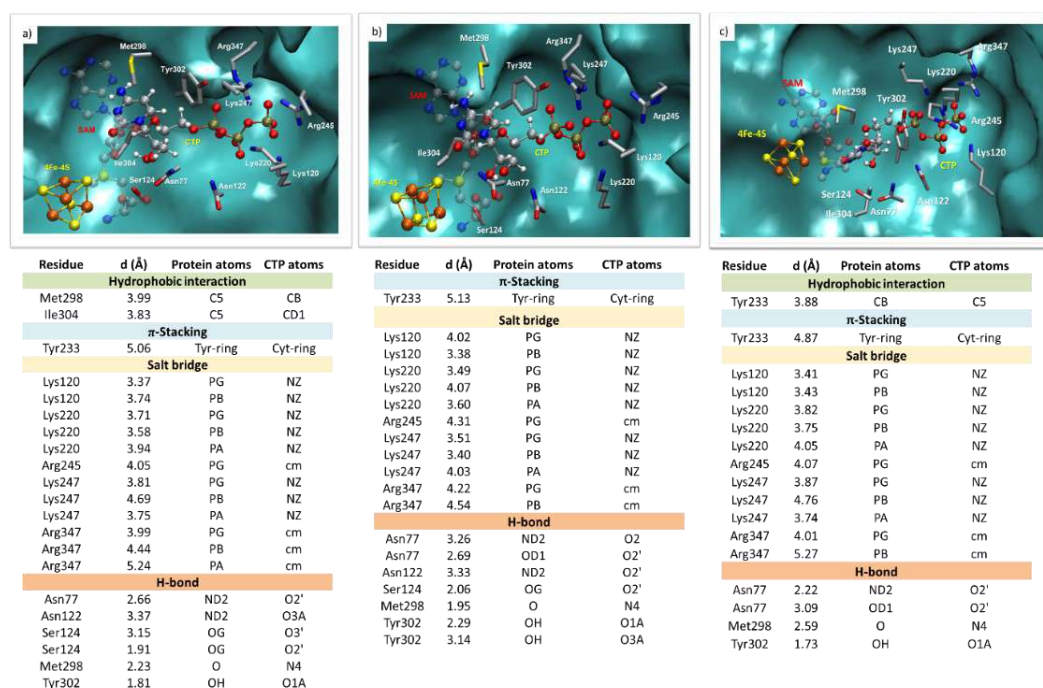
**Figure 3.** Root-Mean-Square-Fluctuations (RMSF) of the structures backbone (Å) with respect to residues of viperin in black and CTP-viperin in red. Some residues have been highlighted in the graph to facilitate discussion.

Figures S2-S3 shows that the most exposed to solvent part of the enzyme is highly more flexible than the internal part, that hosts the active site. As shown in figure 3 (black line), the higher peaks are observed at Lys168 of central domain, Gly260, Ala261, Asp262, Glu275, Glu278, Glu282, Lys285, Glu286 which correspond to alpha helix solvent-exposed of viperin C-terminal domain, Lys359 and Glu360 of the C-terminal extension concerning apo-viperin. The SAM cofactor shows a quite high mobility (8.1 Å) due to the fact that it is not covalently bound to the Fe-S Cluster and its position is stabilized only by neighboring amino acids weak interactions. Some of the most critical conformational changing that enzyme undergoes are those associated with substrate binding, figure 3 emphasizes this aspect. The average value of RMSF, in fact, decreases from 12.4 Å in the apo-form to 9.3 Å in the CTP bounded form. This trend is not only maintained in tree segments of the enzyme, each ranging from Leu112 to Leu118, from Ser180 to Val202 and from Gly329 to Phe336 (reported in Figure S4). Residues interacting with CTP (as reported in Table S2) such as Lys247, Tyr302, Lys220, Lys120, Arg347, His79, Arg245, Asn122, Asn222, Lys319, Met298, Asn77, Ser124, Cys314, Lys299 and Phe249 were observed to be stable.

Different progression of the RMSF curve in the presence or absence of the substrate is represented by a reorganization of the secondary structure of the enzyme. The secondary structures of both viperin and of CTP-viperin structures were analyzed by the DSSP algorithm, that defines secondary structure and geometrical features of proteins, given atomic coordinates along MD trajectory. Figure S5 shows that especially in the C-terminal region of the protein the presence of the substrate imposes a more ordered secondary structure. This is confirmed by the presence of weak interactions between CTP atoms and amino acid residues side chains (Tables S2-S3).

Phosphate groups establish a dense network of hydrogen bonds (Table S3) with amino acid residues pertaining to the C-terminal domain, more specifically the terminal phosphate group is engaged in hydrogen bonds with residues: Arg347, Arg245, Lys247 and Lys220; second phosphate group interacts with Lys120, Asn122 and Arg347; and the phosphate covalently bound to ribose forms H-bond with Tyr302, Lys247, Lys220 and Asn222. Intermolecular bond between Tyr302 and the first P<sub>A</sub>-group of the CTP is seen throughout the whole simulation time (frequency=1.0), candidating Tyr302 as a possible donor through proton-coupled electron transfer step (Scheme S1B) to the radical intermediate for product formation. In agreement with what has recently been studied by Ebrahimi *et al.* for a homologue of viperin, TtRSAD2, where the tyrosine (Tyr252) is the only residue present in the catalytic region arranged to mediate a redox reaction for the formation of

the final product observed by experiments, in which tyrosyl radical is observed by EPR spectroscopy. [55]



**Figure 4.** On top: representation of three most populated structures obtained by clustering analysis on CTP-viperin complex 150 ns MD simulation. Below non-covalent interactions (hydrophobic,  $\pi$ -stacking, salt bridge and hydrogen bond) between viperin and CTP obtained by Protein Ligand Interaction Profiler [35] related to each cluster, reported distances are expressed in Å and are related to distance between two atoms using as atom name as in PDB code 6Q2P, “cm” indicates the charged Arg residue terminal part center of mass. a) b) and c) refer to clusters with population of 23.5%, 18.3% and 17.3% respectively.

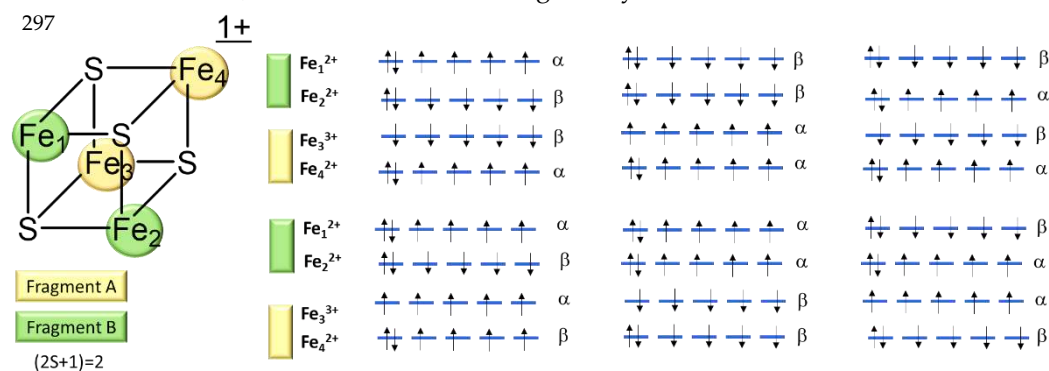
Detection of non-covalent interactions between substrate and the chemical species involved in the catalytic pocket shows some interesting elements. Figure 4 represents the three structures that more represent the MD simulation obtained from the clustering analysis (Tables S4-S5) and the detection of H-bonds,  $\pi$ -stacking, salt bridge and hydrophobic. The series of salt bridge, present in all three most populated cluster structure, between terminal phosphate-group and Lys120, Lys220, Arg245 and Arg347, between PB oxygen atoms and the side chain of Lys120, Lys220, Lys247 and Arg347 and between PA and Lys220, Lys247 and Arg347 give further confirmation of how the presence of CTP goes to structurally stabilize viperin.

The O2' hydroxyl group forms hydrogen bonds with Ser124 and Asn77 in all 3 analyzed clusters (figure 4), the hydrogen bond between O2'-hydroxyl group and NE2 atoms of His79 remarked in the crystallographic structure [22] is not maintained anymore along the MD simulation. Interestingly, electrostatic surface calculations performed on the representative MD snapshots (Figure S6) reveal the presence of a positively charged cavity, rich in arginines and lysines, this pocket acts as an electrostatic trap for the negatively charged nucleotides CTP.

### 3.2. Spin States of the reduced [4Fe-4S]<sup>+</sup> Cluster

The iron-sulfur cluster is one of the most ubiquitous cofactors in nature.[56] The [4Fe-4S]<sup>+</sup> state is the catalytically active oxidation state for the iron-sulfur cluster in the radical SAM enzymes. [57]

The four iron ions of the reduced  $[4\text{Fe-4S}]^+$  cluster are formally divided into two diiron fragments: a fragment in which the formal charges of the metal ions are respectively  $\text{Fe}^{3+}$   $\text{Fe}^{2+}$  and the other fragment with  $\text{Fe}^{2+}\text{Fe}^{2+}$ . We call the FA and FB respectively. In each FA and FB, the two high spin iron ions are ferromagnetically-coupled. The antiferromagnetic coupling of these two fragments ( $S = 9/2$  for FA and  $S = 4$  for FB) leads to a doublet spin multiplicity for the overall cluster. [58,59] (Table S6). There are six different spin states for the doublet state, as shown in Figure 5. Our investigation starts by testing for in cluster  $[4\text{Fe-4S}]^+$  different spin multiplicities. We start with the most representative of the viperin 300 ns MD simulation, optimized QM/MM using Cuby4 software [36]. A SCF (Self-Consistent-Procedure) minimization shows that spin density delocalizes on the sulfur atoms of cysteines that coordinate Fe ions (Table S8). This doesn't correspond to the spin configurations illustrated on Figure 6. To overcome this difficulty, constrained DFT have been used to access each of the spin configurations. More specifically, we imposed for each state, the expected spin charge on each of the four iron cations following the hirshfeld scheme. We have considered three possible QM/MM set-ups, in which the QM region encompasses different number of atoms (8, 60, 156). The results obtained show that for all the functionals used, the doublet state is energetically favored.



**Figure 5.** Six possible diabatic spin doublet states for iron sulfur cluster in its +1 reduced form, in which  $\alpha$  and  $\beta$  indicate spin up and down respectively.

The configurations shown in the figure 5 can be expressed in terms of spin charge as shown in Table S7. The relative energies to six possible antiferromagnetically-coupled spin state are reported in table 1.

**Table 1.** Relative energy in eV of the four iron ions spin states obtained by using cDFT, DZVP basis set, PBE and B3LYP functionals. Different size of QM-region (as shown in figure S6 a, d and e) were considered: only 4Fe-4S cluster (8 atoms), including first coordination shell of cluster (60 atoms) and considering second coordination shell (156 atoms).

QM-region	PBE			B3LYP		
	8 atoms	60 atoms	156 atoms	8 atoms	60 atoms	156 atoms
Spin	$\Delta E$ (eV)					
$\alpha\beta\alpha\beta$	0.18	<b>0.00</b>	<b>0.00</b>	0.15	<b>0.00</b>	<b>0.00</b>
$\alpha\alpha\beta\beta$	0.17	0.11	0.11	0.17	0.11	0.12
$\beta\alpha\alpha\beta$	0.10	0.09	0.11	0.05	0.16	0.16
$\alpha\beta\beta\alpha$	0.26	0.12	0.14	0.15	0.19	0.20
$\beta\beta\alpha\alpha$	<b>0.00</b>	0.15	0.25	<b>0.00</b>	0.17	0.17
$\beta\alpha\beta\alpha$	0.26	0.17	0.25	0.06	0.27	0.30



The ordering of the relative energy of the six different spin states are globally in agreement with each other considering the PBE and or hybrid B3LYP GGA exchange and correlation functionals, although in terms of absolute values there are differences.

For the smaller QM region, encompassing only the cubane, the state  $\alpha\alpha\beta\beta$  in the lower in energy, which is different from what is found with larger QM region. As expected this minimal model is not suitable for accurate evaluation of the electronic properties of the cluster. When the first coordination sphere ligands are included (medium-size model), the  $\alpha\beta\alpha\beta$  state is the most stable. That said, further inclusion of second sphere atoms modify the energy order for some spin-isomers.

Energy values of the six states suggest that all valence isomers are accessible, with the difference in isomers-interconversion energy between the states remaining less than 1 eV, in agreement with other results obtained by broken symmetry method. [25]

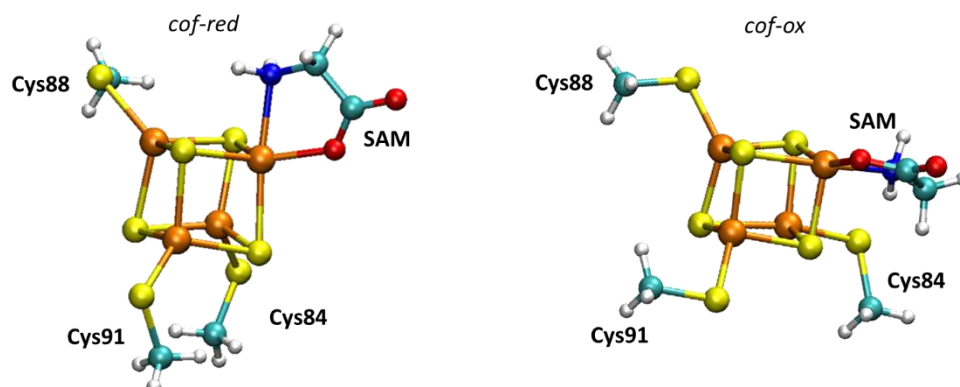
### 3.3. $[4Fe-4S]^+$ ionization potential by QM+MM<sup>pol</sup> scheme

The transition of the Fe-S cluster from the reduced  $[4Fe-4S]^+$  to the oxidized state  $[4Fe-4S]^{2+}$  is in close connection with electronic structure changes which in turn affects the reactivity. Such a change can explain the first step of viperin activation involving an electronic transfer (ET) from the cluster to the S-adenosylmethionine, as experiments highlight.[14,60] Redox potential (IP), in different chemical environments (e.g. vacuum, solution, protein environment) are probably the most relevant electronic properties to understand the involvement of iron ions in electron transfer, unlike other enzymes, the Fe-S cofactor is oxidized/reduced without proton transfer.[61,62] Accurate calculation of ionization potentials (IPs) is one of the most challenging problems in modern computational chemistry. [63]

In order to evaluate the 4Fe-4S cluster redox potential, including the effect of the protein environment, we rely on a QM + MM scheme [27] within Marcus Theory framework. [64,65] In Marcus Theory the collective reaction coordinate is the diabatic energy gap. QM+MM approach provides a separate treatment of inner-sphere contribution (redox cofactors) and out of sphere (protein environment). The QM+MM methodology assumes that the inner-sphere reorganization upon oxidation is decoupled from that of the environment.

The inner-sphere contribution is modelled from DFT geometry optimization of redox cofactor region. In our system the inner sphere contribution was calculated by preliminary QM/MM geometry optimization of the most representative MD structure, including in the QM region the 72 atoms model (Figure S6c), considering the reduced Fe-S cluster. QM/MM optimized structure, which we will call *c0-opt* (coordinate are available in Supporting Material). Starting from *c0-opt* we cropped a 32-atoms model (Figure S6b). Then we performed B3LYP/DZVP gas-phase single point the 32-atoms model, for iron-based cofactor its reduced form (what we will call *cof-red*) and B3LYP/DZVP gas-phase geometry optimization of cluster in its oxidized form (*cof-ox*). Structures of *cof-red* and *cof-ox* are shown in Figure 6. The IP energy of the inner-sphere contribution is equal to **4.3 eV**. This value is calculated as  $\Delta E$  coming from the difference between relative minimum energies of two redox states,  $E(\text{cof-ox})$  and  $E(\text{cof-red})$ .





**Figure 6.** Inner sphere structure optimized by DFT. *cof-red* refers to the  $[4\text{Fe-4S}]^+$  cluster cofactor in its reduced form, *cof-ox* refers to the  $[4\text{Fe-4S}]^{2+}$ . For both models, iron ligands are included.

The outer-sphere contribution is obtained from MD simulations of the solvated viperin. 150 ns MD simulations in the two cluster redox states have been carried out under the conditions described previously (in 2.1 MD simulations set-up). Force field parameters for the iron-sulfur cluster in the two redox states are used. [31]. The root-mean-square deviation of protein during the dynamics is shown in Figure S8., in which the PDB structures after energy minimization is taken as reference. Both RMSD within 4 Å, indicate dynamics are rather stable.

Each 150 ns trajectory (75000 frames) is sampled every 100 frames, for a total of 750 structures derived from MD in the reduced state and 750 structures (MD1-structures) derived from MD in the oxidized state (MD2-structures). For each snapshot belonging to the MD1-structures group two energies are calculated to obtain Marcus theory (MT) parameters. MT relies on the linear response approximation (LRA) which estimates  $\Delta G$  and  $\lambda$  from microscopic simulations by application of the two following equations: [49]

$$\Delta G^0 = \frac{\langle \Delta E \rangle_{1+} - \langle \Delta E \rangle_{2+}}{2} \quad (1)$$

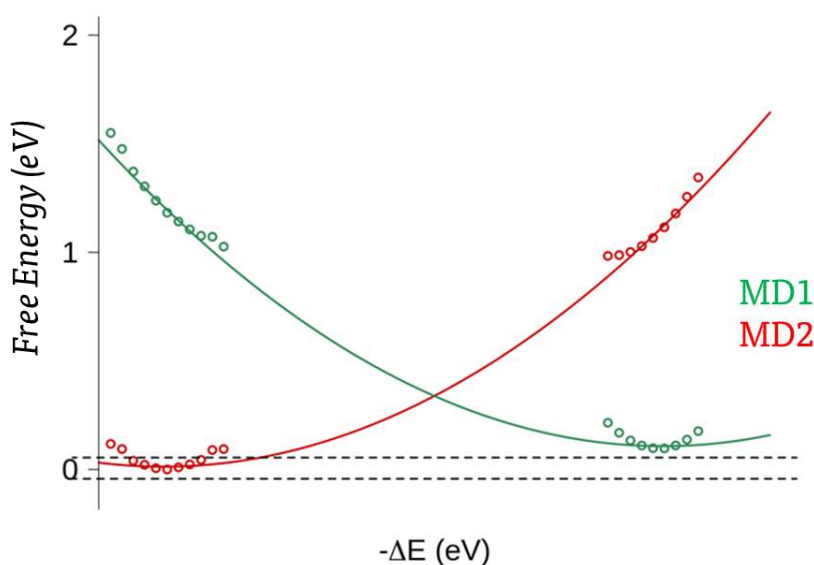
$$\lambda = \frac{\langle \Delta E \rangle_{1+} + \langle \Delta E \rangle_{2+}}{2} = \lambda^{\text{st}} \quad (2)$$

where  $\Delta E = E_2 - E_1$  is the vertical potential energy gap between the final and initial charge transfer states.  $\langle \Delta E \rangle_x$  denotes the thermal average of  $\Delta E$  for the system in electronic state  $x$ , where  $x$  can refer to +1 or +2 cluster charge.  $\lambda^{\text{st}}$  is the Stokes reorganization energy. Using  $\Delta E$  as the global reaction coordinate and defining the probability distribution  $p_x(\epsilon)$  to be the probability to have  $\epsilon = \Delta E$ , the free energy function for the system in each redox state is obtained by the Landau formula:  $g_x(\epsilon) = -\beta \ln(p_x(\epsilon)) + g_x^0$  ( $\beta = 1/k_B T$ ). [49].

MT parameters are calculated using Amberff02/POL3 FF. [54] Importantly the QM+MM approach can be made compatible with polarizable FF as long as the cofactor itself remains non-polarizable. Otherwise the decoupling assumption underlying the QM+MM idea breaks down.

**Table 2.** In the table are reported: PDB code of viperin, inner-sphere and outer-sphere reorganization energies and all MT parameters defined as in ref [66]. All values are in eV.

PDB code	$\Delta E_{in}$	$\Delta E_{out}$	$\lambda^{\text{st}}$	$\lambda_1^{\text{var}}$	$\lambda_2^{\text{var}}$	U
6Q2P	4.3	-0.002	1.56	0.34	0.59	0.0004



**Figure 7.** Diabatic free energy curves initial (green) and final (red) states obtained from the energy gap calculation along 150 ns of MD. The circles correspond to the points obtained directly from MD simulations sampling.

Free energy profile in the Marcus Theory where the collective reaction coordinate is taken to be the diabatic energy gap for viperin enzyme is reported in Figure 7. By including polarization in the force field we obtain the simulated inner sphere energy ( $\Delta E_{in}$ ), outer sphere energy ( $\Delta E_{out}$ ), reported in Table 2. We derive the theoretical ionization potential, that come from the sum of the two, IP=4.3 eV. This value is reasonable when compared to IP values for other enzymes containing Fe-S clusters. [67,68]

## 5. Conclusions

We have achieved 150 ns of MD simulations for viperin and CTP-viperin. The presence of the substrate in the catalytic pocket decreases the flexibility of the protein and causes rearrangement in the secondary structure. The catalytic pocket shows an high affinity for CTP by forming a rich network of hydrogen bonds and salt-bridges between receptor and substrate, maintained along the trajectory. Electron structure of the 4Fe-4S cofactor is studied by using constrain DFT. Considering all six possible anti-ferromagnetically coupled spin states, the  $\alpha\beta\alpha\beta$  state was found to be the most stable. We simulated the redox free energy in the framework of Marcus theory. QM+MM method was applied to evaluate the inner sphere and outer sphere contributions, including the effect of polarizable force field.

**Supplementary Materials:** The following supporting information can be downloaded at: [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1)

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Warshel, A.; Sharma, P.K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M.H.M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106*, 3210–3235, doi:10.1021/cr0503106.

2. Sofia, H.J. Radical SAM, a Novel Protein Superfamily Linking Unresolved Steps in Familiar Biosynthetic Pathways with Radical Mechanisms: Functional Characterization Using New Analysis and Information Visualization Methods. *Nucleic Acids Research* **2001**, *29*, 1097–1106, doi:10.1093/nar/29.5.1097. 424–426
3. Broderick, J.B.; Duffus, B.R.; Duschene, K.S.; Shepard, E.M. Radical S-Adenosylmethionine Enzymes. *Chem. Rev.* **2014**, *114*, 4229–4317, doi:10.1021/cr4004709. 427–428
4. Boudinot, P.; Massin, P.; Blanco, M.; Riffault, S.; Benmansour, A. Vig-1, a New Fish Gene Induced by the Rhabdovirus Glycoprotein, Has a Virus-Induced Homologue in Humans and Shares Conserved Motifs with the MoeA Family. *Journal of Virology* **1999**, *73*, 1846–1852, doi:10.1128/JVI.73.3.1846-1852.1999. 429–431
5. Boudinot, P.; Riffault, S.; Salhi, S.; Carrat, C.; Sedlik, C.; Mahmoudi, N.; Charley, B.; Benmansour, A. 2000 Vesicular Stomatitis Virus and Pseudorabies Virus Induce a Vig1/Cig5 Homologue in Mouse Dendritic Cells via Different Pathways. *Journal of General Virology* **81**, 2675–2682, doi:10.1099/0022-1317-81-11-2675. 432–434
6. Chin, K.-C.; Cresswell, P. Viperin (Cig5), an IFN-Inducible Antiviral Protein Directly Induced by Human Cytomegalovirus. *PNAS* **2001**, *98*, 15125–15130, doi:10.1073/pnas.011593298. 435–436
7. Van der Hoek, K.H.; Eyre, N.S.; Shue, B.; Khantisitthiporn, O.; Glab-Ampi, K.; Carr, J.M.; Gartner, M.J.; Jolly, L.A.; Thomas, P.Q.; Adikusuma, F.; et al. Viperin Is an Important Host Restriction Factor in Control of Zika Virus Infection. *Sci Rep* **2017**, *7*, 4475, doi:10.1038/s41598-017-04138-1. 437–439
8. Smith, D.B.; Vanek, J.; Ramalingam, S.; Johannessen, I.; Templeton, K.; Simmonds, P. Evolution of the Hepatitis E Virus Hypervariable Region. *J Gen Virol* **2012**, *93*, 2408–2418, doi:10.1099/vir.0.045351-0. 440–441
9. Li, W.; Mao, L.; Cao, Y.; Zhou, B.; Yang, L.; Han, L.; Hao, F.; Lin, T.; Zhang, W.; Jiang, J. Porcine Viperin Protein Inhibits the Replication of Classical Swine Fever Virus (CSFV) in Vitro. *Virology Journal* **2017**, *14*, 202, doi:10.1186/s12985-017-0868-4. 442–444
10. Jiang, D.; Weidner, J.M.; Qing, M.; Pan, X.-B.; Guo, H.; Xu, C.; Zhang, X.; Birk, A.; Chang, J.; Shi, P.-Y.; et al. Identification of Five Interferon-Induced Cellular Proteins That Inhibit West Nile Virus and Dengue Virus Infections. *J Virol* **2010**, *84*, 8332–8341, doi:10.1128/JVI.02199-09. 445–447
11. Carlton-Smith, C.; Elliott, R.M. Viperin, MTAP44, and Protein Kinase R Contribute to the Interferon-Induced Inhibition of Bunyamwera Orthobunyavirus Replication. *J Virol* **2012**, *86*, 11548–11557, doi:10.1128/JVI.01773-12. 448–449
12. Cao, J.; Forrest, J.C.; Zhang, X. A Screen of the NIH Clinical Collection Small Molecule Library Identifies Potential Anti-Coronavirus Drugs. *Antiviral Res* **2015**, *114*, 1–10, doi:10.1016/j.antiviral.2014.11.010. 450–451
13. Kurotaki, D.; Osato, N.; Nishiyama, A.; Yamamoto, M.; Ban, T.; Sato, H.; Nakabayashi, J.; Umehara, M.; Miyake, N.; Matsumoto, N.; et al. Essential Role of the IRF8-KLF4 Transcription Factor Cascade in Murine Monocyte Differentiation. *Blood* **2013**, *121*, 1839–1849, doi:10.1182/blood-2012-06-437863. 452–454
14. Gizzi, A.S.; Grove, T.L.; Arnold, J.J.; Jose, J.; Jangra, R.K.; Garforth, S.J.; Du, Q.; Cahill, S.M.; Dulyaninova, N.G.; Love, J.D.; et al. A Naturally Occurring Antiviral Ribonucleotide Encoded by the Human Genome. *Nature* **2018**, *558*, 610–614, doi:10.1038/s41586-018-0238-4. 455–457
15. Shim, J.; Larson, G.; Lai, V.; Naim, S.; Wu, J.Z. Canonical 3'-Deoxyribonucleotides as a Chain Terminator for HCV NS5B RNA-Dependent RNA Polymerase. *Antiviral Research* **2003**, *58*, 243–251, doi:10.1016/S0166-3542(03)00007-X. 458–459
16. Lachowicz, J.C.; Gizzi, A.S.; Almo, S.C.; Grove, T.L. Structural Insight into the Substrate Scope of Viperin and Viperin-like Enzymes from Three Domains of Life. *Biochemistry* **2021**, *60*, 2116–2129, doi:10.1021/acs.biochem.0c00958. 460–462
17. Jiang, D.; Guo, H.; Xu, C.; Chang, J.; Gu, B.; Wang, L.; Block, T.M.; Guo, J.-T. Identification of Three Interferon-Inducible Cellular Enzymes That Inhibit the Replication of Hepatitis C Virus. *Journal of Virology* **2008**, doi:10.1128/JVI.02113-07. 463–465

18. Hinson, E.R.; Cresswell, P. The N-Terminal Amphipathic  $\alpha$ -Helix of Viperin Mediates Localization to the Cytosolic Face of the Endoplasmic Reticulum and Inhibits Protein Secretion \*. *Journal of Biological Chemistry* **2009**, *284*, 4705–4712, doi:10.1074/jbc.M807261200.
19. Hinson, E.R.; Cresswell, P. The Antiviral Protein, Viperin, Localizes to Lipid Droplets via Its N-Terminal Amphipathic  $\alpha$ -Helix. *Proceedings of the National Academy of Sciences* **2009**, *106*, 20452–20457, doi:10.1073/pnas.0911679106.
20. Shaveta, G.; Shi, J.; Chow, V.T.K.; Song, J. Structural Characterization Reveals That Viperin Is a Radical S-Adenosyl-Methionine (SAM) Enzyme. *Biochemical and Biophysical Research Communications* **2010**, *391*, 1390–1395, doi:10.1016/j.bbrc.2009.12.070.
21. Fenwick, M.K.; Li, Y.; Cresswell, P.; Modis, Y.; Ealick, S.E. Structural Studies of Viperin, an Antiviral Radical SAM Enzyme. *Proc Natl Acad Sci USA* **2017**, 201705402, doi:10.1073/pnas.1705402114.
22. Fenwick, M.K.; Su, D.; Dong, M.; Lin, H.; Ealick, S.E. Structural Basis of the Substrate Selectivity of Viperin. *Biochemistry* **2020**, *59*, 652–662, doi:10.1021/acs.biochem.9b00741.
23. Broderick, W.E.; Hoffman, B.M.; Broderick, J.B. Mechanism of Radical Initiation in the Radical S-Adenosyl-Methionine Superfamily. *Acc Chem Res* **2018**, *51*, 2611–2619, doi:10.1021/acs.accounts.8b00356.
24. Blachly, P.G.; Sandala, G.M.; Giammona, D.A.; Bashford, D.; McCammon, J.A.; Noodleman, L. Broken-Symmetry DFT Computations for the Reaction Pathway of IspH, an Iron–Sulfur Enzyme in Pathogenic Bacteria. *Inorg. Chem.* **2015**, *54*, 6439–6461, doi:10.1021/acs.inorgchem.5b00751.
25. Zhou, S.; Wei, W.-J.; Liao, R.-Z. QM/MM Study of the Mechanism of the Noncanonical S-C $\gamma$  Bond Scission in S-Adenosylmethionine Catalyzed by the CmnDph2 Radical Enzyme. *Top Catal* **2021**, doi:10.1007/s11244-021-01420-5.
26. de la Lande, A.; Salahub, D.R. Derivation of Interpretative Models for Long Range Electron Transfer from Constrained Density Functional Theory. *Journal of Molecular Structure: THEOCHEM* **2010**, *943*, 115–120, doi:10.1016/j.theochem.2009.11.012.
27. Blumberger, J. Free Energies for Biological Electron Transfer from QM/MM Calculation: Method, Application and Critical Assessment. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5651–5667, doi:10.1039/B807444E.
28. Orozco, M.; Luque, F.J. Theoretical Methods for the Description of the Solvent Effect in Biomolecular Systems. *Chem. Rev.* **2000**, *100*, 4187–4226, doi:10.1021/cr990052a.
29. Anandakrishnan, R.; Aguilar, B.; Onufriev, A.V. H++ 3.0: Automating PK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res* **2012**, *40*, W537–541, doi:10.1093/nar/gks375.
30. 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.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017; University of California, San Francisco, 2017. 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.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.;

- Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017; University of California, San Francisco, 2017. 507  
508
31. Smith, D.M.A.; Xiong, Y.; Straatsma, T.P.; Rosso, K.M.; Squier, T.C. Force-Field Development and Molecular Dynamics of [NiFe] Hydrogenase. *J. Chem. Theory Comput.* **2012**, *8*, 2103–2114, doi:10.1021/ct300185u. 509  
510
32. Saez, D.A.; Vöhringer-Martinez, E. A Consistent S-Adenosylmethionine Force Field Improved by Dynamic Hirshfeld-I Atomic Charges for Biomolecular Simulation. *J. Comput. Aided Mol. Des.* **2015**, *29*, 951–961, doi:10.1007/s10822-015-9864-1. 511  
512  
513
33. Roe, D.R.; Cheatham, T.E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095, doi:10.1021/ct400341p. 514  
515
34. Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22*, 2577–2637, doi:10.1002/bip.360221211. 516  
517
35. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully Automated Protein–Ligand Interaction Profiler. *Nucleic Acids Research* **2015**, *43*, W443–W447, doi:10.1093/nar/gkv315. 518  
519
36. Řezáč, J. Cuby: An Integrative Framework for Computational Chemistry. *J. Comput. Chem.* **2016**, *37*, 1230–1237, doi:10.1002/jcc.24312. 520  
521
37. A.M. Koster, G. Geudtner, A. Alvarez-Ibarra, P. Calaminici, M.E. Casida, J. Carmona-Espindola, V.D. Dominguez, R. Flores-Moreno, G.U. Gamboa, A. Goursot, T. Heine, A. Ipatov, A. de La Lande, F. Janetzko, J.M. Del Campo, D. Mejia-Rodriguez, J. U. Reveles, J. Vasquez-Perez, A. Vela, B. Zuniga-Gutierrez, and D.R. Salahub, DeMon2k, Version 6, The DeMon Developers, Cinvestav, Mexico City (2018). 522  
523  
524  
525
38. Becke, A.D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652, doi:10.1063/1.464913. 526  
527
39. Godbout, N.; Salahub, D.R.; Andzelm, J.; Wimmer, E. Optimization of Gaussian-Type Basis Sets for Local Spin Density Functional Calculations. Part I. Boron through Neon, Optimization Technique and Validation. *Can. J. Chem.* **1992**, *70*, 560–571, doi:10.1139/v92-079. 528  
529  
530
40. Calaminici, P.; Janetzko, F.; Köster, A.M.; Mejia-Olvera, R.; Zuniga-Gutierrez, B. Density Functional Theory Optimized Basis Sets for Gradient Corrected Functionals: 3d Transition Metal Systems. *The Journal of Chemical Physics* **2007**, *126*, 044108, doi:10.1063/1.2431643. 531  
532  
533
41. Calaminici, P.; Flores–Moreno, R.; Köster, A.M. A Density Functional Study of Structures and Vibrations of Ta3O and Ta3O–. *Computing Letters* **2005**, *1*, 164–171, doi:10.1163/157404005776611420. 534  
535
42. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *PHYSICAL REVIEW LETTERS* **1996**, *77*, 4. 536  
537
43. Perdew, J.P.; Ernzerhof, M.; Burke, K. Rationale for Mixing Exact Exchange with Density Functional Approximations. *J. Chem. Phys.* **1996**, *105*, 9982–9985, doi:10.1063/1.472933. 538  
539
44. Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170, doi:10.1063/1.478522. 540  
541
45. Rudra, I.; Wu, Q.; Van Voorhis, T. Predicting Exchange Coupling Constants in Frustrated Molecular Magnets Using Density Functional Theory. *Inorg. Chem.* **2007**, *46*, 10539–10548, doi:10.1021/ic700871f. 542  
543
46. Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained Density Functional Theory. *Chem. Rev.* **2012**, *112*, 321–370, doi:10.1021/cr200148b. 544  
545
47. Wu, Q.; Van Voorhis, T. Direct Optimization Method to Study Constrained Systems within Density-Functional Theory. *Phys. Rev. A* **2005**, *72*, 024502, doi:10.1103/PhysRevA.72.024502. 546  
547



48. Yang, K.R.; Jalan, A.; Green, W.H.; Truhlar, D.G. Which Ab Initio Wave Function Methods Are Adequate for Quantitative Calculations of the Energies of Biradicals? The Performance of Coupled-Cluster and Multi-Reference Methods Along a Single-Bond Dissociation Coordinate. *J. Chem. Theory Comput.* **2013**, *9*, 418–431, doi:10.1021/ct3009528.
49. de la Lande, A.; Alvarez-Ibarra, A.; Hasnaoui, K.; Cailliez, F.; Wu, X.; Mineva, T.; Cuny, J.; Calaminici, P.; López-Sosa, L.; Geudtner, G.; et al. Molecular Simulations with In-DeMon2k QM/MM, a Tutorial-Review. *Molecules* **2019**, *24*, 1653, doi:10.3390/molecules24091653.
50. Mineva, T.; Russo, N.; Sicilia, E. Solvation Effects on Reaction Profiles by the Polarizable Continuum Model Coupled with the Gaussian Density Functional Method. *Journal of Computational Chemistry* **1998**, *19*, 290–299, doi:10.1002/(SICI)1096-987X(199802)19:3<290::AID-JCC3>3.0.CO;2-O.
51. Tipmanee, V.; Blumberger, J. Kinetics of the Terminal Electron Transfer Step in Cytochrome c Oxidase. *J. Phys. Chem. B* **2012**, *116*, 1876–1883, doi:10.1021/jp209175j.
52. Jiang, X.; Futera, Z.; Ali, Md.E.; Gajdos, F.; von Rudorff, G.F.; Carof, A.; Breuer, M.; Blumberger, J. Cysteine Linkages Accelerate Electron Flow through Tetra-Heme Protein STC. *J. Am. Chem. Soc.* **2017**, *139*, 17237–17240, doi:10.1021/jacs.7b08831.
53. Breuer, M.; Zarzycki, P.; Blumberger, J.; Rosso, K.M. Thermodynamics of Electron Flow in the Bacterial Deca-Heme Cytochrome MtrF. *J Am Chem Soc* **2012**, *134*, 9868–9871, doi:10.1021/ja3027696.
54. Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. Polarization Effects in Molecular Mechanical Force Fields. *J Phys Condens Matter* **2009**, *21*, 333102, doi:10.1088/0953-8984/21/33/333102.
55. Honarmand Ebrahimi, K.; Rowbotham, J.S.; McCullagh, J.; James, W.S. Mechanism of Diol Dehydration by a Promiscuous Radical-SAM Enzyme Homologue of the Antiviral Enzyme Viperin (RSAD2). *ChemBioChem* **2020**, *21*, 1605–1612, doi:10.1002/cbic.201900776.
56. Pandelia, M.-E.; Lanz, N.D.; Booker, S.J.; Krebs, C. Mössbauer Spectroscopy of Fe/S Proteins. *Biochimica et Biophysica Acta* **2015**, *1853*, 1395–1405.
57. Henshaw, T.F.; Cheek, J.; Broderick, J.B. The [4Fe-4S]<sub>1+</sub> Cluster of Pyruvate Formate-Lyase Activating Enzyme Generates the Glycyl Radical on Pyruvate Formate-Lyase: EPR-Detected Single Turnover. *J. Am. Chem. Soc.* **2000**, *122*, 8331–8332, doi:10.1021/ja002012q.
58. Jordanov, J.; Roth, E.K.H.; Fries, P.H.; Noodleman, L. Magnetic Studies of the High-Potential Protein Model [Fe<sub>4</sub>S<sub>4</sub>(S-2,4,6-(Iso-Pr)<sub>3</sub>C<sub>6</sub>H<sub>2</sub>)<sub>4</sub>]- in the [Fe<sub>4</sub>S<sub>4</sub>]<sup>3+</sup> Oxidized State. *Inorg. Chem.* **1990**, *29*, 4288–4292, doi:10.1021/ic00346a025.
59. NOODLEMAN, L. A Model for the Spin States of High-Potential [Fe<sub>4</sub>S<sub>4</sub>]<sup>3+</sup> Proteins. *Inorg. chem. (Print)* **1988**, *27*, 3677–3679.
60. Rivera-Serrano, E.E.; Gizzi, A.S.; Arnold, J.J.; Grove, T.L.; Almo, S.C.; Cameron, C.E. Viperin Reveals Its True Function. *Annual Review of Virology* **2020**, *7*, 421–446, doi:10.1146/annurev-virology-011720-095930.
61. Dempsey, J.L.; Winkler, J.R.; Gray, H.B. Proton-Coupled Electron Flow in Protein Redox Machines. *Chem Rev* **2010**, *110*, 7024–7039, doi:10.1021/cr100182b.
62. Wikström, M.; Jasaitis, A.; Backgren, C.; Puustinen, A.; Verkhovsky, M.I. The Role of the D- and K-Pathways of Proton Transfer in the Function of the Haem-Copper Oxidases. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2000**, *1459*, 514–520, doi:10.1016/S0005-2728(00)00191-2.
63. Bozkaya, U.; Ünal, A. State-of-the-Art Computations of Vertical Ionization Potentials with the Extended Koopmans' Theorem Integrated with the CCSD(T) Method. *J. Phys. Chem. A* **2018**, *122*, 4375–4380, doi:10.1021/acs.jpca.8b01851.



- 
64. Marcus, R.A. Chemical and Electrochemical Electron-Transfer Theory. *Annual Review of Physical Chemistry* **1964**, *15*, 155–196, doi:10.1146/annurev.pc.15.100164.001103. 589  
590
65. Marcus, R.A. Electron Transfer Reactions in Chemistry. Theory and Experiment. *Rev. Mod. Phys.* **1993**, *65*, 599–610, doi:10.1103/RevModPhys.65.599. 591  
592
66. Wu, X.; Hénin, J.; Baciou, L.; Baaden, M.; Cailliez, F.; de la Lande, A. Mechanistic Insights on Heme-to-Heme Transmembrane Electron Transfer Within NADPH Oxydases From Atomistic Simulations. *Frontiers in Chemistry* **2021**, *9*, 271, doi:10.3389/fchem.2021.650651. 593  
594  
595
67. Sweeney, W.V.; Rabinowitz, J.C. Proteins Containing 4Fe-4S Clusters: An Overview. *Annual Review of Biochemistry* **1980**, *49*, 139–161, doi:10.1146/annurev.bi.49.070180.001035. 596  
597
68. Perrin, B.S.; Niu, S.; Ichiye, T. Calculating Standard Reduction Potentials of [4Fe-4S] Proteins. *J Comput Chem* **2013**, *34*, 576–582, doi:10.1002/jcc.23169. 598  
599  
600  
601  
602  
603  
604  
605  
606  
607



**Work in progress**

The catalytic product of the natural antiviral viperin ddhCTP as inhibitor of SARS-CoV-2  
RNA-polymerase-RNA-dependent: a computational insight



# The catalytic product of the natural antiviral viperin, ddhCTP as possible inhibitor of SARS-CoV-2 RNA-polymerase-RNA-dependent: a computational insight

Angela Parise<sup>1,2</sup>, Giada Ciardullo<sup>1</sup>, Mario Pejano<sup>1</sup>, Aurélien de la Lande<sup>2</sup>, Tiziana Marino<sup>1\*</sup>

1 Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy

2 Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France

## Abstract

The novel coronavirus SARS-CoV-2 is the causative agent of the COVID-19 outbreak that has affected the entire planet. As the pandemic is still spreading worldwide, with multiple mutations of the virus, it is of interest and helpful the employment of computational methods for identifying potential inhibitors of the enzymes responsible for viral replication. Attractive antiviral nucleotide analogue RNA-dependent RNA polymerase (RdRp) chain terminator inhibitors are investigated with this purpose. This study aims to discover the binding mode of 3'-deoxy-3',4'-didehydro-cytidine triphosphate (ddhCTP) in complex with RdRp by molecular dynamics (MD) simulation. We built an *in silico* model of RNA strand embedded in RpRd using experimental starting from protein X-ray structure and exploiting the information obtained by spectrometry on the RNA sequence. We determined that the model was stable during the MD simulation time. We analyzed the affinity between ddhCTP and the active site of RdRp by comparison with its natural nucleobase cytidine triphosphate (CTP) substrate.

## Introduction

The world is currently in a state of pandemic emergency due to the spread of a disease called COVID-19 caused by a new viral etiologic agent, SARS CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2). Since the onset of Coronavirus Disease 2019. <sup>1</sup> SARS-CoV-2 is a positive RNA virus. Proliferation of SARS-CoV-2 requires RNA-dependent RNA polymerase (RdRp), a key enzyme that regulates replication and transcription of the viral genome and it is thus the validated target for the development of therapies against COVID-19 disease. The RdRp complex has multiple non-structural protein (nsp) units. The structure of RdRp consists of the nsp7-nsp8 heterodimer, nsp8 subunit, and nsp12 core catalytic unit. <sup>2</sup>

The nsp12 structure (see figure 1) contains a right-hand RdRp domain (residues Ser 367 to Phe 920) and a nidovirus-specific N-terminal extension domain (residues Asp 60 to Arg 249) that adopts a nidovirus RdRp-associated nucleotidyltransferase (NiRAN) architecture.<sup>3</sup> The polymerase domain and NiRAN domain are connected by an interface domain (residues Ala 250 to Arg 365). A further N-terminal  $\beta$  hairpin (residues Asp 29 to Lys 50) inserts into the cleavage blocked by the NiRAN domain and the palm subdomain in the RdRp domain. The polymerase domain adopts the conserved structure of the viral polymerase family<sup>4</sup> and consists of three subdomains: a finger subdomain (residues Leu 366 to Ala581 and Lys 621 to Gly 679), a palm subdomain (residues Thr 582 to Pro620 and Thr 680 to Gln 815), and a thumb subdomain (residues His 816 to Glu 920). The active site, contained in the palm domain, is formed by residues Ser 759, Asp 760 and Asp 761.<sup>5</sup> The carboxylate groups of these aspartates anchor a pair of divalent metal ions ( $Mg^{2+}$ ), which play the major role in catalysis,<sup>4</sup> while two other metal ions ( $Zn^{2+}$ ) play a structural role to stabilize the enzyme. One of the zinc ions is coordinated to amino acid residues His 295, Cys 301, Cys 306, and Cys 310 in the N-terminal domain; the second zinc ion is bonded to Cys 487, His 642, Cys 645, Cys 646 residues located in the finger domain.<sup>6</sup>

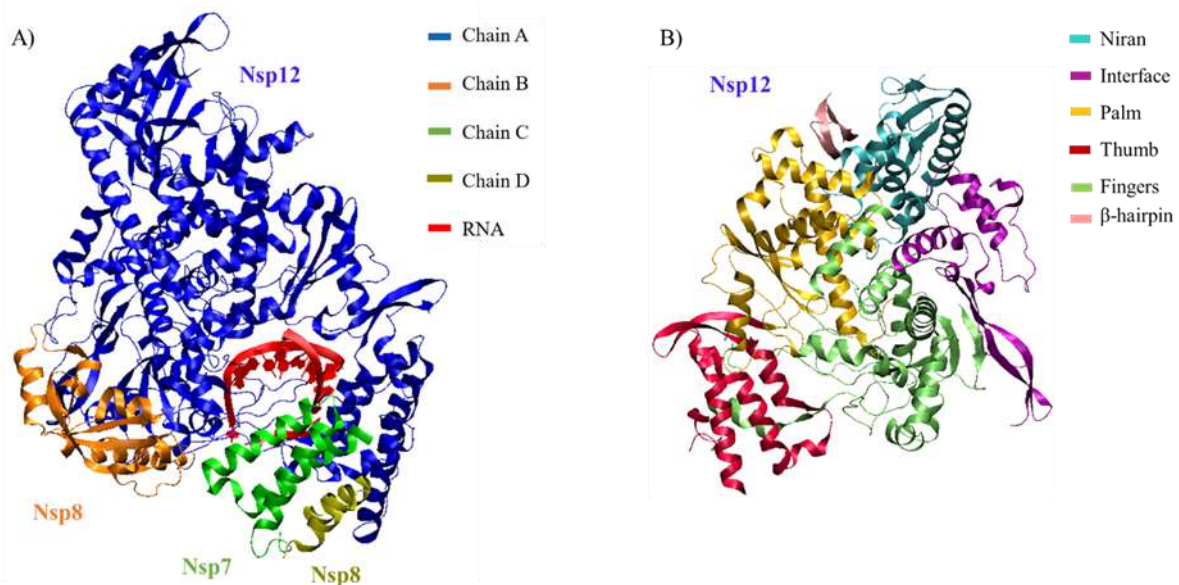


Figure 1 A) Structures of SARS-CoV-2 nsp7-nsp8-nsp12 complex. B) Structure of nsp12 domain organized by color.

RdRp is proposed as a target of a class of antiviral drugs that are nucleotide analogues; this category includes remdesivir<sup>7,8</sup> a nucleotide analogue adenosine mimic, approved by the FDA on August 10, 2020.<sup>9</sup> Several works have proposed remdesivir to induce delayed chain termination.<sup>10-13</sup>

3'-deoxy-3',4'-didehydro-cytidine triphosphate (ddhCTP) is a novel antiviral nucleotide-like compound produced by enzyme viperin as part of the innate immune response.<sup>14,15</sup> ddhCTP moreover, has recently been shown to be easy to synthesize.<sup>16</sup> Its effect as a chain terminator for RNA-dependent polymerases of multiple members of the Flavivirus genus has been shown<sup>17</sup> and recently it has been demonstrated SARS-CoV-2 polymerase to incorporate this cytosine analogue well.<sup>18</sup> In addition, the -OH group loss in 3' position, compared to its natural precursor (cytidine triphosphate, CTP),<sup>19</sup> makes the ddhCTP competitive with CTP at an intracellular concentration of ~100  $\mu$ M and lower than purine analogues.<sup>20</sup>

Encouraged by experimental evidence on ddhCTP<sup>18,21</sup> and the computational work on Remdesivir<sup>22</sup> an *in silico* study has been carried out on RdRp of SARS-CoV-2. Classical Molecular Dynamics (MD) simulations methods are used to study the RdRp inhibition promoted by ddhCTP. To better clarify the binding mode of ddhCTP on the active site, the behavior of the natural nucleotide cytidine triphosphate has been also investigated. Molecular docking has been used to test the affinity between polymerase active site and the examined ligands. MDs of 300 ns have been performed for the following systems: i) RdRp system that includes nsp7, nsp8, nsp12, a partial double stranded RNA represented by a primer nucleotide strand of 6 units and a template strand of 8 nucleotides (RdRp-RNA); ii) RdRp-RNA in complex with ddhCTP and iii) RdRp-RNA with CTP.

The choice of the investigation *in silico* of the endogenous product as antiviral drug represents a good strategy to develop drugs in more reduced times than those needed to obtain new therapeutic molecules from scratch.

### Computational Methods

The investigation started from a crystal complex of RNA polymerase nsp12-nsp7- nsp8 with a resolution of 2.50 Å (PDB code 7AAP) by cryo-EM<sup>23</sup> and with a partial bound template-primer RNA and RDV-TP. Contrarily to previous computational studies on the interaction between Remdesivir and SARS-CoV-2 RdRp that lacked of the RNA template-primer in their RdRp complex models,<sup>24-27</sup> crucial in the nucleotide triphosphate binding, our study included RNA template-primer similarly to Arba et al. work on the remdesivir.<sup>22</sup>

The crystallographic structure includes the subunits nsp7, nsp8 and the catalytic nsp12 bound to a template-primer RNA duplex in complex with favipiravir ribonucleoside triphosphate (favipiravir-RTP). Due to different nature (pyrimidine-like) of the ddhCTP from Favipiravir present in the crystallographic structure, a painstaking process was also required in the selection of the sequence of the RNA fragment to consider.



### A. *In silico* engineered protein

The Favipiravir inhibitor is a purine nucleic acid analogue derived from pyrazine carboxamide (6-fluoro-3-hydroxy-2-pyrazinecarboxamide).<sup>28</sup> Crystallographic RdRp incorporates the inhibitor into the nascent RNA chain. The design of RNA double strand length and the type of bases for cytosine analogue is required. The choice of RNA base sequence has been based from MALDI TOF experiments on CTP-like RpRd inhibitors.<sup>29</sup> The spectrometry study showed that base sequence for ddhCTP and CTP must be: UAAAAU (5' -> 3') and AUUUUAGU (3' -> 5') for primer and template, respectively. To obtain the desired model for computational investigation the phosphodiester bond backbone has been retained from the starting pdb structure (7AAP), while all the already present nucleobases have been removed and scrupulously replaced *in silico* as shown in Figure 2. RNA length has been reduced from 21 to 14 to reduce computational cost. Tleap tool in Amber16<sup>30</sup> has been used to get the final model (Figure 2C).

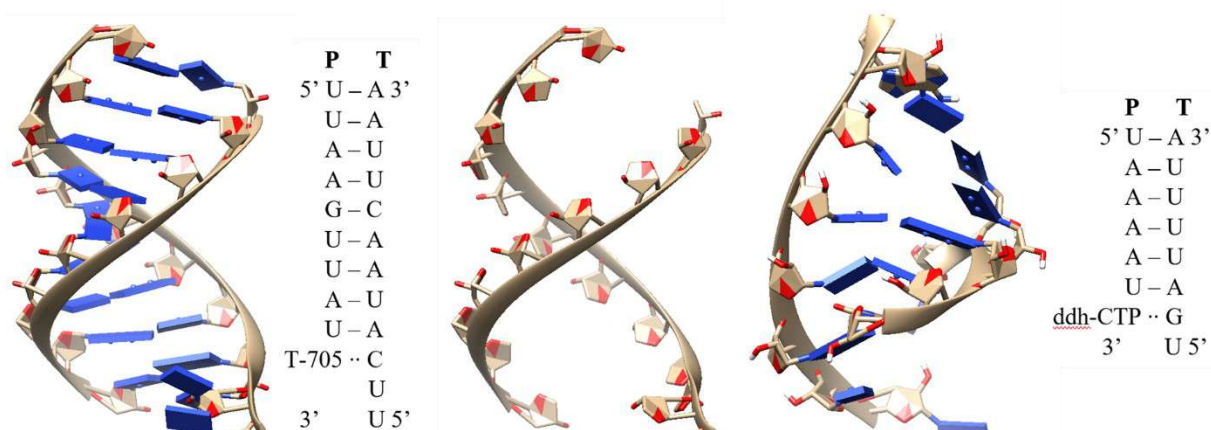


Figure 2 A) cryo-EM structure of favipiravir-RTP at the catalytic site of the SARS-CoV-2 RdRp, in complex with template:primer RNA and 2D representation of template (**T**) and primer (**P**) including T-705 (favipiravir), B) structure of favipiravir-RTP at the catalytic site of the RdRp in which only the phosphate-ribose scaffold is retained from the pdb 7AAP; C) computational model used for RdRp in complex with template:primer RNA and 2D representation of template (**T**) and primer (**P**) incorporating ddhCTP inhibitor.

### B. MD simulations setting

Three different MD simulations were performed, as follows: one 300 ns simulation for RdRp-RNA, 300 ns simulations for the ddhCTP-RdRp complex and 300 ns simulations for the CTP-RdRp complex, for a total of 0.9  $\mu$ s of simulations. The RdRp-RNA system was prepared at physiological pH using H<sup>+</sup> server,<sup>31</sup> with a salt concentration of 0.15 M. The protonated obtained system has been treated by using Amber 16 software.<sup>30</sup> The Amber force field FF14SB<sup>32</sup> and RNA.OL3<sup>33</sup> were used for protein and RNA, respectively. ZAFF force field,<sup>34</sup> specifically parameterized for zinc-containing systems in its commonly coordination in proteins, was selected for two Zn<sup>2+</sup> ions each of them coordinated to three

Cys and one His as above specified. The  $\text{Mg}^{2+}$  ions were treated with the Li et al. force field.<sup>35</sup> The system was solvated in a 12 Å orthorhombic box of water molecules TIP3P.<sup>36</sup> Counter ions (19  $\text{Na}^+$ ) were added appropriately to neutralize the RdRp-RNA system.

The parameters of triphosphate cytidine (CTP) were obtained by mixing AMBER 94/99 force field<sup>37</sup> data of mono-phosphate cytidine and parameters for the two terminal phosphate group of ATP available in AMBER parameter database.

In order to obtain ddhCTP parameters, gas phase geometry optimization has been carried out using B3LYP/6-31G\*. Atomic charges were derived by fitting the electrostatic potential according to the Merz–Singh–Kollman scheme,<sup>38</sup> using the RESP procedure. Antechamber and parmchk modules of Amber16 have been employed for generating preparatory files to perform MM relaxation of the complexes. 20  $\text{Na}^+$  counter ions were added to neutralize the system for both enzyme-RNA-ddhCTP and enzyme-RNA-CTP cases.

The solvated structures were first minimized by applying harmonic restraints on all atoms of enzyme (50 kcal mol<sup>-1</sup> Å<sup>2</sup>) using 5.000 steps of steepest descent algorithm, followed by 5.000 steps of conjugate gradient algorithm (CG). In the second minimization step, we released restraint on hydrogen atoms, with third and fourth minimizations being conducted with and without protein backbone atoms restraint, respectively. We carried out a progressive heating phase from 0 to 310 K for 200 ps using the Langevin thermostat in NVT ensemble, with a time step of 0.002 ps. The production phase of 300 ns for RdRp-RNA and 300 ns for ddhCTP-RdRp/CTP-RdRp complexes was performed under the following conditions: integration step of 2 fs coupling SHAKE algorithm, NPT ensemble at 1 bar pressure using the barostat.<sup>39</sup> The Particle Mesh Ewald summation method<sup>40</sup> was employed for the electrostatic potential and the long-range electrostatic interactions were calculated with 12 Å cut-off distance.

RdRp-RNA dynamic simulation of 300 ns was conducted on to ensure model stability. Trajectories for RdRp-RNA and its complexes with ddhCTP-RdRp and CTP-RdRp were saved every 0.2 ps and analyzed through the PTRAJ module.<sup>41</sup>

MD trajectories so obtained can then be used to assess the magnitude of structural changes (root-mean-square deviation [RMSD]), the propensity for a given residue or region to move (root-mean-square fluctuation [RMSF]), and the evolution of hydrogen bonding (HB) networks. The secondary structures were assigned with DSSP algorithm.<sup>42</sup>

Clustering analysis has been employed to provide a sampling energetically accessible conformational ensemble. RMSD-based clustering of the whole trajectories was performed according to the relaxed complex scheme (RCS) protocol.<sup>30</sup> After removing overall rotations and translations by RMS-fitting the C $\alpha$  atoms' positions of the trajectory, the average

linkage clustering algorithm was applied as implemented in ccpptraj, allowing to identify 10 representative conformations clusters of the protein. The most populated structures were used as a starting point for the docking procedure.

## Results and Discussions

The current work uses molecular dynamics (MD) simulations aimed to explore the conformational behaviour and to obtain atomistic details on the binding region of RpRd system in two different complexed structures, with ddhCTP and CTP (shown in Figure 3), as preliminary step before exploring, at QMMM level of theory, the mechanism for the nucleotides transfer reaction, in presence of CTP, and its corresponding inhibition mechanism promoted by ddhCTP. The mechanistic study is currently under investigation.

All trajectory files obtained from the MD simulations were examined for stability and

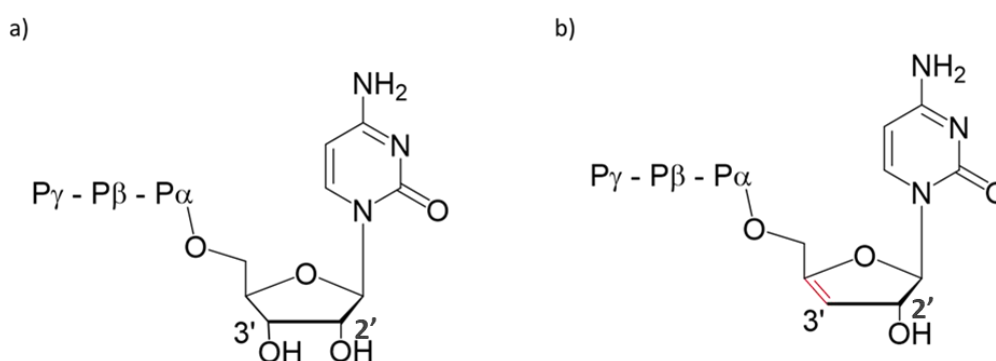


Figure 3. 2D representation of a) cytidine triphosphate (CTP) and b) 3'-deoxy-3',4'-dideohydro-cytidine triphosphate (ddhCTP). P refers to phosphate group.

fluctuation of the systems by monitoring the progress of RMSD and RMSF. The RMSF values, the H-bond interactions and the RMSD values calculated in 300 ns for side chain, C $\alpha$  and backbone of all three systems (RdRp, RdRp-ddhCTP and RdRp-CTP) are reported in Supporting Information (Figures S1-S3). In the present study, major attention will be focused on the comparative analysis of CTP substrate and ddhCTP inhibitor in the binding region such that the valuable information obtained could be further useful in the mechanistic investigation.

In Figure S1, the RMSD plots of the three examined systems related to the protein during the simulation time are reported. For RdRp-ddhCTP and RdRp-CTP the RMSD values converged after about 80 ns, indicating that the systems had reached a stable state. The RdRp system was stabilized with an RMSD value of proximal to 2.5 Å, while the corresponding complexed forms with nucleotide triphosphate had a slightly higher RMSD values, at 3 Å. The same behaviour was observed in the case of MD simulation of RdRp in comparison with that of the RdRp-RDV and RdRp-ATP systems.<sup>21</sup> This indicated that the binding of ddhCTP and CTP induced greater fluctuations in RdRp complex over the course of 300 ns of simulation. The

RMSD values of both ddhCTP and CTP along with those of RNA for all three systems, are reported in Figure S3. It is possible to evidence major deviations of RNA in the case of RdRp containing ddhCTP and CTP than in RdRp, whereas the trend of RMSD referred to ddhCTP and CTP remained constant during the whole simulation time. This finding confirmed their stability inside the catalytic pocket even if the substrate mimic of CTP, ddhCTP, resulted to be more stable than the natural substrate, CTP.

In Figure S4 are reported the RMSD values concerning the residues present in the nucleotide triphosphate (NTP) channel, important also to establish interactions with the negatively charged phosphate groups of ddhCTP and CTP and to potentially stabilize the transition state for pyrophosphate formation. This role is clear for Arg624 whose RMSD values for ddhCTP and CTP change compared with those of RdRp without ligand (line blue of Figure S4e). Similar trend occurs in Arg555 as already observed in the case of another inhibitor.<sup>22</sup> The binding of CTP and ddhCTP instead induced greater fluctuation of Lys798 and Arg553 residues. Asp618 and Asp623 of the NTP channel are in close proximity of magnesium ions which accounts for their consideration. Asp618 does not exhibit any particular deviation during the 300 ns of the three examined systems (Figure S4c), differently from the RMSD of Asp623, which appears rather changeable.

Trajectory clustering analysis applied to the 300 ns simulations of three investigated systems generated structural representatives of the most abundant clusters depicted in Figure S5. The most populated clusters had 45% occupancy, for CTP complex, and 23%, for ddhCTP one (Tables S2-S3).

#### *Active site*

A more detailed analysis was performed on the active site of RdRp to better evidence the different behaviour of ddhCTP and CTP already in the binding phase before the chemical reaction starts. In particular the parts involved in the reaction are: the inhibitor ddhCTP (the substrate CTP), the catalytic triad (Ser759, Asp760 and Asp761), the two magnesium ions 1003 and 1004 and primer uracil 3' terminal that will bind the inhibitor.

The structural stability of the active site of RdRp in presence of ddhCTP and CTP was evaluated by calculating the radius of gyration (Rg) shown in the Figure S6, as measure for the compactness of the structure. Its value appeared lower throughout the simulations of RdRp in presence of ddhCTP and CTP than those of RdRp indicating as the presence of ligands induces a stability inside the active site. This effect is more pronounced in the case of CTP. From the RMSD values plotted in Figure S7 emerges a discriminant behaviour of Ser759 that appears to be the most distant residue from the ddhCTP ligand compared with Asp760 and Asp761, while very similar trend of the three residues takes place for the CTP ligand.

In Figure S8 is presented the distance between the center of mass (COM) of the ddhCTP/CTP and that of each amino acid residue of catalytic triad. In analogy with RMSD above described, for the Ser759 in ddhCTP containing RdRp, the trend of the distance deviates more from that of the aspartate residues, thus suggesting as no H-bond of the residues occurs with the ddhCTP and enhancing the effect of the absence of -OH in position 3' (Figure 3) of the ribose ring. A different trend is observed in the case of CTP for which the COM of Ser759, but also of the Asp760 and Asp761, assumes smaller values as consequence of presence of the -OH on the site 3'.

Further investigation was carried out by calculating the radial distribution function (RDF) obtained as a function of the distance between the water and each active site residue (Ser759, Asp760, and Asp 761) and each magnesium ion (see Figure 4).

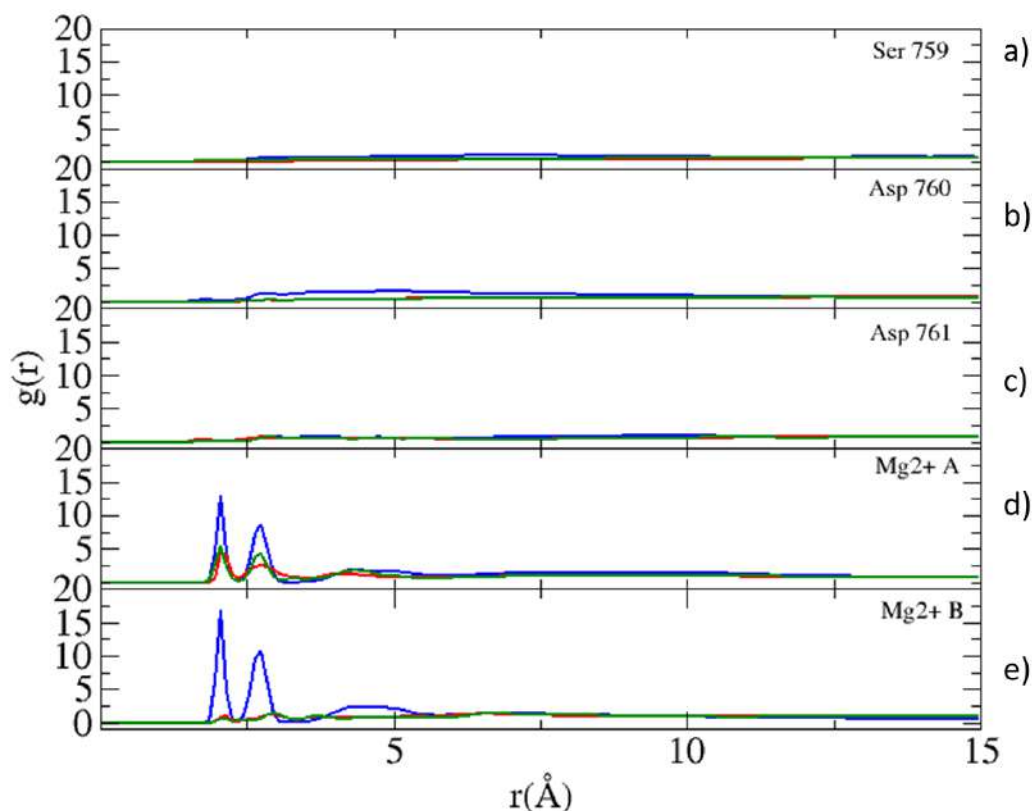


Figure4. Radial distribution function (RDF) of water molecules calculated for 300 ns of MD simulation of systems: RdRp-RNA in blue, RdRp-RNA-ddhCTP in green, RdRp-RNA-CTP in red respect to single residue of active site: a) Ser759, b) Asp760, c) Asp761, d)  $Mg^{2+}$  1003, e)  $Mg^{2+}$  1004.

From this figure it is worth to note that while for the three active site residues the RDF values remain almost the same in the absence and presence of the ligands, for the two  $Mg^{2+}$  ions the situation changes. In the case of Mg1 (1003) in absence of the inhibitor, two peaks relative to the first and second solvation shells within about 2.5 Å can be found. In presence of the inhibitor, these peaks are no longer found because no water molecules are present in the close vicinity of Mg1 (1003) due to the presence of ddhCTP or CTP. The same applies to Mg2 (1004), in fact, in presence of the inhibitor or substrate the first and the second

solvation spheres are not visible. This behaviour evidences as a water displacement from metallic centre must take place in presence of an inhibitor or a substrate.

A more detailed analysis, at atomic level, was performed through the calculation of the number of formed hydrogen bonds involving the protein and Ser759, Asp760 and Asp761 acting as acceptor in all three examined systems. Results are shown in Figure S9. The number of H-bonds involving Asp760 suffers a reduction for the presence of inhibitor/substrate with respect to the RdRp system (for which is equal to 3). Ser759 yet again reports what evidenced before in the case of analysis of RMSD and COM values for ddhCTP-RdRp complex, while for CTP three H-bonds are present. At the end Asp761 exhibits two H-bonds for CTP and only one for ddhCTP.

In Figure S10 the comparative SASA (solvent accessible surface area) profiles for all three systems are presented. From a first glance emerges as the presence of ddhCTP inhibitor induces a reduction of solvent accessible surface owing to lack of -OH moiety on position 3' of the ribose. This finding may prove to be crucial for RNA replication due to the possible role played by water molecules in the RdRp's catalytic mechanism.

The analysis of Mg-Mg distance, carried out by evaluating the COM occurring between the ions (see Figure S11), indicates that, in presence of both CTP and ddhCTP, the cations lie at shorter distance than that in RdRp. This is mainly related to the presence of phosphate moiety that, bridging the ions, imposes a different disposition of Mg1 and Mg2. For ddhCTP this distance assumes values around 3 Å while for CTP it reaches values less than 5 Å. This trend is in agreement with the shrinking of this distance proposed during the reaction.<sup>43</sup> This may lead O3' closer to P $\alpha$ , facilitating the nucleophilic attack. The behaviour of the Uracil 20 (primer 3'-terminal, see Figure 2) that is to bind the ddhCTP inhibitor or the CTP substrate was monitored during the whole simulation time by analysing the COM between the O3' of Uracil 20 and P $\alpha$  of ddhCTP (CTP) as shown in Figure S12. No difference from its trend emerged for CTP and ddhCTP ligands suggesting a similar orientation of their triphosphate groups during the simulation.



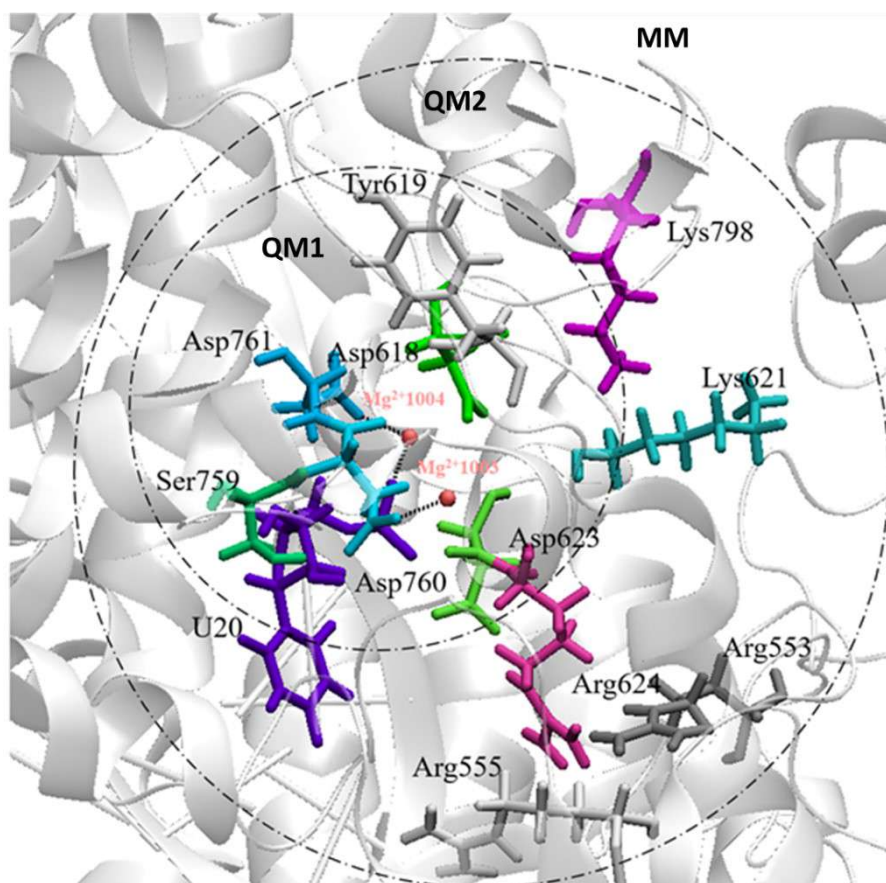


Figure 5. Representation of chosen model for next QM/MM calculations. Dashed circle highlight first (QM1) and second (QM2) metal ions coordination shell.

From the molecular dynamics analysis of the RdRp, RdRp-ddhCTP and RdR-CTP systems, information achieved can drive in the selection of the residues to be considered for the next quantum mechanical investigation of reaction mechanism. In Figure 5 are reported the chosen residues to be included in QM-region containing those closely adjacent to the active site (QM1), and those belonging to the second coordination shell (QM2), in particular positively charged Lys and Arg residues to preserve the electroneutrality of the model.

### Main remarks

This study mainly aims to find out the binding dynamics of ddhCTP to RdRp in comparison with CTP by means of molecular dynamics (MD) simulation.

Based on the presented results, it is important to underline that longer MD simulations will be required for an in-deep analysis of the conformational behaviour of the systems. Further investigation will be also additionally focused on the effect of the double stranded RNA length.

That said, ddhCTP demonstrated to be capable to generate a different behaviour from that of natural CTP nucleotide in proximity of active site that in turn can result fruitful for its inhibition action on RdRp of SARS-CoV-2.

The study of the inhibition mechanism of ddhCTP and of the reaction mechanism of CTP will provide deeper details that will allow to better rationalize the different behaviour of ddhCTP and CTP from both electronic and energetic points of view.

## References

- (1) Riva, L.; Yuan, S.; Yin, X.; Martin-Sancho, L.; Matsunaga, N.; Pache, L.; Burgstaller-Muehlbacher, S.; De Jesus, P. D.; Teriete, P.; Hull, M. V.; Chang, M. W.; Chan, J. F.-W.; Cao, J.; Poon, V. K.-M.; Herbert, K. M.; Cheng, K.; Nguyen, T.-T. H.; Rubanov, A.; Pu, Y.; Nguyen, C.; Choi, A.; Rathnasinghe, R.; Schotsaert, M.; Miorin, L.; Dejosez, M.; Zwaka, T. P.; Sit, K.-Y.; Martinez-Sobrido, L.; Liu, W.-C.; White, K. M.; Chapman, M. E.; Lendy, E. K.; Glynn, R. J.; Albrecht, R.; Rupp, E.; Mesecar, A. D.; Johnson, J. R.; Benner, C.; Sun, R.; Schultz, P. G.; Su, A. I.; García-Sastre, A.; Chatterjee, A. K.; Yuen, K.-Y.; Chanda, S. K. Discovery of SARS-CoV-2 Antiviral Drugs through Large-Scale Compound Repurposing. *Nature* **2020**, 586 (7827), 113–119. <https://doi.org/10.1038/s41586-020-2577-1>.
- (2) Jiang, Y.; Yin, W.; Xu, H. E. RNA-Dependent RNA Polymerase: Structure, Mechanism, and Drug Discovery for COVID-19. *Biochem Biophys Res Commun* **2021**, 538, 47–53. <https://doi.org/10.1016/j.bbrc.2020.08.116>.
- (3) Lehmann, K. C.; Gulyaeva, A.; Zevenhoven-Dobbe, J. C.; Janssen, G. M. C.; Ruben, M.; Overkleeft, H. S.; van Veelen, P. A.; Samborskiy, D. V.; Kravchenko, A. A.; Leontovich, A. M.; Sidorov, I. A.; Snijder, E. J.; Posthuma, C. C.; Gorbalenya, A. E. Discovery of an Essential Nucleotidylating Activity Associated with a Newly Delineated Conserved Domain in the RNA Polymerase-Containing Protein of All Nidoviruses. *Nucleic Acids Res* **2015**, 43 (17), 8416–8434. <https://doi.org/10.1093/nar/gkv838>.
- (4) McDonald, S. M. RNA Synthetic Mechanisms Employed by Diverse Families of RNA Viruses. *Wiley Interdiscip Rev RNA* **2013**, 4 (4), 351–367. <https://doi.org/10.1002/wrna.1164>.
- (5) Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; Ge, J.; Zheng, L.; Zhang, Y.; Wang, H.; Zhu, Y.; Zhu, C.; Hu, T.; Hua, T.; Zhang, B.; Yang, X.; Li, J.; Yang, H.; Liu, Z.; Xu, W.; Guddat, L. W.; Wang, Q.; Lou, Z.; Rao, Z. Structure of the RNA-Dependent RNA Polymerase from COVID-19 Virus. *Science* **2020**, 368 (6492), 779–782. <https://doi.org/10.1126/science.abb7498>.
- (6) Kirchdoerfer, R. N.; Ward, A. B. Structure of the SARS-CoV Nsp12 Polymerase Bound to Nsp7 and Nsp8 Co-Factors. *Nat Commun* **2019**, 10 (1), 2342. <https://doi.org/10.1038/s41467-019-10280-3>.
- (7) Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G. Remdesivir and Chloroquine Effectively Inhibit the Recently Emerged Novel Coronavirus (2019-NCoV) in Vitro. *Cell Res* **2020**, 30 (3), 269–271. <https://doi.org/10.1038/s41422-020-0282-0>.

- (8) Yin, W.; Mao, C.; Luan, X.; Shen, D.-D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; Chang, S.; Xie, Y.-C.; Tian, G.; Jiang, H.-W.; Tao, S.-C.; Shen, J.; Jiang, Y.; Jiang, H.; Xu, Y.; Zhang, S.; Zhang, Y.; Xu, H. E. Structural Basis for Inhibition of the RNA-Dependent RNA Polymerase from SARS-CoV-2 by Remdesivir. *Science* **2020**, 368 (6498), 1499–1504. <https://doi.org/10.1126/science.abc1560>.
- (9) Lamb, Y. N. Remdesivir: First Approval. *Drugs* **2020**, 80 (13), 1355–1363. <https://doi.org/10.1007/s40265-020-01378-w>.
- (10) Gordon, C. J.; Tchesnokov, E. P.; Feng, J. Y.; Porter, D. P.; Götte, M. The Antiviral Compound Remdesivir Potently Inhibits RNA-Dependent RNA Polymerase from Middle East Respiratory Syndrome Coronavirus. *J Biol Chem* **2020**, 295 (15), 4773–4779. <https://doi.org/10.1074/jbc.AC120.013056>.
- (11) Tchesnokov, E. P.; Obikhod, A.; Schinazi, R. F.; Götte, M. Delayed Chain Termination Protects the Anti-Hepatitis B Virus Drug Entecavir from Excision by HIV-1 Reverse Transcriptase. *J Biol Chem* **2008**, 283 (49), 34218–34228. <https://doi.org/10.1074/jbc.M806797200>.
- (12) Salgado-Benvindo, C.; Thaler, M.; Tas, A.; Ogando, N. S.; Bredenbeek, P. J.; Ninaber, D. K.; Wang, Y.; Hiemstra, P. S.; Snijder, E. J.; Hemert, M. J. van. Suramin Inhibits SARS-CoV-2 Infection in Cell Culture by Interfering with Early Steps of the Replication Cycle. *Antimicrobial Agents and Chemotherapy* **2020**. <https://doi.org/10.1128/AAC.00900-20>.
- (13) Ko, W.-C.; Rolain, J.-M.; Lee, N.-Y.; Chen, P.-L.; Huang, C.-T.; Lee, P.-I.; Hsueh, P.-R. Arguments in Favour of Remdesivir for Treating SARS-CoV-2 Infections. *Int J Antimicrob Agents* **2020**, 55 (4), 105933. <https://doi.org/10.1016/j.ijantimicag.2020.105933>.
- (14) Carlton-Smith, C.; Elliott, R. M. Viperin, MTAP44, and Protein Kinase R Contribute to the Interferon-Induced Inhibition of Bunyamwera Orthobunyavirus Replication. *J Virol* **2012**, 86 (21), 11548–11557. <https://doi.org/10.1128/JVI.01773-12>.
- (15) Chin, K.-C.; Cresswell, P. Viperin (Cig5), an IFN-Inducible Antiviral Protein Directly Induced by Human Cytomegalovirus. *PNAS* **2001**, 98 (26), 15125–15130. <https://doi.org/10.1073/pnas.011593298>.
- (16) Wood, J. M.; Evans, G. B.; Grove, T. L.; Almo, S. C.; Cameron, S. A.; Furneaux, R. H.; Harris, L. D. Chemical Synthesis of the Antiviral Nucleotide Analogue DdhCTP. *J Org Chem* **2021**, 86 (13), 8843–8850. <https://doi.org/10.1021/acs.joc.1c00761>.
- (17) Gizzi, A. S.; Grove, T. L.; Arnold, J. J.; Jose, J.; Jangra, R. K.; Garforth, S. J.; Du, Q.; Cahill, S. M.; Dulyaninova, N. G.; Love, J. D.; Chandran, K.; Bresnick, A. R.; Cameron, C. E.; Almo, S. C. A Naturally Occurring Antiviral Ribonucleotide Encoded by the Human Genome. *Nature* **2018**, 558 (7711), 610–614. <https://doi.org/10.1038/s41586-018-0238-4>.
- (18) Seifert, M.; Bera, S. C.; van Nies, P.; Kirchdoerfer, R. N.; Shannon, A.; Le, T.-T.-N.; Meng, X.; Xia, H.; Wood, J. M.; Harris, L. D.; Papini, F. S.; Arnold, J. J.; Almo, S.; Grove, T. L.;

- Shi, P.-Y.; Xiang, Y.; Canard, B.; Depken, M.; Cameron, C. E.; Dulin, D. Inhibition of SARS-CoV-2 Polymerase by Nucleotide Analogs from a Single-Molecule Perspective. *eLife* **2021**, *10*, e70968. <https://doi.org/10.7554/eLife.70968>.
- (19) Gordon, C. J.; Tchesnokov, E. P.; Woolner, E.; Perry, J. K.; Feng, J. Y.; Porter, D. P.; Götte, M. Remdesivir Is a Direct-Acting Antiviral That Inhibits RNA-Dependent RNA Polymerase from Severe Acute Respiratory Syndrome Coronavirus 2 with High Potency. *J Biol Chem* **2020**, *295* (20), 6785–6797. <https://doi.org/10.1074/jbc.RA120.013679>.
- (20) RECOVERY Collaborative Group; Horby, P.; Lim, W. S.; Emberson, J. R.; Mafham, M.; Bell, J. L.; Linsell, L.; Staplin, N.; Brightling, C.; Ustianowski, A.; Elmahi, E.; Prudon, B.; Green, C.; Felton, T.; Chadwick, D.; Rege, K.; Fegan, C.; Chappell, L. C.; Faust, S. N.; Jaki, T.; Jeffery, K.; Montgomery, A.; Rowan, K.; Juszczak, E.; Baillie, J. K.; Haynes, R.; Landray, M. J. Dexamethasone in Hospitalized Patients with Covid-19. *N Engl J Med* **2021**, *384* (8), 693–704. <https://doi.org/10.1056/NEJMoa2021436>.
- (21) Shi, W.; Ye, H.-Q.; Deng, C.-L.; Li, R.; Zhang, B.; Gong, P. A Nucleobase-Binding Pocket in a Viral RNA-Dependent RNA Polymerase Contributes to Elongation Complex Stability. *Nucleic Acids Research* **2020**, *48* (3), 1392–1405. <https://doi.org/10.1093/nar/gkz1170>.
- (22) Arba, M.; Wahyudi, S. T.; Brunt, D. J.; Paradis, N.; Wu, C. Mechanistic Insight on the Remdesivir Binding to RNA-Dependent RNA Polymerase (RdRp) of SARS-Cov-2. *Comput Biol Med* **2021**, *129*, 104156. <https://doi.org/10.1016/j.combiomed.2020.104156>.
- (23) Naydenova, K.; Muir, K. W.; Wu, L.-F.; Zhang, Z.; Coscia, F.; Peet, M. J.; Castro-Hartmann, P.; Qian, P.; Sader, K.; Dent, K.; Kimanius, D.; Sutherland, J. D.; Löwe, J.; Barford, D.; Russo, C. J. Structure of the SARS-CoV-2 RNA-Dependent RNA Polymerase in the Presence of Favipiravir-RTP. *PNAS* **2021**, *118* (7). <https://doi.org/10.1073/pnas.2021946118>.
- (24) Elfiky, A. A. SARS-CoV-2 RNA Dependent RNA Polymerase (RdRp) Targeting: An in Silico Perspective. *J Biomol Struct Dyn* **2021**, *39* (9), 3204–3212. <https://doi.org/10.1080/07391102.2020.1761882>.
- (25) Elfiky, A. A. Anti-HCV, Nucleotide Inhibitors, Repurposing against COVID-19. *Life Sci* **2020**, *248*, 117477. <https://doi.org/10.1016/j.lfs.2020.117477>.
- (26) Kato, K.; Honma, T.; Fukuzawa, K. Intermolecular Interaction among Remdesivir, RNA and RNA-Dependent RNA Polymerase of SARS-CoV-2 Analyzed by Fragment Molecular Orbital Calculation. *J Mol Graph Model* **2020**, *100*, 107695. <https://doi.org/10.1016/j.jmglm.2020.107695>.
- (27) Zhang, L.; Zhou, R. Structural Basis of the Potential Binding Mechanism of Remdesivir to SARS-CoV-2 RNA-Dependent RNA Polymerase. *J. Phys. Chem. B* **2020**, *124* (32), 6955–6962. <https://doi.org/10.1021/acs.jpccb.0c04198>.

- (28) Shannon, A.; Selisko, B.; Le, N.-T.-T.; Huchting, J.; Touret, F.; Piorkowski, G.; Fattorini, V.; Ferron, F.; Decroly, E.; Meier, C.; Coutard, B.; Peersen, O.; Canard, B. Rapid Incorporation of Favipiravir by the Fast and Permissive Viral RNA Polymerase Complex Results in SARS-CoV-2 Lethal Mutagenesis. *Nat Commun* **2020**, *11* (1), 4682. <https://doi.org/10.1038/s41467-020-18463-z>.
- (29) Chien, M.; Anderson, T. K.; Jockusch, S.; Tao, C.; Li, X.; Kumar, S.; Russo, J. J.; Kirchdoerfer, R. N.; Ju, J. Nucleotide Analogues as Inhibitors of SARS-CoV-2 Polymerase, a Key Drug Target for COVID-19. *J. Proteome Res.* **2020**, *19* (11), 4690–4697. <https://doi.org/10.1021/acs.jproteome.0c00392>.
- (30) 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.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017; University of California, San Francisco, 2017. 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.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017; University of California, San Francisco, **2017**.
- (31) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating PK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res* **2012**, *40*, W537-541. <https://doi.org/10.1093/nar/gks375>.
- (32) 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*, 8, 3696–3713 <https://pubs.acs.org/doi/full/10.1021/acs.jctc.5b00255>.
- (33) Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E.; Jurečka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7* (9), 2886–2902. <https://doi.org/10.1021/ct200162x>.
- (34) Yu, Z.; Li, P.; Merz, K. M. Extended Zinc AMBER Force Field (EZAFF). *J. Chem. Theory Comput.* **2018**, *14* (1), 242–254. <https://doi.org/10.1021/acs.jctc.7b00773>.



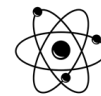
- (35) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, *9* (6), 2733–2748. <https://doi.org/10.1021/ct400146w>.
- (36) Vassetzki, D.; Pagliai, M.; Procacci, P. Assessment of GAFF2 and OPLS-AA General Force Fields in Combination with the Water Models TIP3P, SPCE, and OPC3 for the Solvation Free Energy of Druglike Organic Molecules. *J. Chem. Theory Comput.* **2019**, *15* (3), 1983–1995. <https://doi.org/10.1021/acs.jctc.8b01039>.
- (37) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters, *Proteins*, **2006**, *65*(3), 712-25. <https://pubmed.ncbi.nlm.nih.gov/16981200>.
- (38) 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.* **1993**, *97* (40), 10269–10280. <https://doi.org/10.1021/j100142a004>.
- (39) Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. van; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *The Journal of Chemical Physics* **1998**, *81* (8), 3684. <https://doi.org/10.1063/1.448118>.
- (40) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An  $N \cdot \log(N)$  Method for Ewald Sums in Large Systems. *The Journal of Chemical Physics* **1993**, *98* (12), 10089–10092. <https://doi.org/10.1063/1.464397>.
- (41) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095. <https://doi.org/10.1021/ct400341p>.
- (42) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577–2637. <https://doi.org/10.1002/bip.360221211>.
- (43) Yang, W.; Lee, J. Y.; Nowotny, M. Making and Breaking Nucleic Acids: Two-Mg<sup>2+</sup>-Ion Catalysis and Substrate Specificity. *Mol Cell* **2006**, *22* (1), 5–13. <https://doi.org/10.1016/j.molcel.2006.03.013>.



## Section B

### **Paper V**

The antioxidant capability of higenamine: Insights from theory





Article

# The Antioxidant Capability of Higenamine: Insights from Theory

Isabella Romeo <sup>1</sup>, Angela Parise <sup>1</sup>, Annia Galano <sup>2</sup>, Nino Russo <sup>1,\*</sup>,  
Juan Raúl Alvarez-Idaboy <sup>3</sup> and Tiziana Marino <sup>1,\*</sup>

<sup>1</sup> Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Arcavacata di Rende, Italy; isabella.romeo@unical.it (I.R.); angela.parise@unical.it (A.P.)

<sup>2</sup> Departamento de Química, Universidad Autónoma Metropolitana-Iztapalapa, Ciudad de México 09340, Mexico; annia.galano@gmail.com

<sup>3</sup> Facultad de Química, Departamento de Física y Química Teórica, Universidad Nacional Autónoma de México, Ciudad de Mexico 04510, Mexico; jdaboy@unam.mx

\* Correspondence: nrusso@unical.it (N.R.); tmarino@unical.it (T.M.)

Received: 20 March 2020; Accepted: 23 April 2020; Published: 25 April 2020



**Abstract:** Density functional theory was employed to highlight the antioxidant working mechanism of higenamine in aqueous and lipid-like environments. Different reaction mechanisms were considered for the reaction of higenamine with the  $\bullet\text{OOH}$  radical. The pH values and the molar fraction at physiological pH were determined in aqueous solution. The results show that the preferred reaction mechanism was the hydrogen atom transfer from the catecholic ring. The computed kinetic constants revealed that, in order to obtain reliable results, it is important to consider all the species present in water solution derived from acid–base equilibria. From the present investigation, it emerges that at physiological pH (7.4), the scavenging activity of higenamine against the  $\bullet\text{OOH}$  radical is higher than that of Trolox, chosen as a reference antioxidant. Furthermore, higenamine results to be more efficient for that purpose than melatonin and caffeine, whose protective action against oxidative stress is frequently associated with their reactive oxygen species (ROS) scavenging activity.

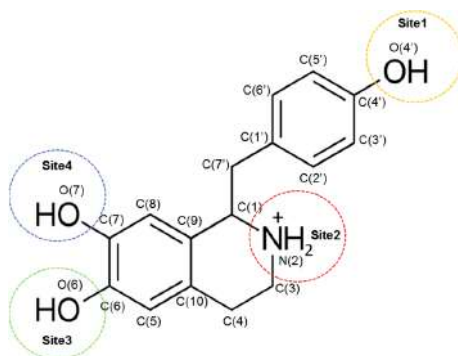
**Keywords:** DFT; antioxidant mechanism; kinetic constants; higenamine; acid–base equilibria

## 1. Introduction

Higenamine [1-(4'-hydroxybenzyl)-6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline] (Figure 1), also known as norcoclaurine or dl-demethylcoclaurine, is a plant-based alkaloid belonging to the structural class of protoberberines. It is present in many plants such as *Aconitum japonicum*, *Nandina Domestica*, *Gnetum Parvifolium*, *Asarum heterotropoides*, *Nelumbo nucifera*, *Galium divaricatum*, *Annona squamosa*, and *Aconitum carmichaelii* [1,2].

For 40 years, higenamine has been known as a cardiostimulant due to its beta-agonist activity and inotropic and chronotropic properties [3]. More recently, a revived interest in this compound has been motivated by its possible applications in many other therapeutic fields. Indeed, several studies have highlighted that higenamine exerts a hypotensive effect (it is a  $\alpha_1$ -adrenergic receptor antagonist) [4] and a protective effect on ischemia/reperfusion injuries (it activates the Phosphatidylinositol-3-kinase–Protein kinase B also known as Akt (PI3K/AKT) pathway) [5] and, more recently, it has been proposed as pharmacological stress agent for myocardial perfusion imaging. In addition to these properties, higenamine exhibits pharmacological activity towards other diseases, such as sepsis, heart failure, breathing difficulties, erectile dysfunction (ED), bradyarrhythmia, arthritis, and disseminated intravascular coagulation [1]. This multi-target activity led researchers to focus on the mechanisms and on the pathways implicated in higenamine action in different diseases.

Recent literature [6–10] indicates that the common denominator could be its antioxidant activity against reactive oxygen species (ROS). These highly reactive species may originate endogenously and exogenously, such as from metabolic pathways and external influences (e.g., smoking, radiation, drugs, and other environmental contaminants). Although under normal conditions the organism is able to maintain a good balance between production and removal of free radicals, their overproduction leads to oxidative stress, necrosis, apoptosis, and damage to biological macromolecules and compromises homeostasis and cellular function [11–13].



**Figure 1.** The 2D structure of higenamine.

Antioxidant compounds act by detoxifying and reducing ROS and controlling oxidative stress through different reaction mechanisms, such as electron transfer (ET), proton transfer (PT), sequential proton loss electron transfer (SPLET), hydrogen atom transfer (HAT) and radical adduct formation (RAF). Considering that natural products offer a wide range of antioxidant compounds, the identification of the specific role of each chemical portion of these compounds and the associated reaction mechanism is paramount in order to overcome oxidative stress in a targeted manner [14,15].

The presence of an OH group with phenolic nature proves to be essential for these properties, particularly, in the alkoxy, carboxyl, ester, and carbonyl groups in which the O atom provides acidity or neutrality. Uncommonly, alkaline phenolic compounds are studied for their antioxidant activity [16–18].

N-containing compounds, such as alkaloids, are abundantly present in natural products. It has been shown that these alkaloids exhibit protective action against free radicals due to the elimination of the cation radical 2,2'-azinobis-(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS<sup>•+</sup>) [19,20] as well as to their scavenging activity against the 1,1-diphenyl-2-picryl-hydrazyl (•DPPH) radical and their inhibitory effects on hydrogen peroxide radical [21].

Recently, among the phenolic alkaloids, Xie et al. [22] studied higenamine as an attractive scaffold to test its antioxidant effect. In particular, as shown in Figure 1, the peculiarity of higenamine structure is the presence of a protonated N-atom (site 2) which provides it with a strong electron-withdrawing capacity, thus resulting in an electron density change and suggesting an effect of pH on the antioxidant power of higenamine. For all these reasons and in order to rationalize the existing experimental results [22], we decided to carefully investigate the antioxidant properties of higenamine by employing a density functional theory (DFT)-based computational protocol previously and successfully used for a series of natural antioxidants [23–25]. The hydroperoxyl radical (•OOH) was chosen because its half-life allows the best interception by chemical scavengers [26,27]. Several reaction mechanisms (HAT, single-electron transfer (SET), and RAF) were considered, and the overall kinetic behaviour was evaluated.

## 2. Materials and Methods

All the calculations were carried out with the Gaussian 09 package of programs [28]. Full geometry optimizations and frequency calculations were done by using the DFT. The M06-2X functional coupled

with the extended 6-311+G(d) basis set was chosen because of the good performance of this level of theory for kinetic calculations [29] and was successfully used in previous works for modelling chemical reactions between antioxidant species and free radicals [30–36]. Unrestricted calculations were used for open-shell systems. The solvent effects, in water and pentylethanoate (PE) environments, were taken into account by using the Solvent Model based on Density SMD [37], which has been proven to estimate the solvation free energies for charged or uncharged solutes with relatively low errors. Local minima and transition states (TS) were identified by the number of imaginary frequencies (0 or 1, respectively). Intrinsic reaction coordinate calculations (IRC) were performed to verify if the located TS properly connected the relative minima along the reaction coordinate [38]. Thermodynamic corrections at 298.15 K were included in the calculation of relative energies. The used computational protocol is in line with the quantum mechanics-based test for overall free-radical scavenging activity (QM-ORSA), [39,40] which was validated by comparison with experimental results. Spin density computations were performed for the most stable open-shell species. Natural bond orbital (NBO) analysis [41–43], as implemented in the Gaussian 09 package, was to evaluate net charges, bond order, and conjugation.

### 3. Results and Discussion

As it is well known, in the aqueous phase, knowledge of the acid–base equilibrium is crucial for the individuation of the chemical species present in physiological conditions. For this reason, our preliminary calculations were devoted to the computation of the acid dissociation constants (pKas) of the investigated compound, using the parameter fitting method [44], and to the quantification of the relative molar fractions at pH 7.4 (see Figure 2). Considering all the possible deprotonation paths (see Figure 3), our results indicated that the first deprotonation occurred at the OH in position C6 (Figure 1), and the relative pKa1 value was 8.2, in agreement with experimental results [45]. The second deprotonation at pKa2 = 9.1 (see Figure 3) involved the NH<sub>2</sub> group, giving rise to the H<sub>2</sub>A<sup>−</sup> species. Furthermore, the loss of H<sup>+</sup> from the OH in position C4 generated the HA<sup>2−</sup> anion (pKa3 = 10.1), while the last deprotonation (pKa4 = 13.2) generated the A<sup>3−</sup> species. Looking at Table 1, it is possible to evidence that at physiological pH, in addition to the dominant H<sub>4</sub>A<sup>+</sup> species, also the H<sub>3</sub>A and H<sub>2</sub>A<sup>−</sup> species are present in aqueous solution, in molar fractions of 0.136 and 0.002, respectively.

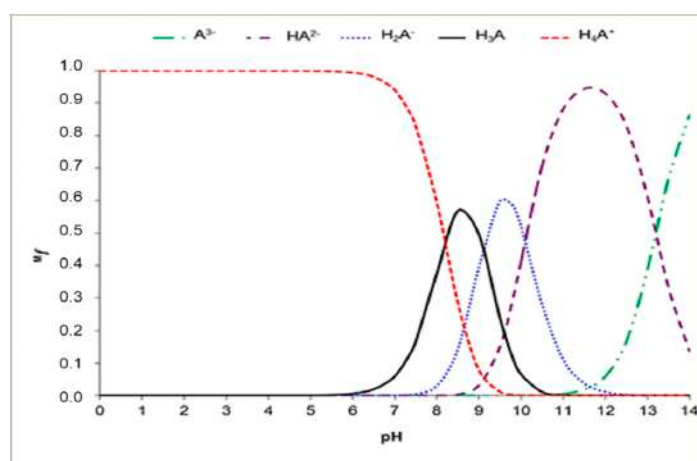
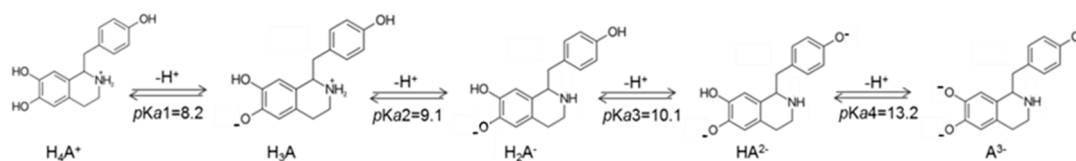


Figure 2. Distribution diagram of higenamine as a function of pH.

Table 1. Molar fractions (*M<sub>f</sub>*) of the different acid–base species of higenamine at physiological pH.

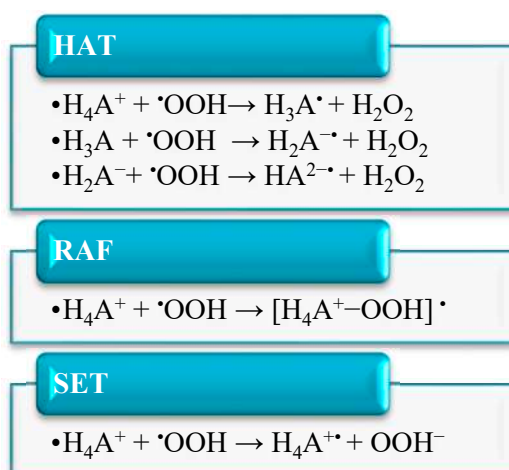
H <sub>4</sub> A <sup>+</sup>	H <sub>3</sub> A	H <sub>2</sub> A <sup>−</sup>	HA <sup>2−</sup>	A <sup>3−</sup>
0.861	0.136	0.002	0.000	0.000



**Figure 3.** pKa values of the relative deprotonation paths of higenamine at physiological pH.

These results indicated that for higenamine, it is necessary to take into account the species produced by the first two acid–base equilibria for the determination of the antioxidant power of higenamine toward the  $\cdot\text{OOH}$  radical in an aqueous environment.

The investigated reaction mechanisms are summarized in Scheme 1:



**Scheme 1.** Schematic representation of the considered mechanisms for higenamine and its acid–base forms. Gibbs free energies of reaction ( $\Delta\text{G}$ ) and activation ( $\Delta\text{G}^\ddagger$ ), for all populated species in aqueous and pentylethanoate (PE) solvents, involved in the studied mechanisms are summarized in Table 2. HAT, hydrogen atom transfer, RAF, radical adduct formation, SET, single-electron transfer.

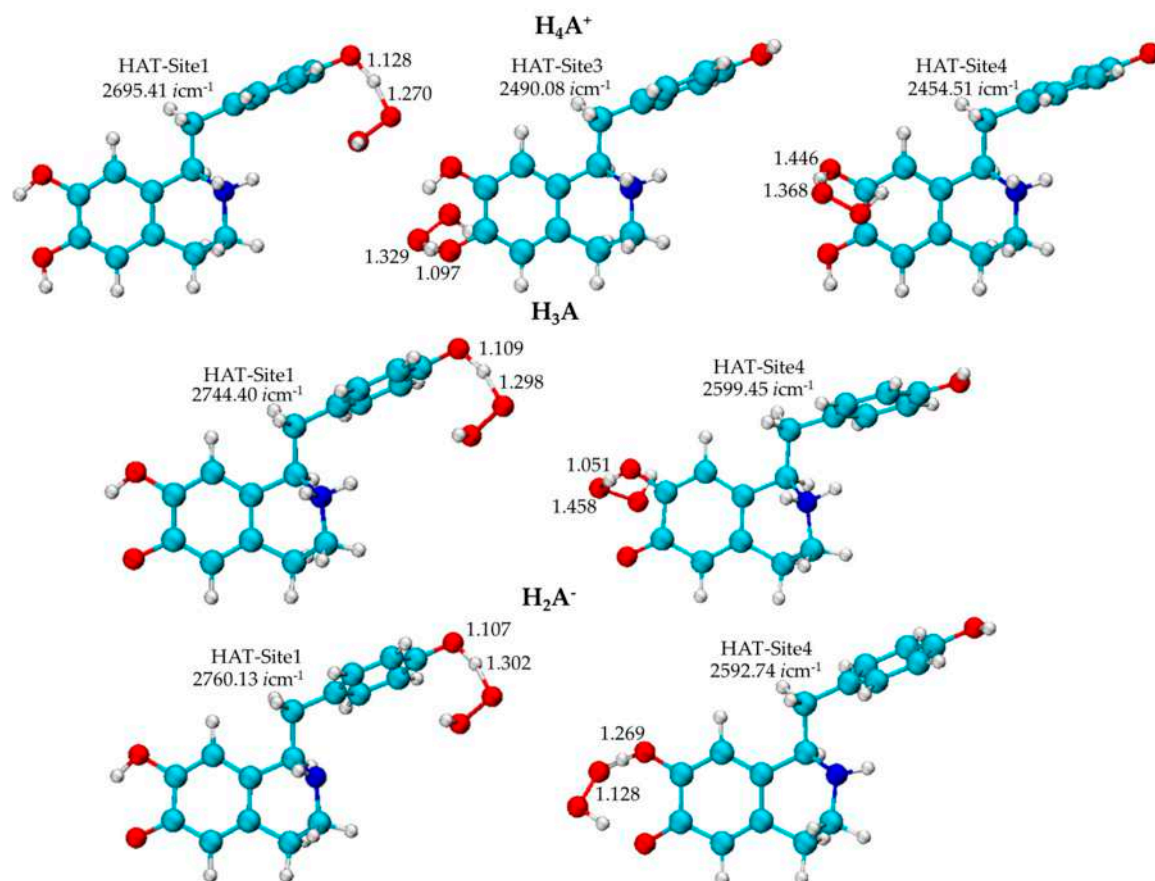
**Table 2.** Gibbs free energies of reaction ( $\Delta\text{G}$ ) and activation ( $\Delta\text{G}^\ddagger$ ), expressed in  $\text{kcal mol}^{-1}$ , at 298.15 K. in aqueous solution involved in HAT and SET. The  $\Delta\text{G}$  values calculated in PE medium are reported in parenthesis.

	$\text{H}_4\text{A}^+$		$\text{H}_3\text{A}$		$\text{H}_2\text{A}^-$	
	$\Delta\text{G}$	$\Delta\text{G}^\ddagger$	$\Delta\text{G}$	$\Delta\text{G}^\ddagger$	$\Delta\text{G}$	$\Delta\text{G}^\ddagger$
HAT-O4'	1.82	23.35	−0.26	21.80	−3.81	21.36
HAT-O6	−3.62	19.99				
HAT-O7	−2.58	20.93	−13.34	8.49	−15.63	1.33
HAT-C1	2.43		−0.85	13.39		
HAT-C3	16.46		15.58			
HAT-C4	3.19		2.55			
HAT-C7	6.33		7.57			
SET	31.19 (91.53)		4.42 (36.60)		−1.03 (15.11)	
RAF-C7	15.58	23.89				
RAF-C8	19.14	21.61				
RAF-C9	22.73	24.11				
RAF-C10	15.93	20.17				
RAF-C5	19.14	21.61				
RAF-C6	15.65	24.37				
RAF-C1'	19.84	26.42				
RAF-C6'	22.21	28.95				
RAF-C2'	18.17	28.09				
RAF-C5'	18.65	24.36				

Concerning the HAT mechanism, in the  $H_4A^+$  form, only the  $\bullet OOH$  attack on hydrogens at site 3 (O6) and site 4 (O7) of the catechol moiety results to be exergonic by 3.62 and 2.58 kcal mol<sup>-1</sup>, respectively, while at the  $H_3A$  sites, 1 (O4') and 4 (O7) are thermodynamically favoured. In the  $H_2A^-$  species, the Gibbs energies for the reaction at sites 1 and 4 are -3.81 and -15.63 kcal mol<sup>-1</sup>, respectively. In addition, for the catecholic moiety in higenamine going from the neutral form to the anionic one, the trend of the obtained  $\Delta G$  values well reproduces that observed for catechol [46].

The most exergonic reaction path is HAT from site 4 of the catechol moiety, for both  $H_2A^-$  and  $H_3A$  species.

The optimized geometries of the transition states are illustrated in Figure 4.



**Figure 4.** Geometries of the transition states of the  $H_4A^+$ ,  $H_3A$ , and  $H_2A^-$  forms obtained in water as a result of hydrogen atom transfer mechanism at the M062X/6-311+G(d) level of theory.

From Table 2, it is possible to note that the activation energy for the reaction of  $H_4A^+$  species for which the TS have been characterized, assumes values ranging from about 20 to 28 kcal mol<sup>-1</sup>, whereas for the reaction of  $H_3A$  and  $H_2A^-$  species, the  $\Delta G^\ddagger$  values range from 1 to 21 kcal mol<sup>-1</sup>.

In pentylethanoate solvent, that mimics the lipid environment and in which only the  $H_4A^+$  form is present, the obtained  $\Delta G$  for the  $\bullet OOH$  attack at the different sites had positive values (0.50, 6.43, 0.58 kcal mol<sup>-1</sup> for O4', O6, O7, respectively), indicating that the HAT process can hardly take place.

Concerning the RAF mechanism for the most abundant species  $H_4A^+$ , all  $\bullet OOH$  addition channels were found to be endergonic (Table 2), with  $\Delta G$  values larger than 20 kcal mol<sup>-1</sup>. Generally, the kinetic calculations for endergonic channels are excluded because, although they might occur at significant rates, the reaction is reversible, and no products are observed. Instead, if the products are able to further react quickly producing a driving force and barriers are small, these processes also need to be considered [44]. For this reason, they were considered in our kinetic computation.



In a water environment, SET reactions for all the investigated species, with the exception of the anionic form, gave also endergonic  $\Delta G$  (see Table 2). The degree of deprotonation contributed to increasing the thermochemical viability of the SET process, thus pointing out the role of all the existing species at physiological pH in the determination of the antioxidant power of a chemical species. In PE medium, the SET mechanism probably does not contribute to the overall reactivity of higenamine towards  $\bullet\text{OOH}$ , since such an environment does not provide the necessary solvation of the intermediate ionic species yielded by this mechanism. Accordingly, the calculated  $\Delta G$  for the SET mechanism were found to be largely endergonic.

The computed rate constants in aqueous solution for all the considered mechanisms are shown in Table 3, which also reports the overall rate coefficient calculated as the sum of the rate constants of each path.

**Table 3.** Rate constants ( $\text{M}^{-1} \text{s}^{-1}$ ) and branching ratios ( $\Gamma$ ) computed at the M062x level of theory at 298.15 K.

	$\text{H}_4\text{A}^+$		$\text{H}_3\text{A}$		$\text{H}_2\text{A}^-$	
	$k$	$\Gamma(\%)$	$k$	$\Gamma(\%)$	$k$	$\Gamma(\%)$
HAT-O4'	$1.62 \times 10^1$	11.02	$3.11 \times 10^1$	~0.00	$6.85 \times 10^1$	~0.00
HAT-O6	$1.01 \times 10^2$	68.67	~	~	~	~
HAT-O7	$2.12 \times 10^1$	14.41	$5.03 \times 10^8$	65.58	$2.79 \times 10^9$	29.25
HAT-C1	~	~	$1.25 \times 10^5$	0.02	~	~
SET	$1.04 \times 10^{-13}$	~0.00	$2.64 \times 10^8$	34.40	$6.75 \times 10^9$	70.75
RAF-C7	$1.53 \times 10^{-2}$	0.01				
RAF-C8	$4.07 \times 10^{-3}$	~0.00				
RAF-C9	$9.68 \times 10^{-3}$	0.01				
RAF-C10	$7.85 \times 10^1$	5.34				
RAF-C5	$7.22 \times 10^{-1}$	0.49				
RAF-C6	$7.15 \times 10^{-3}$	~0.00				
RAF-C1'	$2.07 \times 10^{-4}$	~0.00				
RAF-C6'	$1.46 \times 10^{-5}$	~0.00				
RAF-C2'	$3.43 \times 10^{-6}$	~0.00				
RAF-C5'	$5.69 \times 10^{-2}$	0.04				
RAF-C3'	$7.37 \times 10^{-3}$	0.01				
RAF-C4'	$2.10 \times 10^{-5}$	~0.00				
Overall	$1.47 \times 10^2$		$7.67 \times 10^8$		$9.54 \times 10^9$	

The rate constants showed that for the  $\text{H}_4\text{A}^+$  species, the HAT mechanism was favoured with respect to SET and RAF. In particular, the rate constants for H abstraction from site 3 of catechol were 6.23 and 4.76 times higher than those for site 1 and site 4, respectively. For the  $\text{H}_3\text{A}$  species, the rate constant associated with the SET mechanism was feasible, with a value equal to  $2.64 \times 10^8 \text{ L mol}^{-1} \text{ s}^{-1}$ . In general, the results indicated that the rate constant increased in the presence of the anionic form of higenamine.

In order to assert the substantial contribution of each species, the calculated overall rate coefficients were corrected by considering the population of each acid–base form at physiological pH. The derived sum of the corrected-by-fraction total rate coefficients are reported in Table 4.

**Table 4.** Molar fractions ( $f$ ), total rate constants ( $k_{\text{tot}}$ ,  $\text{M}^{-1} \text{s}^{-1}$ ), and corrected-by-fraction total rate coefficients ( $f k_{\text{tot}}$ ,  $\text{M}^{-1} \text{s}^{-1}$ ) at 298.15 K, in aqueous solution at pH 7.4.

	$f$	$k_{\text{tot}}$	$f k_{\text{tot}}$
$\text{H}_4\text{A}^+$	0.861	$1.47 \times 10^2$	$1.26 \times 10^2$
$\text{H}_3\text{A}$	0.136	$7.67 \times 10^8$	$1.04 \times 10^8$
$\text{H}_2\text{A}^-$	0.002	$9.54 \times 10^9$	$1.91 \times 10^7$

From Table 4, it is evident that the overall reactivity of higenamine against the  $\bullet\text{OOH}$  radical is almost entirely due to the presence in aqueous solution of the  $\text{H}_3\text{A}$  form. Furthermore, even if the total rate coefficient of the anionic form is higher than those of the other two forms,  $\text{H}_2\text{A}^-$  is present in very small quantities and cannot be considered the driving force of the antioxidant activity. On the other hand, at pH 7.4, the most abundant species remains  $\text{H}_4\text{A}^{+}$  with a total rate constant of  $1.26 \times 10^2 \text{ M}^{-1} \text{ s}^{-1}$ .

To put into perspective the potential role of higenamine as an antioxidant, the overall rate coefficient of its reaction with  $\bullet\text{HOO}$ , in aqueous solution at physiological pH, was compared with those obtained for known antioxidants using a similar methodology. According to the estimated data, higenamine ( $k_{\text{overall}} = 1.23 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$ , this work) is more efficient for scavenging hydroperoxyl radicals than Trolox ( $8.96 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ ) [47], which is frequently used as a reference antioxidant. It is also more efficient for that purpose than melatonin ( $2.0 \times 10^1 \text{ M}^{-1} \text{ s}^{-1}$ ) and caffeine ( $3.3 \times 10^{-1} \text{ M}^{-1} \text{ s}^{-1}$ ) [48,49], whose protective action against oxidative stress is frequently associated with their ROS scavenging activity [50–52].

The results underline that the antioxidant power of higenamine is strongly influenced by pH. Higher pH values increase the protective effects against oxidative stress due to the increase of the  $\text{H}_3\text{A}$  and  $\text{H}_2\text{A}^-$  forms. In particular, at more basic pHs, SET and SPLET become the favored mechanisms.

#### 4. Conclusions

A systematic study of the reactivity of higenamine toward  $\bullet\text{OOH}$  was carried out in aqueous and lipid environments considering hydrogen transfer, single-electron transfer, and radical adduct formation. In aqueous solution, the species coming from acid–base equilibria were taken into account. It was found that the hydrogen atom transfer from the catecholic ring was the main reaction channel for the  $\text{H}_4\text{A}^+$  species. In contrast, the  $\text{H}_3\text{A}$  and  $\text{H}_2\text{A}^-$  forms,  $\bullet\text{OOH}$  scavenging activity took place almost exclusively via SET. Our computed kinetic constants revealed that in water solution it is important to consider all the species derived from acid–base equilibria to obtain reliable results. Comparisons with other species considered good antioxidants revealed that the  $\bullet\text{OOH}$  scavenging activity of higenamine is higher than that of Trolox. In addition, higenamine resulted to be more efficient than melatonin and caffeine for that purpose.

**Author Contributions:** T.M. and N.R. conceived the presented idea and supervised the project. I.R. and A.P. performed the computations. I.R. and T.M. verified and analyzed the data. I.R. wrote the manuscript with support from N.R., A.G., T.M. and J.R.A.-I. contributed to the final version of the manuscript, drafted the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** I.R.: A.P., N.R., T.M. gratefully acknowledge the Dipartimento di Chimica e Tecnologie Chimiche dell'Università della Calabria and thank the Ministero degli Affari Esteri e della Cooperazione Internazionale-Italia. A.G. and J.R.A.-I. gratefully acknowledge the Laboratorio de Visualizacion y Computo Paralelo at Universidad Autonoma Metropolitana Iztapalapa.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Zhang, N.; Lian, Z.; Peng, X.; Li, Z.; Zhu, H. Applications of Higenamine in pharmacology and medicine. *J. Ethnopharmacol.* **2017**, *196*, 242–252. [[CrossRef](#)]
2. Bai, J.; Mao, J.; Yang, H.; Khan, A.; Fan, A.; Liu, S.; Zhang, J.; Wang, D.; Gao, H.; Zhang, J. Sucrose non-ferment 1 related protein kinase 2 (SnRK2) genes could mediate the stress responses in potato (*Solanum tuberosum* L.). *BMC Genet.* **2017**, *18*, 41. [[CrossRef](#)] [[PubMed](#)]
3. Kosuge, T.; Yokota, M. Studies on cardiac principle of aconite root. *Chem. Pharm. Bull.* **1976**, *24*, 176–178. [[CrossRef](#)] [[PubMed](#)]
4. Zhang, N.; Qu, K.; Wang, M.; Yin, Q.; Wang, W.; Xue, L.; Fu, H.; Zhu, H.; Li, Z. Identification of higenamine as a novel  $\alpha 1$ -adrenergic receptor antagonist. *Phytother. Res.* **2019**, *33*, 708–717. [[CrossRef](#)] [[PubMed](#)]

5. Lee, Y.S.; Kang, Y.J.; Kim, H.J.; Park, M.K.; Seo, H.G.; Lee, J.H.; Yun-Choi, H.S.; Chang, K.C. Higenamine reduces apoptotic cell death by induction of heme oxygenase-1 in rat myocardial ischemia-reperfusion injury. *Apoptosis* **2006**, *11*, 1091–1100. [[CrossRef](#)]
6. Ramalakshmi, K.; Kubra, I.R.; Rao, L.J.M. Antioxidant potential of low-grade coffee beans. *Food Res. Int.* **2008**, *41*, 96–103. [[CrossRef](#)]
7. Dandlen, S.A.; Lima, A.S.; Mendes, M.D.D.S.; Miguel, M.G.; Faleiro, M.; Sousa, M.J.; Pedro, L.G.; Barroso, J.G.; Figueiredo, A.C. Antioxidant activity of six Portuguese thyme species essential oils. *Flavour Fragr. J.* **2010**, *25*, 150–155. [[CrossRef](#)]
8. Lobo, V.; Patil, A.; Phatak, A.; Chandra, N. Free radicals, antioxidants and functional foods: Impact on human health. *Pharmacogn. Rev.* **2010**, *4*, 118–126. [[CrossRef](#)]
9. Pandey, K.B.; Rizvi, S.I. Plant polyphenols as dietary antioxidants in human health and disease. *Oxidative Med. Cell. Longev.* **2009**, *2*, 270–278. [[CrossRef](#)]
10. Duthie, G.G.; Duthie, S.; Kyle, J.A.M. Plant polyphenols in cancer and heart disease: Implications as nutritional antioxidants. *Nutr. Res. Rev.* **2000**, *13*, 79–106. [[CrossRef](#)]
11. Betteridge, D.J. What is oxidative stress? *Metab. Clin. Exp.* **2000**, *49*, 3–8. [[CrossRef](#)]
12. Pham-Huy, L.A.; He, H.; Pham-Huy, C. Free Radicals, Antioxidants in Disease and Health. *Int. J. Biomed. Sci.* **2008**, *4*, 89–96. [[PubMed](#)]
13. Seifried, H.E.; Anderson, D.E.; Fisher, E.I.; Milner, J.A. A review of the interaction among dietary antioxidants and reactive oxygen species. *J. Nutr. Biochem.* **2007**, *18*, 567–579. [[CrossRef](#)] [[PubMed](#)]
14. Ichikawa, K.; Sasada, R.; Chiba, K.; Gotoh, H. Effect of Side Chain Functional Groups on the DPPH Radical Scavenging Activity of Bisabolane-Type Phenols. *Antioxidants* **2019**, *8*, 65. [[CrossRef](#)] [[PubMed](#)]
15. Devasagayam, T.P.; Tilak, J.C.; Bloor, K.K.; Sane, K.S.; Ghaskadbi, S.S.; Lele, R.D. Free radicals and antioxidants in human health: Current status and future prospects. *J. Assoc. Phys. India* **2004**, *52*, 4.
16. Galano, A.; Alvarez-Idaboy, J.R. Computational strategies for predicting free radical scavengers' protection against oxidative stress: Where are we and what might follow? *Int. J. Quantum Chem.* **2018**, *119*, e25665. [[CrossRef](#)]
17. Lee, S.E.; Hwang, H.J.; Ha, J.-S.; Jeong, H.-S.; Kim, J.H. Screening of medicinal plant extracts for antioxidant activity. *Life Sci.* **2003**, *73*, 167–179. [[CrossRef](#)]
18. Li, X.; Lin, J.; Chen, B.; Xie, H.; Chen, D. Antioxidant and Cytoprotective Effects of Kukoamines A and B: Comparison and Positional Isomeric Effect. *Molecules* **2018**, *23*, 973. [[CrossRef](#)]
19. Herraiz, T.; Galisteo, J. Tetrahydro- $\beta$ -carboline Alkaloids Occur in Fruits and Fruit Juices. Activity as Antioxidants and Radical Scavengers. *J. Agric. Food Chem.* **2003**, *51*, 7156–7161. [[CrossRef](#)]
20. Herraiz, T.; Galisteo, J.; Chamorro, C. L-Tryptophan Reacts with Naturally Occurring and Food-Occurring Phenolic Aldehydes To Give Phenolic Tetrahydro- $\beta$ -carboline Alkaloids: Activity as Antioxidants and Free Radical Scavengers. *J. Agric. Food Chem.* **2003**, *51*, 2168–2173. [[CrossRef](#)]
21. Yang, Z.; Liu, C.; Xiang, L.; Zheng, Y. Phenolic alkaloids as a new class of antioxidants in *Portulaca oleracea*. *Phytother. Res.* **2009**, *23*, 1032–1035. [[CrossRef](#)] [[PubMed](#)]
22. Xie, Y.; Li, X.; Chen, J.; Deng, Y.; Lu, W.; Chen, D. pH Effect and Chemical Mechanisms of Antioxidant Higenamine. *Molecules* **2018**, *23*, 2176. [[CrossRef](#)] [[PubMed](#)]
23. Marino, T.; Russo, N.; Galano, A. A deeper insight on the radical scavenger activity of two simple coumarins toward OOH radical. *Comput. Theor. Chem.* **2016**, *1077*, 133–138. [[CrossRef](#)]
24. Galano, A.; Mazzone, G.; Alvarez-Diduk, R.; Marino, T.; Alvarez-Idaboy, J.R.; Russo, N. Food Antioxidants: Chemical Insights at the Molecular Level. *Annu. Rev. Food Sci. Technol.* **2016**, *7*, 335–352. [[CrossRef](#)]
25. Leopoldini, M.; Russo, N.; Toscano, M. The molecular basis of working mechanism of natural polyphenolic antioxidants. *Food Chem.* **2011**, *125*, 288–306. [[CrossRef](#)]
26. Terpin, P.; Abramovič, H. A kinetic approach for evaluation of the antioxidant activity of selected phenolic acids. *Food Chem.* **2010**, *121*, 366–371. [[CrossRef](#)]
27. Sies, H. Oxidative stress: Oxidants and antioxidants. *Exp. Physiol.* **1997**, *82*, 291–295. [[CrossRef](#)]
28. 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.; et al. Gaussian 09, revision B. 01, Gaussian, Inc., Wallingford CT, 2009 Search PubMed;(b) NM O'Boyle, AL Tenderholt and KM Langner. *J. Comput. Chem.* **2008**, *29*, 839.

29. Zhao, Y.; Truhlar, D. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06 functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *119*, 525. [[CrossRef](#)]
30. Kumar, M.; Busch, D.H.; Subramaniam, B.; Thompson, W.H. Role of Tunable Acid Catalysis in Decomposition of  $\alpha$ -Hydroxyalkyl Hydroperoxides and Mechanistic Implications for Tropospheric Chemistry. *J. Phys. Chem. A* **2014**, *118*, 9701–9711. [[CrossRef](#)]
31. Alarcón, P.; Bohn, B.; Zetzsch, C.; Rayez, M.-T. Reversible addition of the OH radical to p-cymene in the gas phase: Multiple adduct formation. Part 2. *Phys. Chem. Chem. Phys.* **2014**, *16*, 17315–17326. [[CrossRef](#)] [[PubMed](#)]
32. Denegri, B.; Matić, M.; Kronja, O. A DFT-based model for calculating solvolytic reactivity. The nucleofugality of aliphatic carboxylates in terms of N f parameters. *Org. Biomol. Chem.* **2014**, *12*, 5698. [[CrossRef](#)] [[PubMed](#)]
33. Parkhomenko, D.; Edeleva, M.V.; Kiselev, V.; Bagryanskaya, E.G. pH-Sensitive C–ON Bond Homolysis of Alkoxyamines of Imidazoline Series: A Theoretical Study. *J. Phys. Chem. B* **2014**, *118*, 5542–5550. [[CrossRef](#)] [[PubMed](#)]
34. Da Silva, G. Reaction of Methacrolein with the Hydroxyl Radical in Air: Incorporation of Secondary O2 Addition into the MACR + OH Master Equation. *J. Phys. Chem. A* **2012**, *116*, 5317–5324. [[CrossRef](#)]
35. Zhao, Y.; Truhlar, D.G. How Well Can New-Generation Density Functionals Describe the Energetics of Bond-Dissociation Reactions Producing Radicals? *J. Phys. Chem. A* **2008**, *112*, 1095–1099. [[CrossRef](#)]
36. Galano, A.; Alvarez-Idaboy, J.R. Kinetics of radical-molecule reactions in aqueous solution: A benchmark study of the performance of density functional methods. *J. Comput. Chem.* **2014**, *35*, 2019–2026. [[CrossRef](#)]
37. Marenich, A.V.; Cramer, C.J.; Truhlar, D. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [[CrossRef](#)]
38. Maeda, S.; Harabuchi, Y.; Ono, Y.; Taketsugu, T.; Morokuma, K. Intrinsic reaction coordinate: Calculation, bifurcation, and automated search. *Int. J. Quantum Chem.* **2014**, *115*, 258–269. [[CrossRef](#)]
39. Galano, A.; Medina, M.E.; Tan, D.X.; Reiter, R.J. Melatonin and its metabolites as copper chelating agents and their role in inhibiting oxidative stress: A physicochemical analysis. *J. Pineal Res.* **2014**, *58*, 107–116. [[CrossRef](#)]
40. Galano, A.; Alvarez-Idaboy, J.R. A computational methodology for accurate predictions of rate constants in solution: Application to the assessment of primary antioxidant activity. *J. Comput. Chem.* **2013**, *34*, 2430–2445. [[CrossRef](#)]
41. Foster, J.P.; Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218. [[CrossRef](#)]
42. Reed, A.E.; Weinhold, F. Natural bond orbital analysis of near-Hartree–Fock water dimer. *J. Chem. Phys.* **1983**, *78*, 4066–4073. [[CrossRef](#)]
43. Reed, A.E.; Weinstock, R.B.; Weinhold, F. Natural population analysis. *J. Chem. Phys.* **1985**, *83*, 735–746. [[CrossRef](#)]
44. Galano, A.; Perez-Gonzalez, A.; Castañeda-Arriaga, R.; Muñoz-Rugeles, L.; Mendoza-Sarmiento, G.; Romero-Silva, A.; Ibarra-Escutia, A.; Rebollar-Zepeda, A.M.; Leon-Carmona, J.R.; Hernández-Olivares, M.A.; et al. Empirically Fitted Parameters for Calculating pKa Values with Small Deviations from Experiments Using a Simple Computational Strategy. *J. Chem. Inf. Model.* **2016**, *56*, 1714–1724. [[CrossRef](#)]
45. Grunewald, G.L.; Seim, M.R.; Lu, J.; Makhoul, M.; Criscione, K.R. Application of the Goldilocks Effect to the Design of Potent and Selective Inhibitors of Phenylethanolamine N-Methyltransferase: Balancing pKa and Steric Effects in the Optimization of 3-Methyl-1,2,3,4-tetrahydroisoquinoline Inhibitors by  $\beta$ -Fluorination. *J. Med. Chem.* **2006**, *49*, 2939–2952. [[CrossRef](#)]
46. Castañeda-Arriaga, R.; Perez-Gonzalez, A.; Reina, M.; Alvarez-Idaboy, J.R.; Galano, A. Comprehensive Investigation of the Antioxidant and Pro-oxidant Effects of Phenolic Compounds: A Double-Edged Sword in the Context of Oxidative Stress? *J. Phys. Chem. B* **2018**, *122*, 6198–6214. [[CrossRef](#)]
47. Alberto, M.E.; Grand, A.; Russo, N.; Galano, A. A physicochemical examination of the free radical scavenging activity of Trolox: Mechanism, kinetics and influence of the environment. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4642. [[CrossRef](#)]
48. Leon-Carmona, J.R.; Galano, A. Is Caffeine a Good Scavenger of Oxygenated Free Radicals? *J. Phys. Chem. B* **2011**, *115*, 4538–4546. [[CrossRef](#)]

49. Galano, A. On the direct scavenging activity of melatonin towards hydroxyl and a series of peroxy radicals. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7178. [[CrossRef](#)]
50. Carrillo-Vico, A.; Guerrero, J.; Lardone, P.J.; Reiter, R.J. A Review of the Multiple Actions of Melatonin on the Immune System. *Endocrine* **2005**, *27*, 189–200. [[CrossRef](#)]
51. Tan, D.X.; Chen, L.D.; Poeggeler, B.; Manchester, L.C.; Reiter, R.J.; Poeggler, B. Melatonin a potent endogenous hydroxyl radical scavenger. *Endocr. J.* **1993**, *1*, 57–60.
52. Shi, X.; Dalal, N.; Jain, A. Antioxidant behaviour of caffeine: Efficient scavenging of hydroxyl radicals. *Food Chem. Toxicol.* **1991**, *29*, 1–6. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).







**Paper VI**  
Quantum mechanical predictions of the antioxidant capability of moracin C isomers





# Quantum Mechanical Predictions of the Antioxidant Capability of Moracin C Isomers

Angela Parise<sup>1,2</sup>, Bruna Clara De Simone<sup>1</sup>, Tiziana Marino<sup>1</sup>, Marirosa Toscano<sup>1</sup> and Nino Russo<sup>1\*</sup>

<sup>1</sup> Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Rende, Italy, <sup>2</sup> Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France

## OPEN ACCESS

### Edited by:

Jorge Ignacio Martínez-Araya,  
Andrés Bello University, Chile

### Reviewed by:

José Pedro Cerón-Carrasco,  
Catholic University San Antonio of  
Murcia, Spain

Daniel Glossman-Mitnik,  
Advanced Materials Research Center  
(CIMAV), Mexico

### \*Correspondence:

Nino Russo  
nrusso@unical.it

### Specialty section:

This article was submitted to  
Theoretical and Computational  
Chemistry,  
a section of the journal  
Frontiers in Chemistry

**Received:** 10 February 2021

**Accepted:** 25 March 2021

**Published:** 21 April 2021

### Citation:

Parise A, De Simone BC, Marino T,  
Toscano M and Russo N (2021)  
Quantum Mechanical Predictions of  
the Antioxidant Capability of Moracin  
C Isomers. *Front. Chem.* 9:666647.  
doi: 10.3389/fchem.2021.666647

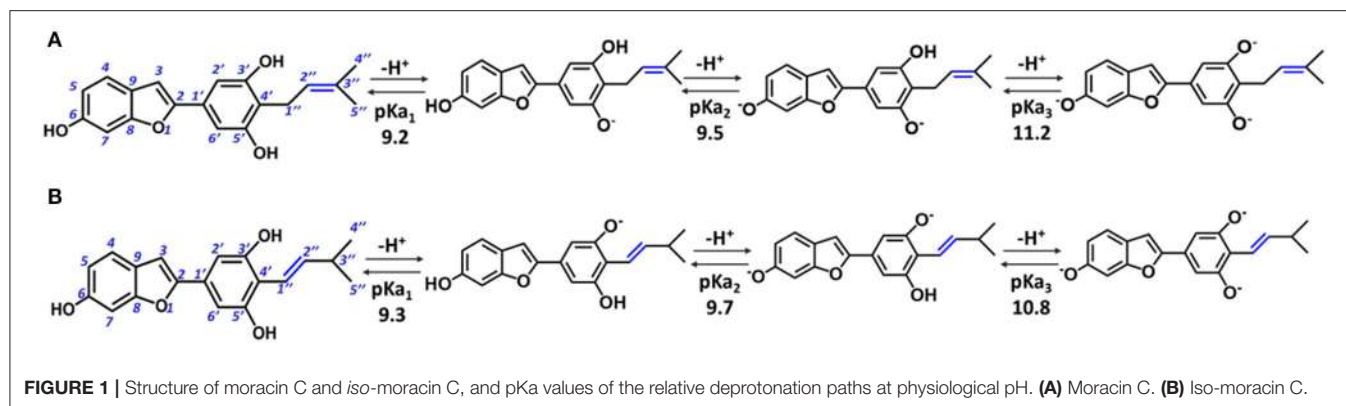
The antioxidant capability of moracin C and *iso*-moracin C isomers against the OOH free radical was studied by applying density functional theory (DFT) and choosing the M05-2X exchange-correlation functional coupled with the all electron basis set, 6-311++G(d,p), for computations. Different reaction mechanisms [hydrogen atom transfer (HAT), single electron transfer (SET), and radical adduct formation (RAF)] were taken into account when considering water- and lipid-like environments. Rate constants were obtained by applying the conventional transition state theory (TST). The results show that, in water, scavenging activity mainly occurs through a radical addition mechanism for both isomers, while, in the lipid-like environment, the radical addition process is favored for *iso*-moracin C, while, redox- and non-redox-type reactions can equally occur for moracin C. The values of pKa relative to the deprotonation paths at physiological pH were predicted in aqueous solution.

**Keywords:** moracin, antioxidants, DFT, kinetic constants, reaction mechanisms

## INTRODUCTION

In the last decades, 2-phenyl-benzofuran-containing molecules, found in a variety of plants (*Morus alba*, *Artocarpus champeden*, *Erythrina addisoniae*, and *Calpocalyx dinklagei*) (Hakim et al., 1999; Na et al., 2007; Naik et al., 2015; Kapche et al., 2017; Pel et al., 2017), have attracted considerable interest both for their massive use in pharmacology and for their ancient use in traditional medicine in Asia, Africa, and America (Fashing, 2001; Venkatesh and Seema, 2008; Kapche et al., 2009; Kuete et al., 2009). A rich source of natural products with a 2-phenyl-benzofuran basic scaffold is the *Moraceae* family (e.g., *M. alba*, *Morus mesozygia*, *Morus lhou*, and *Morus macroura*) (Sang-Hee et al., 2002), from which more than 24 molecules (moracin A–Z) have already been isolated and characterized (Nguyen et al., 2009). Many of them showed a variety of biological and pharmacological activities and were tested as potent antioxidants (Kapche et al., 2009; Seong et al., 2018), anti-cancer agents (Nguyen et al., 2009), anti-inflammatories, and anti-microbial agents (Kuete et al., 2009; Zelová et al., 2014; Lee et al., 2016). Furthermore, they were proven to act as cholinesterase (Delogu et al., 2016; Seong et al., 2018) and  $\beta$ -site amyloid precursor protein cleaving enzyme 1 (BACE1) (Jeon et al., 2007; Seong et al., 2018) inhibitors *in vitro*.

In particular, moracin C {2-[3',5'-dihydroxy-4'-(3-methylbut-2-enyl)phenyl]-6-hydroxybenzofuran} and its *iso*-moracin C isomer {2-[3',5'-dihydroxy-4'-(3-methylbut-1-enyl)phenyl]-6-hydroxybenzofuran} (see **Figure 1**), extracted from *M. alba* and *Artocarpus heterophyllus*, exhibit antioxidant capabilities (Li et al., 2018; Seong et al., 2018) and other biological

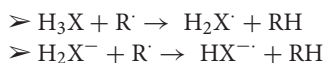


functions correlated with oxidative stress (Zelová et al., 2014; Naik et al., 2015; Li et al., 2018; Seong et al., 2018). The only structural difference between the two isomers is the position of the C=C double bond in the methylbut-enyl moiety (see **Figure 1**). This apparent small structural difference may have significant consequences on the electronic and reactivity properties of the two isomers. In fact, when the C=C bond is close to the phenyl ring (as occurs in *iso* moracin), the electronic delocalization between the two groups increases, stabilizing accordingly the radical that is formed as a result of O–H abstraction reaction. On the contrary, the localization of the double bond in position 2'' prevents conjugation with the phenolic ring and, in principle, would favor radical attack reactions.

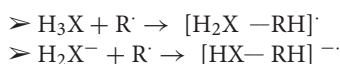
Very recently, in an accurate experimental study (Li et al., 2018), the authors attempted to correlate the estimated antioxidant properties with the position of the C=C bond in the two isomers, concluding that “[B]oth moracin C and *iso*-moracin C can inhibit ROS, likely through redox-related pathways (especially ET and H<sup>+</sup>-transfer) and a non-redox-related RAF pathway. In the redox-related pathways, a double bond at the conjugation position can enhance the ET and H<sup>+</sup>-transfer potential. However, in the non-redox-related pathway, the double bond position hardly affected the RAF potential.”

We have conducted an accurate theoretical study on the thermodynamic and kinetic properties of moracin C and *iso*-moracin C when reacting with the OOH free radical by considering the following most common antioxidant scavenging reaction mechanisms (Leopoldini et al., 2011; Alberto et al., 2013; Mazzone et al., 2015; Galano et al., 2016; Markovic et al., 2016; Ahmadi et al., 2018; Castaneda-Arriaga et al., 2020; Romeo et al., 2020):

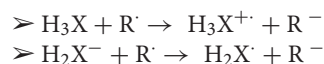
HAT: hydrogen atom transfer



RAF: radical adduct formation



SET: single electron transfer



## COMPUTATIONAL DETAILS

All calculations were performed with the Gaussian 09 code (Frisch et al., 2009) by applying the density functional theory. Following a well-consolidated protocol that was proven to be reliable in a large amount of antioxidant systems (Galano et al., 2016; Pérez-González et al., 2020), the M05-2X functional (Zhao et al., 2006) and the all electron basis set, 6-311++G(d,p) were chosen for all computations. Geometry optimization without any constraint was followed by frequency calculations to verify if the obtained structures were local minima (0 imaginary frequency) and transition states (TSs) (1 imaginary frequency) and to obtain zero-point energy corrections. Furthermore, for the TSs, it was verified that the imaginary frequency matched with the expected motion along the reaction coordinate. The solvation model based on density (SMD) (Marenich et al., 2009) was used to mimic the aqueous and lipid-like environments (water and pentyl ethanoate, respectively). Intrinsic reaction coordinate computations were performed to verify if the intercepted TSs properly connected to the relative minima in a given path.

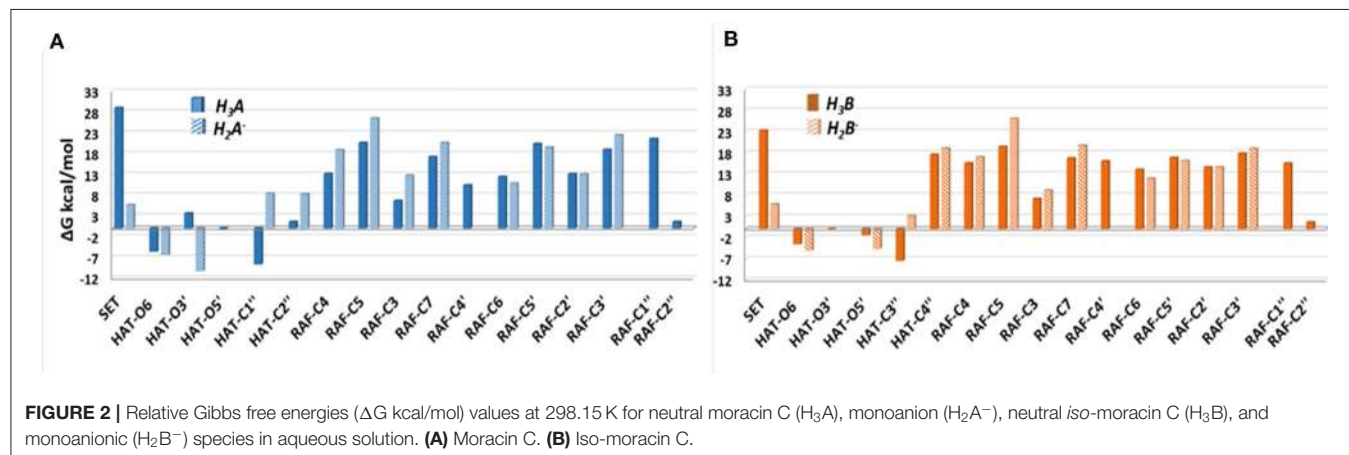
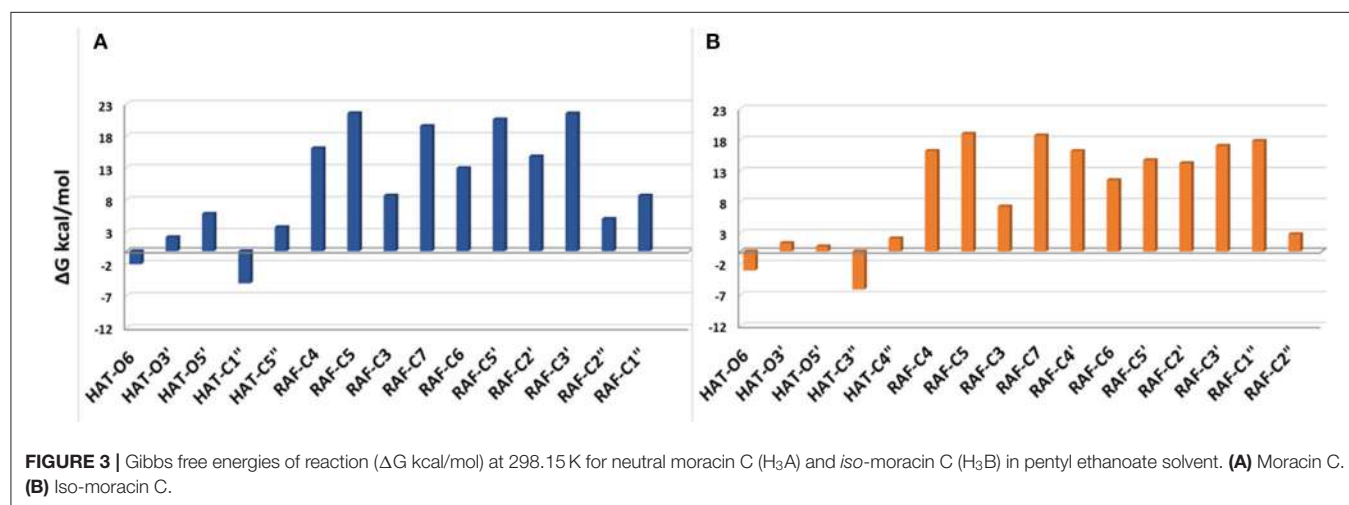
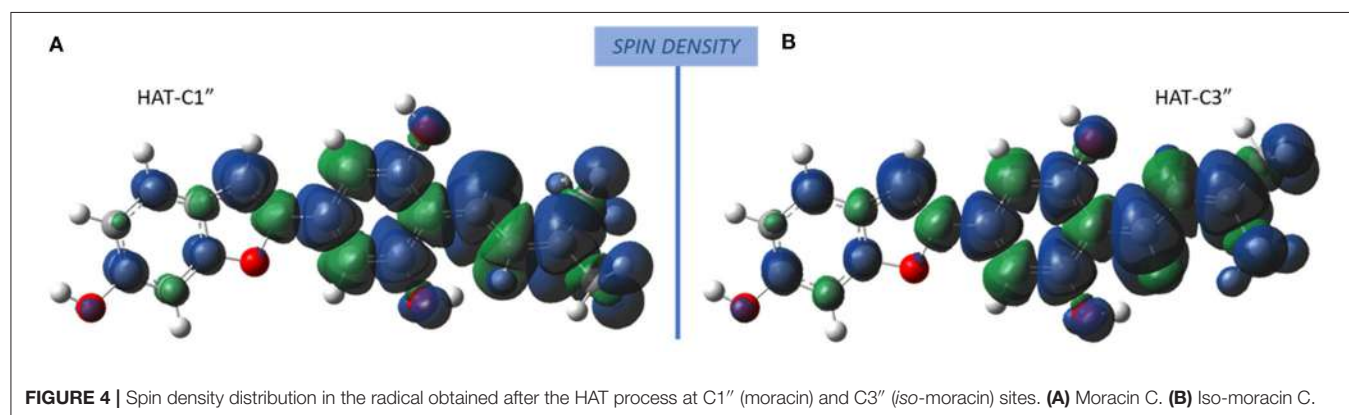
Relative energies were computed with respect to the sum of separate reactants, and the thermodynamics corrections at 298.15 K were taken into account following the quantum mechanics-based test for the overall free radical scavenging activity (QM-ORSA) procedure (Galano and Alvarez-Idaboy, 2013, 2019). Rate constants, *k*, were determined by applying the conventional transition state theory (TST) at the 1M standard state (Truhlar et al., 1996). For the mechanism involving SETs, the barriers of reaction were computed using the Marcus theory (Marcus, 1957). For rate constants, close to the diffusion limit, the Collins–Kimball theory (Collins and Kimball, 1949) was applied.

## RESULT AND DISCUSSION

For the study in water environment, knowledge of the acid-base equilibria under physiological conditions (pH = 7.4) is

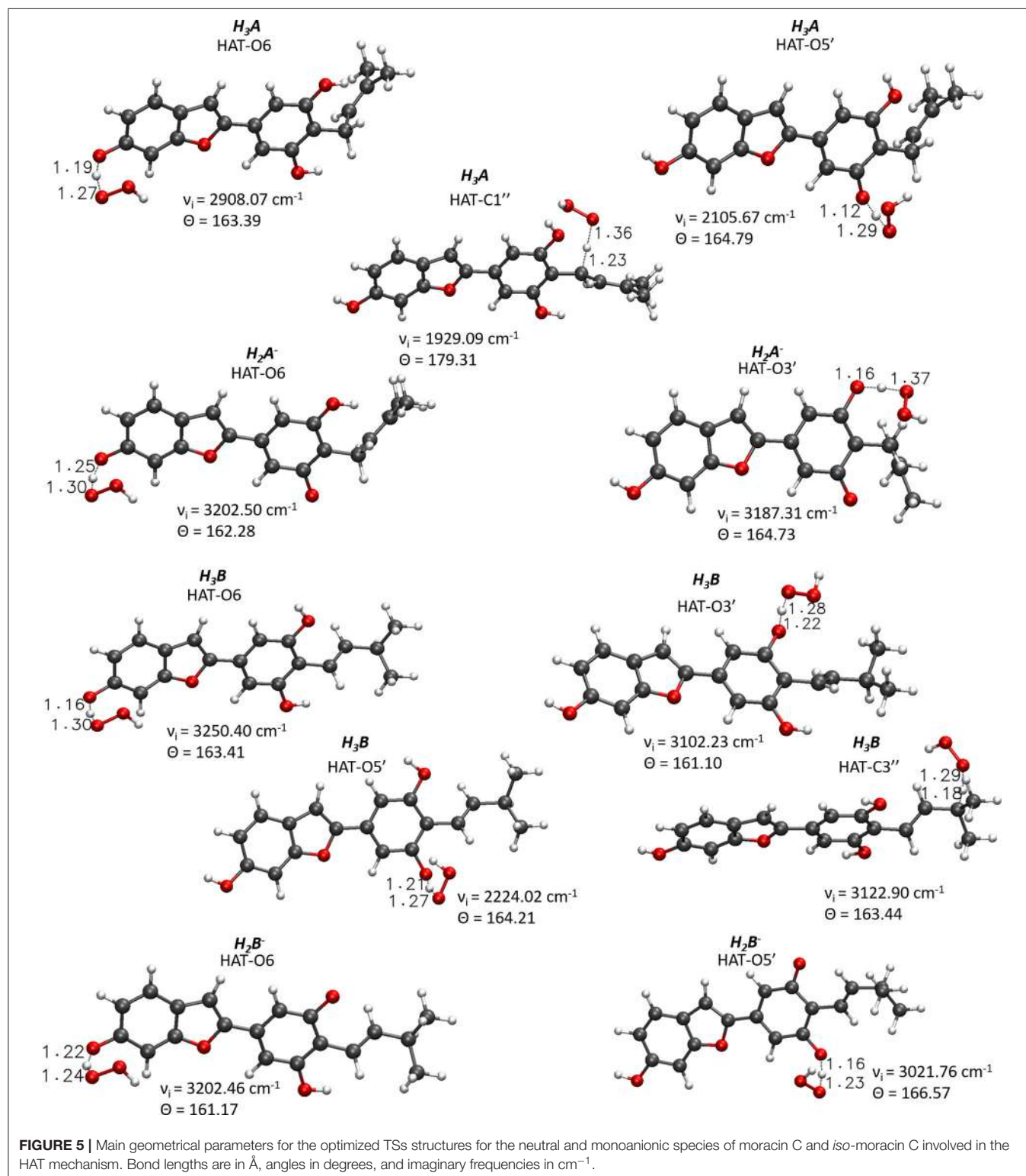
**TABLE 1** | pKa value and molar fractions (Mf) of the different acid–base species of moracin C and *iso*-moracin C, at physiological pH.

Molecule	pKa <sub>1</sub>	pKa <sub>2</sub>	pKa <sub>3</sub>	Mf (H <sub>3</sub> X)	Mf (H <sub>2</sub> X) <sup>-</sup>	Mf (HX) <sup>2-</sup>	Mf (HX) <sup>3-</sup>
Moracin C	9.2	9.5	11.2	$9.8 \times 10^{-1}$	$1.6 \times 10^{-2}$	$1.2 \times 10^{-4}$	$2.0 \times 10^{-8}$
Iso-moracin C	9.3	9.7	10.8	$9.9 \times 10^{-1}$	$1.2 \times 10^{-2}$	$6.2 \times 10^{-5}$	$2.5 \times 10^{-8}$

**FIGURE 2** | Relative Gibbs free energies ( $\Delta G$  kcal/mol) values at 298.15 K for neutral moracin C (H<sub>3</sub>A), monoanion (H<sub>2</sub>A<sup>-</sup>), neutral *iso*-moracin C (H<sub>3</sub>B), and monoanionic (H<sub>2</sub>B<sup>-</sup>) species in aqueous solution. **(A)** Moracin C. **(B)** Iso-moracin C.**FIGURE 3** | Gibbs free energies of reaction ( $\Delta G$  kcal/mol) at 298.15 K for neutral moracin C (H<sub>3</sub>A) and *iso*-moracin C (H<sub>3</sub>B) in pentyl ethanoate solvent. **(A)** Moracin C. **(B)** Iso-moracin C.**FIGURE 4** | Spin density distribution in the radical obtained after the HAT process at C1'' (moracin) and C3'' (*iso*-moracin) sites. **(A)** Moracin C. **(B)** Iso-moracin C.

very important. Because of lack of experimental information on both studied isomers, the relative pKa values were obtained (Table 1) using the parameters fitting method, which was previously proven to give results that are in good agreement

with experimental data (Pérez-González et al., 2018). The deprotonation path of the two study systems is shown in Figure 1. The preferred deprotonation site in moracin C is the OH in the C5' position, followed by those in C6 and C3'. On





**TABLE 2** | Gibbs free energies of reaction ( $\Delta G$ ) and activation ( $\Delta G^\ddagger$  kcal/mol) at 298.15 K in aqueous solution for neutral and monoanion moracin C and *iso*-moracin C species for the considered mechanisms.

Mechanism	$H_3A$		$H_2A^-$		$H_3B$		$H_2B^-$		$H_3A^{PE}$		$H_3B^{PE}$	
	$\Delta G$	$\Delta G^\ddagger$	$\Delta G$	$\Delta G^\ddagger$	$\Delta G$	$\Delta G^\ddagger$	$\Delta G$	$\Delta G^\ddagger$	$\Delta G$	$\Delta G^\ddagger$	$\Delta G$	$\Delta G^\ddagger$
SET	29.03		5.65		23.21		5.89					
HAT-O6	-5.62	19.78	-6.43	19.81	-3.62	21.04	-5.12	19.52	-0.02	17.23	-3.05	15.39
HAT-O3'	3.63		-10.32	17.04	0.01	19.86					0.84	17.23
HAT-O5'	-0.14	20.94			-1.81	26.27	-4.71	18.15				
HAT-C1''	-8.66	19.06							-5.02	6.08		
HAT-C3''					-7.52	17.85					-6.12	11.95

Apex PE refers to the neutral moracin C and *iso*-moracin C in pentyl ethanoate solvent.

the contrary, in *iso*-moracin C, the preferred deprotonation site is the OH in the C3' position, while the second and the third ones involve sites C6 and C5', respectively. In both conformers, all deprotonation sites are found in the benzene ring. A look at the molecular electrostatic potential, whose maps are reported in **Supplementary Figure 1**, shows that, in the case of *iso*-moracin C, the presence of the double bond in position C1'-C2' increases the  $\pi$  delocalization as proven by a great negative charge on the oxygen of the hydroxyl group on C5' position. The charge distribution reported in **Supplementary Table 1** further underlines that the localization of the double bond of methylbut-enyl moiety can influence the acid-base equilibrium of the two isomers. The calculated pKa values, at pH = 7.4 (see **Table 1**), indicate that, for both isomers, the neutral species are prevalent (molar fractions are 0.98 and 0.99 for moracin and *iso*-moracin, respectively). The monoanionic forms were not negligible in both isomers (see **Supplementary Figure 2**), so the  $H_3X$  and  $H_2X^-$  species were considered in the water environment study.

The Gibbs free energies of reaction ( $\Delta G$ ), computed for the two investigated mechanisms in water and lipid-like environments, are reported in **Figures 2, 3**. As can be seen, for both molecules and environments,  $\Delta G$  values for the RAF mechanism are all positive. However, since a recent experimental study (Li et al., 2018) suggested that this kind of mechanism might happen instead, we have also considered the addition of the OOH free radical to the C2'' sites, in which the obtained Gibbs reaction energies assume the less positive values.

Although the  $\Delta G$  values obtained for SET are always positive, we have also considered this mechanism that was found active in several systems that had been previously studied (Galano et al., 2016; Castaneda-Arriaga et al., 2020; Romeo et al., 2020). From **Figure 2**, it is clear that HAT in the aqueous solution occurs preferentially at C1'', O6, and O5' sites of the moracin C neutral form and O6 and O3' sites of the corresponding monoanion. For *iso*-moracin, HAT is favored at C3'', O6, O5', and O3' sites of the neutral form and at O6 and O5' sites of the monoanion one.

In the pentyl ethanoate solvent, where only the neutral species are present, the HAT process is favored, and the lowest  $\Delta G$  values are obtained for the OOH attack at the C1'' site followed by O6 for moracin C and at C3'' and O6 for *iso*-moracin C.

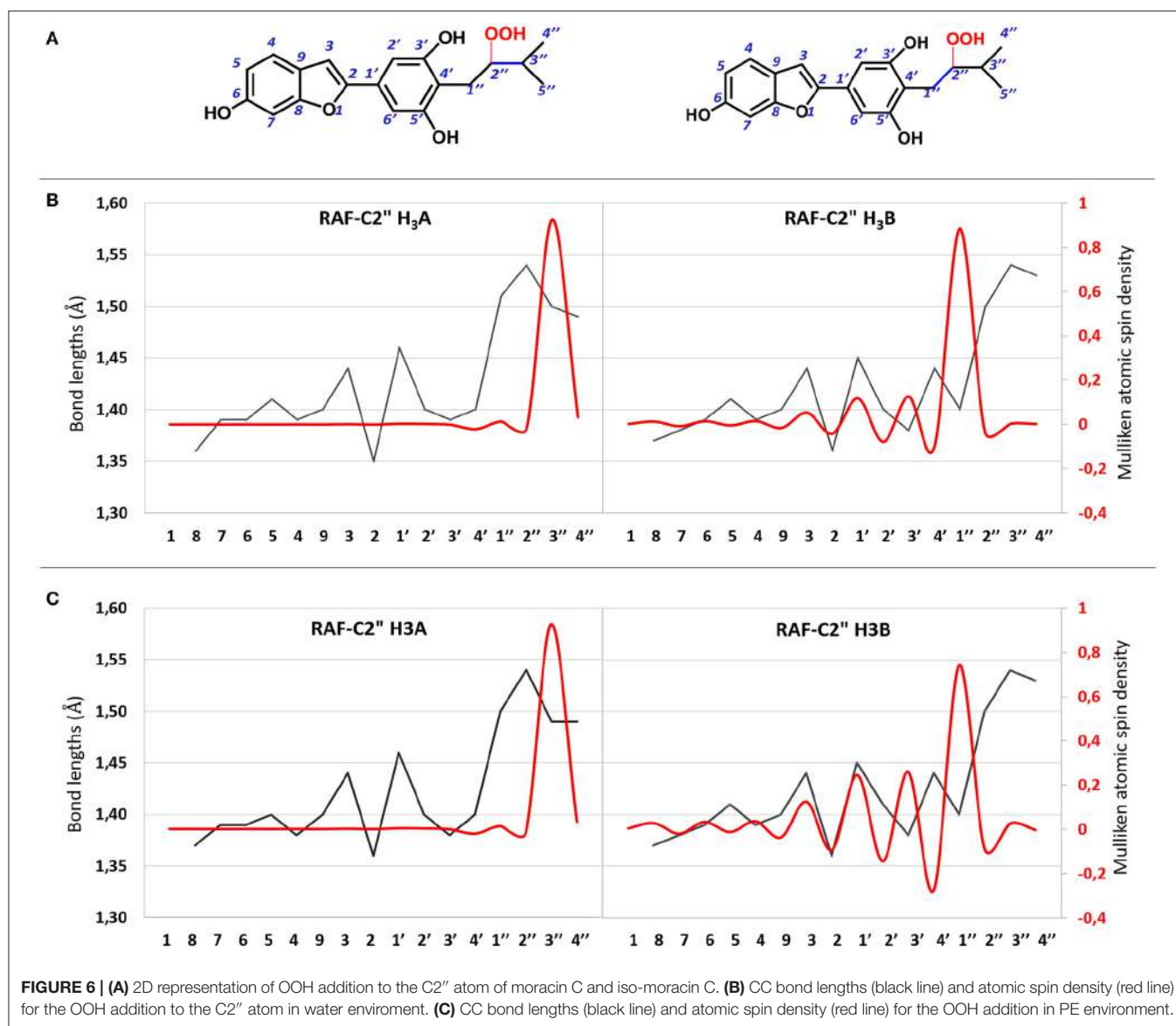
The radicals obtained following the abstraction of the proton by OOH free radical have a spin density that is distributed over almost the entire molecular structure, as it is reported in **Figure 4** that the spin density plots of moracin C deprotonated on C1'' and *iso*-moracin C deprotonated on C3.'' In particular, due to the C=C double bond proximity to the phenyl ring, electron delocalization appears to be slightly more extended in *iso*-moracin C. In any case, this trend suggests good stability of the formed radical species for both molecules.

For the processes that show exergonic, almost isergonic, and moderate endergonic behaviors, we have computed the kinetic constants. To do this, it was necessary to locate the TSs to obtain the activation energies for the given reaction mechanism. The structures of all TSs obtained for the HAT process and their relative imaginary frequencies are shown in **Figure 5** for neutral and monoanionic species in the aqueous environment, while the structures of TSs of the neutral systems in the lipid-like environment are shown in **Supplementary Figure 3**. The obtained energy barriers ( $\Delta G^\ddagger$ ) are reported in **Table 2** together with the Gibbs free energies of reaction.

Inspection of the last Table reveals that, in aqueous solution, the  $\Delta G^\ddagger$  values of moracin C fall in the range of 19–21 kcal/mol for the neutral forms and become slightly lower for the charged ones. A similar behavior can be noted for *iso*-moracin C species. In the pentyl ethanoate solvent, the result is different, and for some sites, the barriers are sensibly smaller (e.g., 6.08 kcal/mol for the C1'' site). We would like to underline that the C–H bonds of the 3-methylbut-2-enyl in moracin C and in *iso*-moracin C are involved in the HAT process, making these two natural molecules interesting antioxidant agents.

All attempts to locate the relevant TS for the radical attack to the C2'' site for both molecules failed. However, this is not unusual since this type of radical attack often occurs without energy barriers. The structures derived from the OOH radical attack on the C2'' atom for both molecules are reported in **Supplementary Figure 4**. The C=C bond variation and atomic spin density for the moracin C–OOH and *iso*-moracin C–OOH radical adduct in both considered environments are shown in **Figure 6**, and the corresponding values are reported in **Supplementary Table 2**. An inspection of **Figure 6** shows that the addition of the OOH radical on





**FIGURE 6 | (A)** 2D representation of OOH addition to the C2'' atom of moracin C and iso-moracin C. **(B)** CC bond lengths (black line) and atomic spin density (red line) for the OOH addition to the C2'' atom in water environment. **(C)** CC bond lengths (black line) and atomic spin density (red line) for the OOH addition in PE environment.

the C2'' atom induces different effects in the two tautomers. In fact, in moracin C, in both the considered solvents, the spin density is essentially located at C3'' and the bond length results of C2''-C3'' needs to be elongated by assuming values close to those of a single bond (1.513 and 1.491 Å in water and PE, respectively). In *iso*-moracin C, in the water environment, the addition of the radical in the same position induces a large spin density in the C1'' atom, a smaller but significant one in C3' and C1' atoms, and a very small density in C3 and C4 atoms. The C2''-C1'' distance is now 1.504 Å. This means that this radical, with a more extended spin density distribution result, would be more stable than the corresponding in moracin C. Similar relationships have previously been observed in other theoretical investigations on the antioxidant power of carotenoid derivatives (Ceron-Carrasco et al., 2010).

**TABLE 3 |** Ionization potential (IP), electron affinity (AE), electrodonating ( $\omega^-$ ), and electroaccepting ( $\omega^+$ ) indices of moracin C and *iso*-moracin C in water and PE (in parentheses) environments.

Molecule	IP	AE	$\omega^-$	$\omega^+$
Moracin C (H <sub>3</sub> A)	5.04 (4.98)	1.32 (1.12)	4.54 (4.18)	1.36 (1.13)
Iso-moracin C (H <sub>3</sub> B)	5.14 (4.79)	1.24 (1.25)	4.45 (4.31)	1.26 (1.29)

All values are in eV.

The computation of the electrodonating ( $\omega^-$ ) and electroaccepting ( $\omega^+$ ) values, as proposed by Gázquez et al. (2007), allows the verification of the possible correlation between these reactivity indices and the RAF antioxidant capability of the investigated molecule. Results are shown in Table 3. Since low values of  $\omega^-$  indicate greater antioxidant activity, the

**TABLE 4** | Rate constants ( $M^{-1}s^{-1}$ ) and branching ratios ( $\Gamma$ ) computed at the M05-2x level of theory at 298.15 K, **(A)** in aqueous and **(B)** in pentyl ethanoate solvent.

Mechanism	$H_3A$		$H_2A^-$		$H_3B$		$H_2B^-$	
	$k (M^{-1}s^{-1})$	$\Gamma (%)$	$k (M^{-1}s^{-1})$	$\Gamma (%)$	$k (M^{-1}s^{-1})$	$\Gamma (%)$	$k (M^{-1}s^{-1})$	$\Gamma (%)$
<b>(A)</b>								
SET	$1.03 \times 10^{-8}$	~0.00	$1.83 \times 10^9$	100.0	$1.08 \times 10^{-9}$	~0.00	$8.23 \times 10^8$	100.0
HAT-O6	$4.57 \times 10^2$	~0.00	$4.45 \times 10^2$	~0.0	$2.45 \times 10^2$	~0.00	$9.74 \times 10^2$	~0.0
HAT-O3'			$7.49 \times 10^7$	~0.0	$2.79 \times 10^9$	~0.00		
HAT-O5'	$7.49 \times 10^1$	~0.00			$4.39 \times 10^{-2}$	~0.00	$3.97 \times 10^1$	~0.00
HAT-C1''	$2.99 \times 10^2$	~0.00						
HAT-C3''					$1.09 \times 10^3$	~0.00		
RAF-C2''	$2.15 \times 10^9$	100.00			$2.15 \times 10^9$	100.00		
Total	$2.15 \times 10^9$		$1.83 \times 10^9$		$2.15 \times 10^9$		$8.23 \times 10^8$	
Overall	$2.11 \times 10^9$		$2.93 \times 10^7$		$2.13 \times 10^9$		$9.88 \times 10^6$	
Mechanism	$H_3A^{PE}$		$H_3B^{PE}$					
	$k (M^{-1}s^{-1})$	$\Gamma (%)$	$k (M^{-1}s^{-1})$	$\Gamma (%)$				
<b>(B)</b>								
HAT-O6	$8.71 \times 10^2$	~0.00	$1.94 \times 10^4$	~0.00				
HAT-O5'			$3.42 \times 10^1$	~0.00				
HAT-C1''	$2.88 \times 10^9$	56.68						
HAT-C3''			$1.57 \times 10^6$	0.07				
RAF-C2''	$2.20 \times 10^9$	43.32	$2.22 \times 10^9$	99.93				
Total	$5.08 \times 10^9$		$2.22 \times 10^9$					
Overall	$4.98 \times 10^9$		$2.20 \times 10^9$					

analysis of **Table 3** shows how *iso*-moracin C seems to have greater scavenging power in the aqueous environment. On the contrary, in the PE solvent, the antioxidant action of moracin C is greater. Considering the average of the values obtained in the two solvents as previously suggested by some authors (Ceron-Carrasco et al., 2012), values of  $\omega^-$  being very close to each other are obtained (4.36 and 4.38 eV for moracin C and *iso*-moracin C, respectively), making it difficult to reliably predict their correlation with the antioxidant activity of the two molecules. The calculation of the kinetic constants can shed further light on the antioxidant activity of the two systems.

Using the data from **Table 2** and following the QM-ORSA computational protocol (Galano and Alvarez-Idaboy, 2013), we computed the individual, as well as the total kinetic, constants that are reported in **Table 4**.

For neutral moracin C ( $H_3A$ ) in the water medium, the faster process is the RAF mechanism in the C2'' site ( $k = 2.15 \times 10^9 M^{-1}s^{-1}$ ; branching ratio  $\Gamma = 100\%$ ), while for the corresponding monoanion ( $H_2A^-$ ), the SET mechanism is the faster process ( $k = 1.83 \times 10^9 M^{-1}s^{-1}$ , branching ratio  $\Gamma = 100\%$ ). In the *iso*-moracin neutral system ( $H_3B$ ), the calculated kinetic constants for RAF and HAT mechanisms are similar ( $k = 2.79 \times 10^9 M^{-1}s^{-1}$  for HAT on the O3' site and  $k = 2.15 \times 10^9 M^{-1}s^{-1}$  for RAF on the C2'' atom), but the branching ratio for the former is 100%. For both the H2X- species, the SET mechanism is preferred ( $k = 1.83 \times 10^9 M^{-1}s^{-1}$  for moracin and  $8.23 \times 10^8 M^{-1}s^{-1}$  for *iso*-moracin).

In the PE environment, for  $H_3B$ , RAF is the preferred mechanism on the C2'' site with a  $k$  value of  $2.22 \times 10^9 M^{-1}s^{-1}$ , while, for  $H_3A$ , a competition between RAF on C2'' ( $k = 2.20 \times 10^9 M^{-1}s^{-1}$ ;  $\Gamma = 43.32\%$ ) and HAT on the C1'' site ( $k = 2.88 \times 10^9 M^{-1}s^{-1}$ ;  $\Gamma = 56.68\%$ ) mechanisms was found.

From the obtained individual and total kinetic constants, it is clear that the experimental mass spectrometric suggestion, according to which the RAF mechanism is best possible solution (Li et al., 2018), is theoretically confirmed. The data indicate that the scavenging activity in the water solution of both moracin C and *iso*-moracin neutral forms is carried out through the RAF mechanism. In the lipid-like environment, the situation appears to be different since mainly the non-redox RAF reaction on C3'' site can occur through the OOH attacking *iso*-moracin C, while moracin C can undergo the attack through both redox (SET)- and non-redox- (RAF) like reactions.

## CONCLUSION

From the density functional computations on the antioxidant potential of moracin C and its isomer *iso*-moracin C, the following conclusions can be outlined:

- pKa calculations in water environment evidence that, for both systems, the neutral form is dominant with the

- monoanionic species being present in lower percentage but not negligible;
- the preferred atomic sites for the different reaction mechanisms were established;
  - the attack of the OOH free radical, for both the isomers, on the most abundant neutral species in a water solvent mainly occurs through a radical addition mechanism;
  - for *iso*-moracin C, the radical addition process is favored in the lipid-like environment, while, for moracin C, both redox- and non-redox-type reactions can occur equally.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author.

## REFERENCES

- Ahmadi, S., Marino, T., Prejanò, M., Russo, N., and Toscano, M. (2018). Antioxidant properties of the Vam3 derivative of resveratrol. *Molecules* 23:2446. doi: 10.3390/molecules23102446
- Alberto, M. E., Grand, A., Russo, N., and Galano, A. (2013). A physicochemical examination of the free radical scavenging activity of Trolox: mechanism, kinetics and influence of the environment. *Phys. Chem. Chem. Phys.* 15:4642. doi: 10.1039/c3cp43319f
- Castaneda-Arriaga, R., Marino, T., Russo, N., Alvarez-Idaboy, J. R., and Galano, A. (2020). Chalcogen effects on the primary antioxidant activity of chrysin and quercetin. *New J. Chem.* 44, 9073–9082. doi: 10.1039/D0NJ01795G
- Ceron-Carrasco, J., Bastida, A., Requena, A., and Zuniga, J. (2010). A theoretical study of the reaction of beta-carotene with the nitrogen dioxide radical in solution. *J. Phys. Chem. B* 114, 4366–4372. doi: 10.1021/jp911846h
- Ceron-Carrasco, J., Zuniga, J., Bastida, A., and Requena, A. (2012). Antioxidant properties of beta-carotene isomers and their role in photosystems: insights from Ab initio simulations. *J. Phys. Chem. A* 116, 3498–3506. doi: 10.1021/jp301485k
- Collins, F. C., and Kimball, G. E. (1949). Diffusion in chemical reaction processes and in the growth of colloid particles. *J. Colloid Sci.* 4:425. doi: 10.1016/0095-8522(49)90023-9
- Delogu, G. L., Matos, M. J., Fanti, M., Era, B., Medda, R., Pieroni, E., et al. (2016). 2-Phenylbenzofuran derivatives as butyrylcholinesterase inhibitors: synthesis, biological activity and molecular modeling. *Bioorg. Med. Chem. Lett.* 26, 2308–2313. doi: 10.1016/j.bmcl.2016.03.039
- Fashing, P. J. (2001). Feeding ecology of Guerezas in the Kakamega forest Kenya: the importance of Moraceae fruit in their diet. *Int. J. Primatol.* 22, 579–609. doi: 10.1023/A:1010737601922
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2009). *Gaussian 09*. Pittsburgh, PA: Gaussian Inc.
- Galano, A., and Alvarez-Idaboy, J. R. (2013). A computational methodology for accurate predictions of rate constants in solution: application to the assessment of primary antioxidant activity. *J. Comput. Chem.* 34, 2430–2445. doi: 10.1002/jcc.23409
- Galano, A., and Alvarez-Idaboy, J. R. (2019). Computational strategies for predicting free radical scavengers' protection against oxidative stress: where are we and what might follow? *Int. J. Quantum. Chem.* 119:25665. doi: 10.1002/qua.25665
- Galano, A., Mazzone, G., Alvarez-Diduk, R., Marino, T., Alvarez-Idaboy, J. R., and Russo, N. (2016). Food antioxidants: chemical insights at the molecular level. *Annu. Rev. Food Sci. Technol.* 7, 335–352. doi: 10.1146/annurev-food-041715-033206
- Gázquez, J. L., Cedillo, A., and Vela, A. (2007). Electrodonating and electroaccepting powers. *J. Phys. Chem. A* 111, 1966–1970. doi: 10.1021/jp065459f

## AUTHOR CONTRIBUTIONS

AP, BD, TM, MT, and NR made equal contributions to the study and the publication of this work. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation (Grant no. MAE00643232020-06-16).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2021.666647/full#supplementary-material>

- Hakim, E. H., Fahriyati, A., Kau, M. S., Achmad, S. A., Makmur, L., Ghisalberti, E. L., et al. (1999). Artoindonesianins A and B, two new prenylated flavones from the root of *Artocarpus champeden*. *J. Nat. Prod.* 62, 613–615. doi: 10.1021/np980279l
- Jeon, S. Y., Kwon, S. H., Seong, Y. H., Bae, K., Hur, J. M., Lee, Y. Y., et al. (2007). beta-secretase (BACE1)-inhibiting stilbenoids from *Smilax Rhizoma*. *Phytomedicine* 14, 403–408. doi: 10.1016/j.phymed.2006.09.003
- Kapche, D., Lekane, N. M., Kulabas, S. S., Ipek, H., Tok, T. T., Ngadjui, B. T., et al. (2017). Aryl benzofuran derivatives from the stem bark of *Calpocalyx dinklagei* attenuate inflammation. *Phytochemistry* 141, 70–79. doi: 10.1016/j.phytochem.2017.05.007
- Kapche, G. D. W. F., Fozing, C. D., Donfack, J. H., Fotso, G. W., Amadou, D., Tchana, A. N., et al. (2009). Prenylatedarylbenzofuran derivatives from *Morus mesozygia* with antioxidant activity. *Phytochemistry* 70, 216–221. doi: 10.1016/j.phytochem.2008.12.014
- Kuete, V., Fozing, D. C., Kapche, W. F. G. D., Mbavengd, A. T., Kuiatea, J. R., Ngadjuib, B. T., et al. (2009). Antimicrobial activity of the methanolic extract and compounds from *Morus mesozygia* stem bark. *J. Ethnopharmacol.* 124, 551–555. doi: 10.1016/j.jep.2009.05.004
- Lee, J. H., Ko, H. J., Woo, E. R., Lee, S. K., Moon, B. S., Lee, C. W., et al. (2016). Moracin M inhibits airway inflammation by interrupting the JNK/ c-Jun and NF-κB pathways *in vitro* and *in vivo*. *Eur. J. Pharmacol.* 783, 64–72. doi: 10.1016/j.ejphar.2016.04.055
- Leopoldini, M., Russo, N., and Toscano, M. (2011). The molecular basis of working mechanism of natural polyphenolic antioxidants. *Food Chem.* 125, 288–306. doi: 10.1016/j.foodchem.2010.08.012
- Li, X., Xie, H., Zhan, R., and Chen, D. (2018). Effect of double bond position on 2-phenyl-benzofuran antioxidants: a comparative study of moracin C and iso-moracin C. *Molecules* 23:754. doi: 10.3390/molecules23040754
- Marcus, R. (1957). On the theory of oxidation-reduction reactions involving electron transfer. III. Applications to data on the rates of organic redox reactions. *J. Chem. Phys.* 26:872. doi: 10.1063/1.1743424
- Marenich, A. V., Cramer, C. J., and Truhlar, D. G. (2009). Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* 113, 6378–6396. doi: 10.1021/jp810292n
- Markovic, Z., Tošovic, J., Milenkovic, D., and Markovic, S. (2016). Revisiting the solvation enthalpies and free energies of the proton and electron in various solvents. *Comput. Theor. Chem.* 1077, 11–17. doi: 10.1016/j.comptc.2015.09.007
- Mazzone, G., Malaj, N., Galano, A., Russo, N., and Toscano, M. (2015). Antioxidant properties of several coumarin–chalcone hybrids from theoretical insights. *RSC Adv.* 5, 565–575. doi: 10.1039/C4RA11733F
- Na, M., Hoang, D. M., Njamen, D., Mbafor, J. T., Fomum, Z. T., Thuong, P. T., et al. (2007). Inhibitory effect of 2-arylbenzofurans from *Erythrina addisoniae*

- on protein tyrosine phosphatase-1B. *Bioorg. Med. Chem. Lett.* 17, 3868–3871. doi: 10.1016/j.bmcl.2007.05.005
- Naik, R., Harmalkar, D. S., Xu, X., Jang, K., and Lee, K. (2015). Bioactive benzofuran derivatives: moracins A-Z in medicinal chemistry. *Eur. J. Med. Chem.* 90, 379–393. doi: 10.1016/j.ejmech.2014.11.047
- Nguyen, T. D., Jin, X., Lee, K., Hog, Y. S., Young, H. K., and Jung, J. L. (2009). Hypoxia-inducible factor-1 inhibitory benzofurans and chalcone-derived diels-alder adducts from *Morus* species. *J. Nat. Prod.* 72, 39–43. doi: 10.1021/np800491u
- Pel, P., Chae, H. S., Nhoek, P., Kim, Y. M., and Chin, Y. W. (2017). Chemical constituents with proprotein convertase subtilisin/kexin type 9 mRNA expression inhibitory activity from dried immature *Morus alba* fruits. *J. Agric. Food Chem.* 65, 5316–5321. doi: 10.1021/acs.jafc.7b02088
- Pérez-González, A., Castañeda-Arriaga, R., Verastegui, B., Carreón-González, M., Alvarez-Idaboy, J. R., and Galano, A. (2018). Estimation of empirically fitted parameters for calculating pK<sub>a</sub> values of thiols in a fast and reliable way. *Theor. Chem. Acc.* 137:5. doi: 10.1007/s00214-017-2179-7
- Pérez-González, A., García-Hernández, E., and Chigo-Anota, E. (2020). The antioxidant capacity of an imidazole alkaloids family through single-electron transfer reactions. *J. Mol. Model.* 26:321. doi: 10.1007/s00894-020-04583-2
- Romeo, I., Parise, A., Galano, A., Russo, N., Alvarez-Idaboy, J. R., and Marino, T. (2020). The antioxidant capability of higenamine: insights from theory. *Antioxidants* 9:358. doi: 10.3390/antiox9050358
- Sang-Hee, L., Sang-Yoon, C., Hocheol, K., Jae-Sung, H., Byeong-Gon, L., Jian-Jun, G., et al. (2002). Mulberroside F isolated from the leaves of *Morus alba* inhibits melanin biosynthesis. *Biol. Pharm. Bull.* 25, 1045–1048. doi: 10.1248/bpb.25.1045
- Seong, S. H., Ha, M. T., Min, B. S., Jung, H. A., and Choi, J. S. (2018). Moracin derivatives from *Morus radix* as dual BACE1 and cholinesterase inhibitors with antioxidant and anti-glycation capacities. *Life Sci.* 210, 20–28. doi: 10.1016/j.lfs.2018.08.060
- Truhlar, D. G., Garrett, B. C., and Klippenstein, S. J. (1996). Current status of transition-state theory. *J. Phys. Chem.* 100:12771. doi: 10.1021/jp953748q
- Venkatesh, K. R., and Seema, C. (2008). Mulberry: life enhancer. *J. Med. Plants Res.* 2, 271–278. doi: 10.5897/JMPR.9000005
- Zelová, H., Hanáková, Z., Cermáková, Z., Šmejkal, K., Dall, A. S., Babula, P., et al. (2014). Evaluation of anti-inflammatory activity of prenylated substances isolated from *Morus alba* and *morusnigra*. *J. Nat. Prod.* 77, 1297–1303. doi: 10.1021/np401025f
- Zhao, Y., Schultz, N. E., and Truhlar, D. G. (2006). Assessment of model chemistries for noncovalent interactions. *J. Chem. Theory Comput.* 2, 364–382. doi: 10.1021/ct0502763

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Parise, De Simone, Marino, Toscano and Russo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## Abbreviations and Acronyms

3CL <sup>pro</sup>	3-Chymotrypsin-like protease
AMBER	Assisted Model Building and Energy Refinement
AP1	Arsenoplatin-1
B3LYP	Becke , 3-parameter, Lee- Yang - Parr functional
CHARMM	Chemistry at HARvard Macromolecular Mechanics
CTP	Cytidine triphosphate
ddhCTP	3'-deoxy-3',4'-didehydro-CTP
deMon2k	density of Montréal
DFT	Density Functional Theory
DNA	Deoxyribonucleic acid
E	Enzyme
EBS	Ebselen
EBS-OH	Benzisoselenazol-3(2H)-one ebselen
EC	Enzyme Commission number
EI	Enzyme-inibitor non-covalent complex
E-I	Enzyme-inibitor covalent complex
EP	Enzyme-product complex
ES	Enzyme-substrate complex
ESI	Enzyme-substrate-inhibitor complex
FF	Force Field
FFT	Fast Fourier Transform
GGA	Generalized Gradient Approximation
GROMOS	GRONingen MOlecular Simulation
HAT	Hydrogen-atom transfer
HEWL	Hen Egg White Lysozyme
HF	Hartree-Fock
HIV	Human Immunodeficiency Virus
I	Inhibitor
K <sub>m</sub>	Michaelis-Menten constant
KS	Kohn-Sham
LSDA	Local Spin Density Approximation
MD	Molecular Dynamics
MM	Molecular Mechanics
MM-PBSA	Poisson-Boltzmann Surface Area Continuum Solvation
M <sup>pro</sup>	Main protease
NADH	Nicotinamide Adenine Dinucleotide
ONIOM	our own N-layered integrated molecular orbital and molecular mechanics
P	Product
PBC	Periodic Boundary Condition
PBE	Perdew, Burke, Ernzerhof functional
PCET	Proton-coupled electron transfer
PIPF	Polarizable Intermolecular Potential Function

PL <sup>pro</sup>	Papain-like protease
QM	Quantum Mechanics
QM-ORSA	Quantum Mechanics-based Test for Overall Free Radical Scavenging Activity
RAF	Radical adduct formation
RdRp	RNA-dependent RNA polymerase
RNA	Ribonucleic Acid
RNase	Ribonuclease
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
S	Substrate
SAM	S-Adenosyl-methionine
SARS-CoV-2	Severe Acute Aespiratory Syndrome CoronaVirus 2
SCF	Self Consistent Field
SET	Single electron transfer
SM	Shape Matching
SO	Stress
SPLET	Sequential proton-loss electron transfer
TS	Transition State
UFF	Universal Force Field

### ***List of amino acids***

Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
Cys	Cysteine
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Pyl	Pyrrolysine
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine



## **Addendum**



## Supplementary Information

### The generation of the oxidant agent of a mononuclear nonheme Fe(II) biomimetic complex by oxidative decarboxylation. A DFT investigation

Angela Parise, Maria Costanza Muraca, Nino Russo, Marirosa Toscano and Tiziana Marino,

*Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Cubo I4C, Via P. Bucci, 87036 Arcavacata di Rende, CS, Italy;*

Table S1. Calculated spin contamination for all investigated complexes under all three spin states co spin states examined.

	<b>3</b>	<b>5</b>	<b>7</b>
<b>a</b>	2.0442	6.0127	12.0233
adduct	3.5769	6.6643	12.0201
TS1	2.0210	6.6945	12.0188
hydroperoxo	2.0294	6.1325	12.0182
TS2	2.0361	6.1380	-
Prod	2.0334	6.0156	12.0297

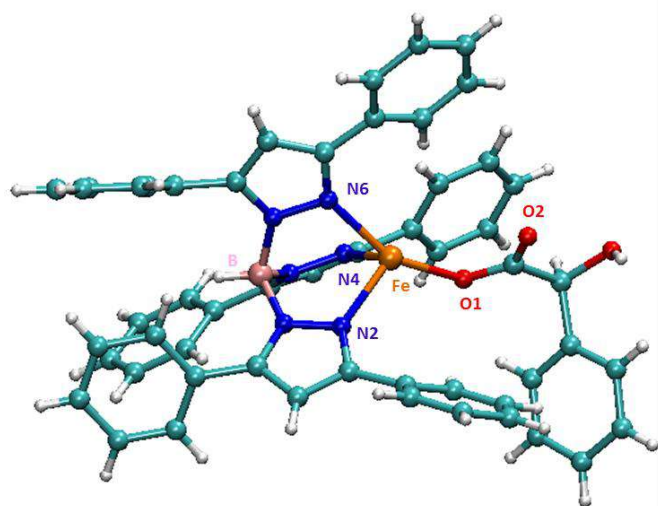
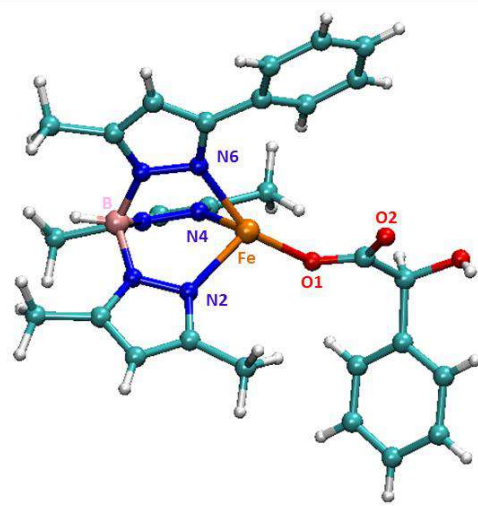
**X-Ray****Model  
(a)**

Figure S1 Crystal structure of hydrotris(3,5-diphenylpyrazole-1-yl)borate) Fe(II)-mandelate and its computational model (a) used in this study.

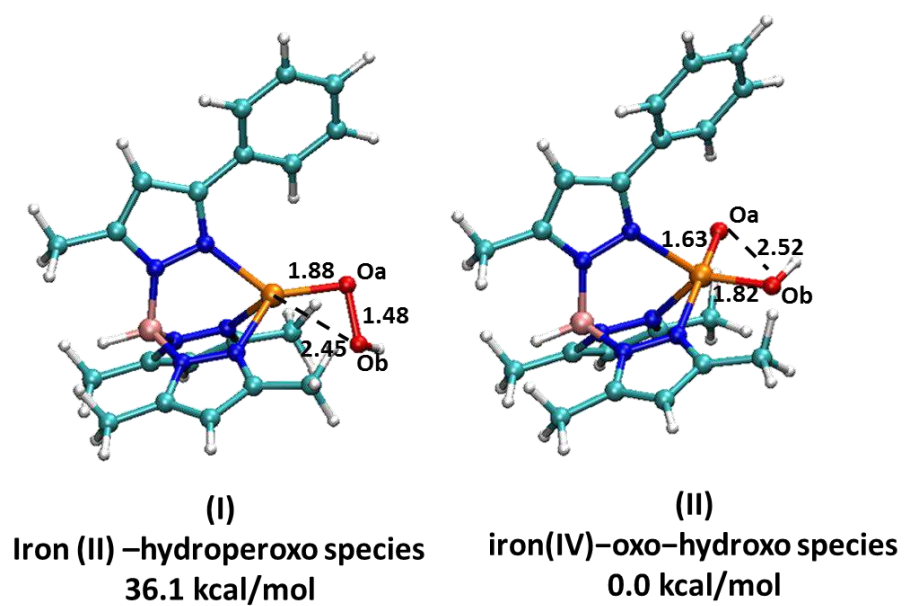


Figure S2 The two possible forms of the product of the oxidative decarboxylation process in quintet species.

## SUPPORTING INFORMATION

# **Platination mechanism of RNase A by Arsenoplatin: Insights from theory**

A. Parise<sup>ab</sup>, N. Russo<sup>a</sup> and T. Marino<sup>a\*</sup>

<sup>a</sup> Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy

<sup>b</sup> Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France

\* Correspondence: [tiziana.marino65@unical.it](mailto:tiziana.marino65@unical.it)



**Table S1** NBO charge values in |e| for every stationary point along the potential energy surface.

	<b>APw</b>	<b>EAPw</b>	<b>TS</b>	<b>P</b>
<b>Pt</b>	0.200	0.214	0.145	0.139
<b>As</b>	2.137	2.142	2.179	2.113
<b>O<sub>w</sub></b>	-0.841	-0.844	-0.906	-0.962
<b>N<sub>His</sub></b>	-	-0.533	-0.583	-0.514
<b>N<sub>1</sub></b>	-0.715	-0.712	-0.703	-0.707
<b>N<sub>2</sub></b>	-0.720	-0.725	-0.715	-0.709
<b>O<sub>1</sub></b>	-0.759	-0.763	-0.763	-0.774
<b>O<sub>2</sub></b>	-0.757	-0.759	-0.762	-0.767
<b>O<sub>3</sub></b>	-0.966	-0.977	-0.981	-1.024
<b>O<sub>4</sub></b>	-0.967	-0.969	-0.979	-0.972

MD simulation. We started from enzyme structure ((PDB id: 5NJ7) [1] prepared at physiological pH using PROPKA. [2] After removing water and co-crystallized ligand, the molecular dynamics simulation was carried out with only free protein. Antechamber modules of Amber16 [3] were used for generating preparatory files. MD simulation of 300 ns was performed by using FF14SB. [4] The free protein was protonated and counter ions (3Cl<sup>-</sup>) were added appropriately to make the total charge zero. The system was solvated in an orthorhombic box with a buffer of 10 Å, using TIP3P water model. The solvated structure was first minimized by applying positional harmonic restraints on all atoms (50 kcal mol<sup>-1</sup> Å<sup>2</sup>) using 5000 steps of steepest descent followed by 5000 steps of conjugate gradient (CG). In the second minimization step, the whole system was released without any restraint and then a progressive heating phase was carried out from 0 to 310 K for 50 ps, followed by 50 ps at 310 K using the Langevin thermostat in NVT ensemble. The production phase was performed for 100 ns of MDs under the following conditions: integration step of 2 fs coupling SHAKE algorithm [5],

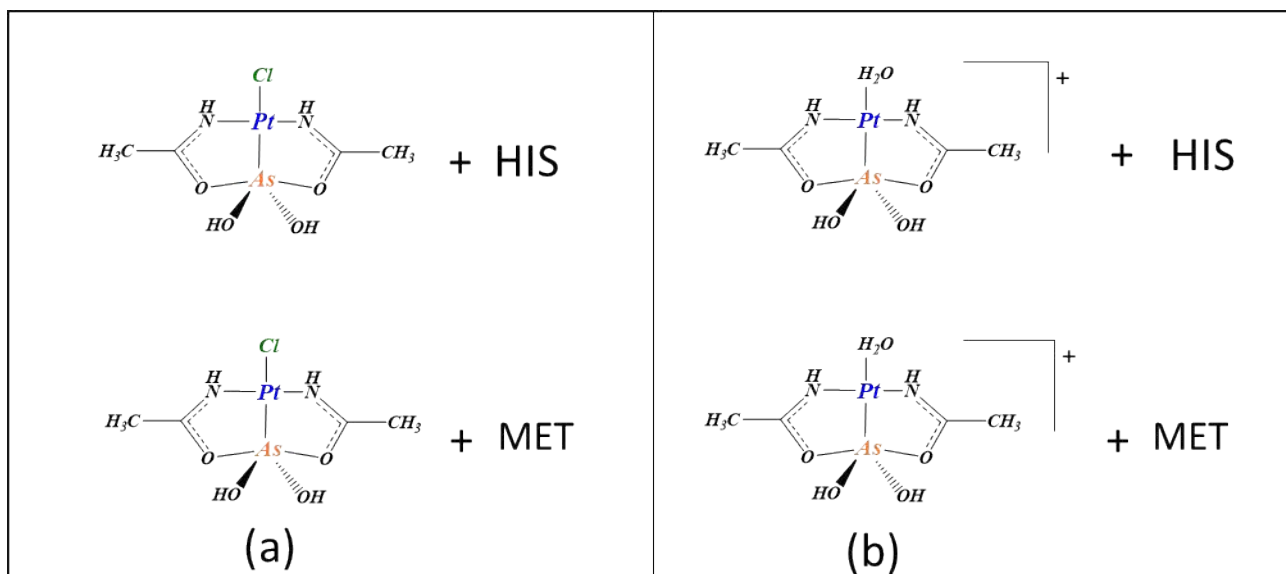
NPT ensemble at 1 bar pressure using the Berendsen barostat with a time constant  $\tau_p = 2.0$  ps. The Particle mesh Ewald summation method was employed for the electrostatic potential and the long-range electrostatic interactions were calculated with 12 Å cut-off distance.

**Table S2:** Calculated pKa for ionizable residues of RNase A. Residues fully protonated (positively charged) or deprotonated (negatively charged) are highlighted in blue and red, respectively. Residues belonging to the catalytic triad are underlined

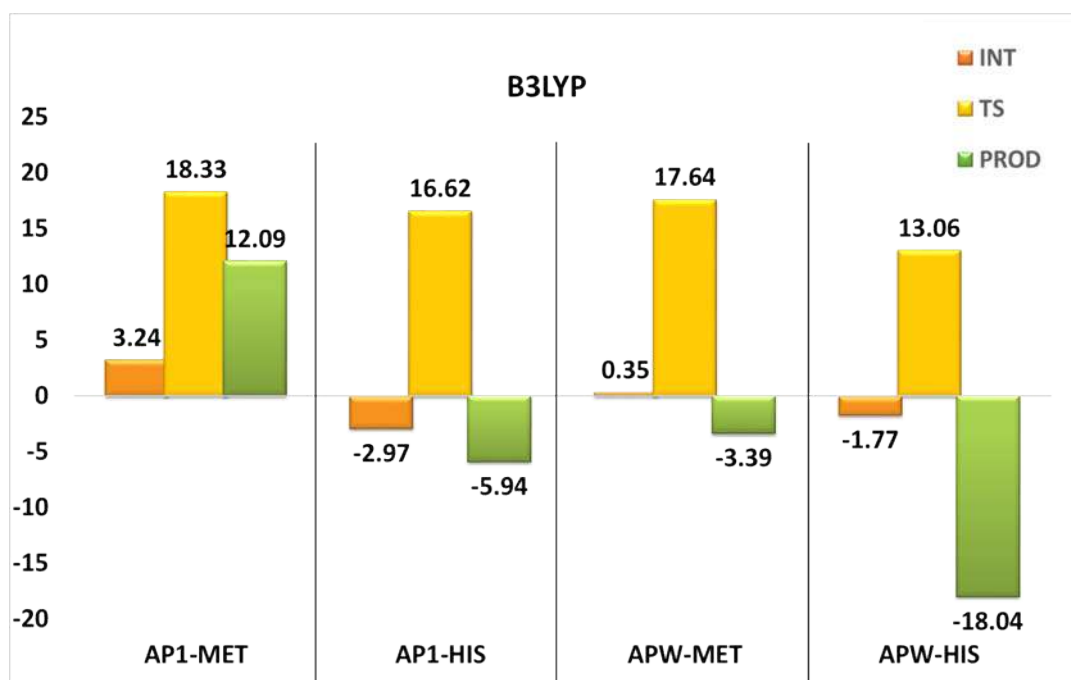
<i>Residue</i>	<i>pKa</i>	<i>Residue</i>	<i>pKa</i>
<b>GLU2</b>	2.87	TYR25	13.51
<b>GLU9</b>	4.71	TYR73	11.18
<b>ASP14</b>	0.12	TYR76	10.83
<b>ASP38</b>	4.22	TYR92	10.67
<b>GLU49</b>	4.04	TYR97	11.48
<b>ASP53</b>	4.70	TYR115	9.17
<b>ASP83</b>	3.88	<b>LYS31</b>	10.50
<b>GLU86</b>	4.05	<b>LYS37</b>	10.50
<b>GLU111</b>	4.57	<b>LYS41</b>	9.10
<b>ASP121</b>	2.62	<b>LYS61</b>	10.29
HIS12	1.37	<b>LYS66</b>	10.29
HIS48	5.02	<b>LYS91</b>	10.50
HIS105	7.19	<b>LYS98</b>	10.15
<b>HIS119</b>	5.83	<b>LYS104</b>	10.15
CYS26	-	<b>LYS7</b>	10.43
CYS40	-	<b>ARG10</b>	11.73
CYS58	-	<b>ARG33</b>	13.00
CYS65	-	<b>ARG39</b>	12.22
CYS72	-	<b>ARG85</b>	12.36
CYS84	-		
CYS95	-		
CYS110	-		

**Table S3.** H-bond intramolecular in apo-protein RNase A during MD simulation 100 ns.

<b>Acceptor</b>	<b>Donor</b>	<b>Freq (%)</b>	<b><math>\bar{d}</math> (Å)</b>
PHE-8	HIS-12	68.48	2.869
LYS-7	GLN-11	48.89	2.879
ALA-5	GLU-9	43.98	2.881
ALA-6	ARG-10	25.91	2.891
ALA-4	PHE-8	24.62	2.885
THR-3	LYS-7	18.82	2.875
THR-3	ALA-6	13.23	2.896
ALA-4	LYS-7	8.49	2.897
ALA-4	ALA-6	2.11	2.863
LYS-7	ARG-10	1.67	2.911
ALA-5	PHE-8	1.14	2.896
PHE-8	GLN-11	1.01	2.921
ALA-6	GLU-9	0.22	2.893
LYS-1	THR-3	0.19	2.861
THR-3	ALA-5	0.18	2.851
HIS-119	ASP-121	0.09	2.901
ALA-5	LYS-7	0.08	2.882
GLU-2	ALA-4	0.08	2.871
GLU-9	HIS-12	0.04	2.891
PHE-8	ARG-10	0.01	2.731



Scheme S1.



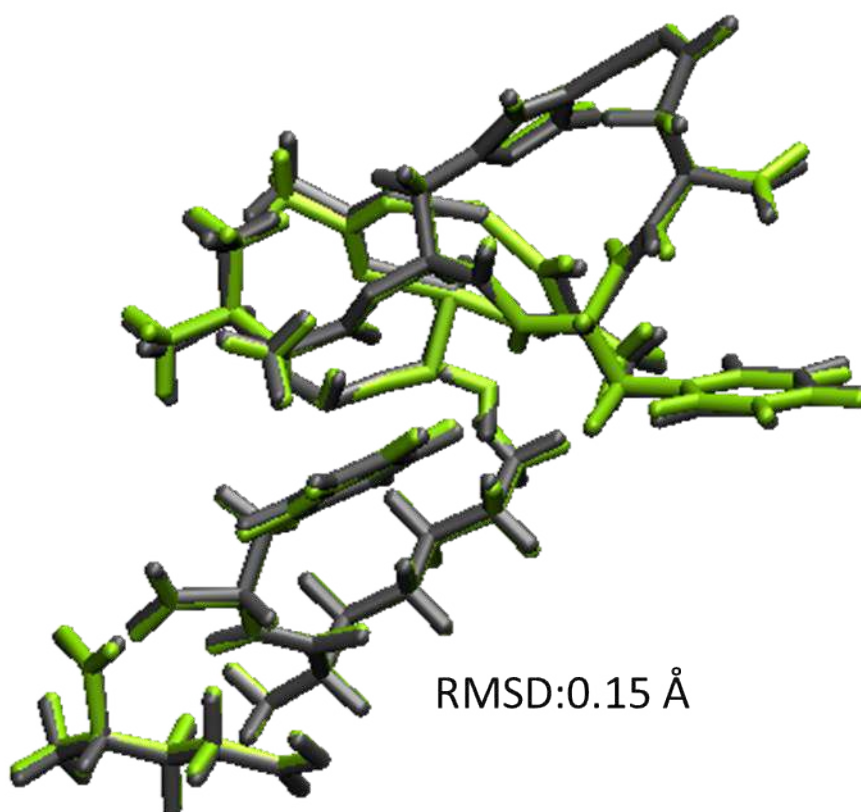
a)



b)

Figure S1

Prod\_B3LYP + D3 (green)  
Prod\_M062X (Grey)



Overlap of the two structures  
optimized in water with  
B3LYP+D3 and M062X

Figure S2



## References

- [1] Đ. Miodragović, A. Merlino, E. P. Swindell, A. Bogachkov, R. W. Ahn, S. Abuhadba, G. Ferraro, T. Marzo, A. P. Mazar, L. Messori and T. V. O'Halloran, *J. Am. Chem. Soc.*, **2019**, 141, 6453–6457.
- [2] D. C. Bas, D. M. Rogers, J. H. Jensen, *Proteins*, **2008**, 73, 765–783.
- [3] D. A. Case, D. S. Cerutti, T. E. III Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman. **2017**, AMBER 2017, University of California, San Francisco.
- [4] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simemrlin. *J. Chem. Theory Comput.* **2015**, 11, 3696–3713.
- [5] J-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, 23, 327-341.

## Coordinate of optimized structure

EAPw

N	11.03900000	4.88600000	101.11200000
H	10.74400000	3.99000000	101.47800000
H	11.72200000	5.28300000	101.74400000
H	10.23900000	5.50100000	101.04600000
C	11.65500000	4.69800000	99.76700000
H	10.88900000	4.40100000	99.05000000
C	12.36000000	5.97600000	99.29800000
H	11.76300000	6.83500000	99.60900000
H	13.31600000	6.03900000	99.82300000
C	12.57500000	6.10500000	97.79700000
H	12.94400000	5.16300000	97.39800000
H	11.62900000	6.33000000	97.30300000
C	13.62800000	7.17000000	97.49900000
H	14.22700000	7.38800000	98.38400000
H	14.31500000	6.78800000	96.74100000
C	13.02600000	8.47600000	96.99300000
H	12.16800000	8.74800000	97.60900000
H	13.77800000	9.26300000	97.07300000
N	12.61800000	8.34900000	95.56600000
H	11.91500000	7.63000000	95.47000000
H	12.24000000	9.22800000	95.23900000
H	13.42000000	8.10800000	95.00000000
C	12.69300000	3.60400000	99.90200000
O	13.14700000	3.33900000	101.02600000
N	13.08100000	2.99900000	98.77300000
C	11.74800000	1.02800000	99.33800000
C	14.21600000	1.26400000	100.21800000
C	14.63800000	-0.20900000	100.21500000
C	15.63700000	-0.54700000	101.30800000
O	15.94500000	0.23300000	102.19300000
O	16.16800000	-1.77400000	101.27600000
N	15.51600000	-3.73900000	99.47700000
C	14.08900000	-4.06700000	99.67500000
C	13.70700000	-5.55600000	99.59300000
O	12.81400000	-6.00200000	100.31700000
C	13.24400000	-3.24200000	98.68700000
C	11.80100000	-3.00200000	99.13600000
C	11.07800000	-2.12400000	98.11400000
C	9.69100000	-1.69000000	98.57900000
N	9.22200000	-0.53700000	97.72800000
N	14.34500000	-6.30800000	98.65700000
C	14.28600000	-7.77500000	98.65000000
C	15.49100000	-8.27100000	97.79700000
O	15.55400000	-9.52600000	97.59200000
C	12.90600000	-8.24100000	98.13200000
C	12.57100000	-9.72000000	98.14800000
C	11.60600000	-10.18400000	97.23500000
C	13.14100000	-10.64200000	99.03600000
C	11.24000000	-11.53500000	97.19700000
C	12.76700000	-11.98600000	99.00500000
C	11.82200000	-12.44300000	98.08200000

N	14.30100000	-9.97600000	95.36000000
C	14.60800000	-8.66500000	94.68000000
C	13.43100000	-7.91600000	94.06500000
O	13.63900000	-6.72500000	93.77800000
C	15.77600000	-8.81900000	93.67600000
C	17.04500000	-9.30100000	94.39600000
C	15.41500000	-9.69400000	92.46800000
N	12.25200000	-8.52500000	93.86400000
C	11.09100000	-7.83200000	93.31400000
C	10.08200000	-7.51900000	94.43300000
O	9.74800000	-8.38400000	95.24400000
C	10.39900000	-8.71600000	92.24200000
C	9.07300000	-8.13800000	91.86700000
N	8.68813860	-6.88544302	92.14588313
C	7.95598507	-8.80336565	91.42729987
C	7.42792031	-6.72259564	91.87837171
N	6.92648672	-7.89030662	91.42267836
N	9.54800000	-6.28100000	94.38600000
C	8.52700000	-5.82400000	95.33000000
C	7.66200000	-4.81500000	94.56600000
O	8.03100000	-3.62400000	94.44000000
C	9.20900000	-5.22200000	96.57300000
C	8.32800000	-4.97500000	97.77800000
C	7.09200000	-4.31500000	97.68800000
C	8.79500000	-5.35300000	99.04700000
C	6.35500000	-4.03400000	98.84200000
C	8.06300000	-5.06400000	100.20000000
C	6.83800000	-4.39800000	100.10200000
N	6.55400000	-5.24400000	93.97500000
C	5.72300000	-4.31100000	93.21200000
C	5.05900000	-3.22000000	94.12300000
O	4.91500000	-3.47900000	95.34200000
C	4.65100000	-5.04800000	92.37500000
C	3.95700000	-6.18900000	93.14500000
O	4.68000000	-7.24500000	93.31400000
O	2.80200000	-6.02400000	93.57300000
C	14.31200000	-3.49900000	94.66600000
C	8.36600000	-1.85900000	91.79900000
C	7.05300000	-1.39100000	91.24400000
Pt	10.70300000	-3.62400000	92.20300000
As	10.51034173	-1.82421789	93.63220982
O	12.33400000	-2.18800000	94.48500000
O	8.84100000	-1.14700000	92.78000000
O	9.72400000	-1.87300000	95.21900000
O	11.09700000	-0.15100000	93.49200000
N	12.43700000	-4.06300000	93.17500000
N	8.97800000	-2.92800000	91.34500000
C	12.94700000	-3.26000000	94.07800000
O	12.45300000	-5.57400000	96.21600000
H	15.10200000	1.89400000	100.07800000
H	13.77800000	1.51900000	101.18800000
H	15.07400000	-0.48600000	99.24600000
H	13.75100000	-0.83600000	100.35500000
H	13.88200000	3.37100000	98.25400000
O	11.54400000	0.20200000	100.24000000

H	13.73900000	-2.27400000	98.55600000
H	13.24400000	-3.72900000	97.70500000
H	11.80300000	-2.49600000	100.10900000
H	11.27800000	-3.95500000	99.26700000
H	11.68100000	-1.22700000	97.93600000
H	11.00500000	-2.63900000	97.15200000
H	9.72700000	-1.31200000	99.60100000
H	8.95500000	-2.49100000	98.49700000
H	8.23700000	-0.32200000	97.89700000
H	9.79600000	0.32000000	97.96700000
H	9.33500000	-0.76900000	96.72700000
H	13.83200000	-3.77300000	100.69800000
H	15.88700000	-2.35200000	100.49200000
H	16.12500000	-4.44800000	99.88400000
H	15.75200000	-3.67100000	98.48800000
H	12.76000000	-7.84900000	97.12000000
H	12.16400000	-7.71300000	98.74300000
H	11.11300000	-9.48100000	96.56600000
H	13.90300000	-10.32000000	99.73600000
H	10.49800000	-11.86800000	96.47500000
H	13.22800000	-12.68400000	99.69900000
H	11.54400000	-13.49300000	98.05500000
H	14.41500000	-8.13400000	99.67900000
H	15.20500000	-5.97800000	98.21800000
H	15.94800000	-7.79800000	93.32000000
H	17.24500000	-8.69500000	95.28500000
H	17.90000000	-9.23600000	93.71600000
H	16.95900000	-10.34800000	94.71300000
H	14.52400000	-9.32400000	91.94800000
H	16.24500000	-9.69800000	91.75600000
H	15.23300000	-10.74000000	92.74800000
H	14.95500000	-8.00600000	95.47700000
H	13.35400000	-10.04600000	95.75900000
H	14.93300000	-10.00900000	96.23500000
H	14.44800000	-10.79100000	94.75900000
H	10.02800000	-5.89100000	96.84300000
H	9.69000000	-4.28900000	96.26200000
H	6.68100000	-4.00900000	96.73000000
H	9.74900000	-5.86900000	99.13400000
H	5.40100000	-3.52300000	98.75000000
H	8.44900000	-5.35900000	101.17200000
H	6.26400000	-4.17100000	100.99600000
H	7.93600000	-6.69900000	95.61300000
H	9.96100000	-5.58400000	93.77500000
H	5.10500000	-5.46300000	91.46700000
H	3.92200000	-4.29700000	92.06300000
H	6.20600000	-6.19600000	94.09100000
O	4.72000000	-2.17700000	93.50000000
H	11.45800000	-6.91700000	92.84900000
H	11.07400000	-8.81100000	91.38400000
H	10.24300000	-9.71600000	92.66100000
H	7.80789470	-9.83489617	91.14742306
H	6.83362661	-5.83149945	92.01529715
H	5.90844038	-8.15959980	91.37733697
H	15.00200000	-2.76400000	94.23700000

H	14.28000000	-3.33800000	95.74400000
H	14.68000000	-4.50300000	94.44900000
H	7.10800000	-0.31800000	91.04000000
H	6.27000000	-1.56100000	91.99900000
H	6.78500000	-1.92400000	90.32800000
H	9.00200000	-2.57700000	95.09700000
H	11.99700000	-0.13600000	93.87200000
H	8.51700000	-3.39200000	90.56900000
H	13.14400000	-5.75600000	96.87900000
H	12.87000000	-5.78800000	95.36400000
O	10.70914550	-5.44226317	91.03299845
H	9.74982377	-6.19077655	91.25551069
H	10.73369424	-5.26932941	89.88688872
H	12.12800000	-9.50400000	94.08100000
H	12.96200000	-4.91600000	92.98900000
C	13.20800000	1.52700000	99.10500000
H	13.54500000	1.06800000	98.17000000
H	6.39100000	-3.78300000	92.53000000
N	16.24900000	-7.24700000	98.15000000
H	16.16800000	-6.35600000	97.67900000
C	17.22700000	-7.38800000	99.24600000
H	17.84600000	-8.25300000	99.04500000
C	18.13100000	-6.16900000	99.36900000
H	17.54000000	-5.28500000	99.61100000
H	18.82400000	-6.35500000	100.19000000
C	18.91400000	-5.89200000	98.11500000
H	19.43600000	-6.80700000	97.82800000
H	18.26000000	-5.63300000	97.30500000
C	19.95800000	-4.80900000	98.30400000
O	20.50500000	-4.66700000	99.42700000
O	20.24900000	-4.12600000	97.30200000
C	16.53500000	-7.58800000	100.56700000
O	16.93400000	-8.49500000	101.33900000
N	15.49700000	-6.77600000	100.81700000
H	15.24300000	-6.06000000	100.15400000
C	14.65800000	-6.98700000	101.98300000
H	15.29300000	-6.86200000	102.86100000
C	13.51400000	-5.96500000	102.09100000
H	13.94600000	-4.96300000	102.06200000
H	12.85600000	-6.07300000	101.22700000
C	12.70100000	-6.10800000	103.39000000
H	12.17300000	-7.06100000	103.38900000
H	13.38200000	-6.09300000	104.24200000
C	11.67800000	-4.98300000	103.57100000
H	12.06442906	-4.23942757	104.26761296
H	11.50666826	-4.49891512	102.60847038
N	10.41552381	-5.52720689	104.05417465
H	9.70081934	-5.69551205	103.35621127
C	10.13135770	-5.82001606	105.32505521
N	10.99279931	-5.61788606	106.31265474
H	11.87069135	-5.18938166	106.09063324
H	10.74012496	-5.81476899	107.26826490
N	8.96443070	-6.34129943	105.60954956
H	8.74014435	-6.62636444	106.54713080
H	8.28160956	-6.43453954	104.86564394

C	14.04700000	-8.37600000	102.05000000
O	14.08400000	-9.02100000	103.10600000
N	13.43000000	-8.82500000	100.96700000
H	13.77381392	-8.58353087	100.05194937
C	12.49729550	-9.71939673	101.32922692
H	12.04720124	-9.31927942	102.23929027
C	11.44639590	-9.66050136	100.26622627
H	11.88191677	-9.98300694	99.31981689
H	10.64848787	-10.35736879	100.53148571
C	10.83256285	-8.27965087	100.11134923
H	10.50793898	-7.94061949	101.07868116
H	11.54768513	-7.54789737	99.75485667
C	9.62296993	-8.27747086	99.19689081
O	8.62116791	-7.61902507	99.49061663
N	9.70096803	-9.02743215	98.09477631
H	8.91845155	-9.07762386	97.46192740
H	10.53527859	-9.56720858	97.91294862
C	12.82435555	-11.20654728	101.62681888
O	12.10798088	-12.10183767	101.16722292
N	13.86884581	-11.49134748	102.40094997
H	14.42720704	-10.74822380	102.79373990
C	14.40136457	-12.84551458	102.42735542
H	13.73509807	-13.51916721	101.89304083
C	15.72665530	-12.85808303	101.66471100
C	15.94383756	-14.14061669	100.93017354
N	16.02542437	-15.34760223	101.58137897
H	16.00218146	-15.46555954	102.58478360
C	16.18503141	-16.31979841	100.70065757
H	16.35122927	-17.36219614	100.92848040
N	16.18684689	-15.78553120	99.49481792
C	16.02228230	-14.42273157	99.60673612
H	15.96852912	-13.71125968	98.79708660
C	14.58462714	-13.50156562	103.76497714
O	14.84255180	-14.67543090	103.80574690
H	16.52791155	-12.71651713	102.35958329
H	15.72996477	-12.05120028	100.96198113
O	14.27849467	-12.75268266	105.02809124
O	10.92298677	1.28042662	98.35223432
H	16.29097603	-16.28903145	98.63712064

## TS

N	11.03900000	4.88600000	101.11200000
H	10.74400000	3.99000000	101.47800000
H	11.72200000	5.28300000	101.74400000
H	10.23900000	5.50100000	101.04600000
C	11.65500000	4.69800000	99.76700000
H	10.88900000	4.40100000	99.05000000
C	12.36000000	5.97600000	99.29800000
H	11.76300000	6.83500000	99.60900000
H	13.31600000	6.03900000	99.82300000
C	12.57500000	6.10500000	97.79700000
H	12.94400000	5.16300000	97.39800000
H	11.62900000	6.33000000	97.30300000



C	13.62800000	7.17000000	97.49900000
H	14.22700000	7.38800000	98.38400000
H	14.31500000	6.78800000	96.74100000
C	13.02600000	8.47600000	96.99300000
H	12.16800000	8.74800000	97.60900000
H	13.77800000	9.26300000	97.07300000
N	12.61800000	8.34900000	95.56600000
H	11.91500000	7.63000000	95.47000000
H	12.24000000	9.22800000	95.23900000
H	13.42000000	8.10800000	95.00000000
C	12.69300000	3.60400000	99.90200000
O	13.14700000	3.33900000	101.02600000
N	13.08100000	2.99900000	98.77300000
C	11.74800000	1.02800000	99.33800000
C	14.21600000	1.26400000	100.21800000
C	14.63800000	-0.20900000	100.21500000
C	15.63700000	-0.54700000	101.30800000
O	15.94500000	0.23300000	102.19300000
O	16.16800000	-1.77400000	101.27600000
N	15.51600000	-3.73900000	99.47700000
C	14.08900000	-4.06700000	99.67500000
C	13.70700000	-5.55600000	99.59300000
O	12.81400000	-6.00200000	100.31700000
C	13.24400000	-3.24200000	98.68700000
C	11.80100000	-3.00200000	99.13600000
C	11.07800000	-2.12400000	98.11400000
C	9.69100000	-1.69000000	98.57900000
N	9.22200000	-0.53700000	97.72800000
N	14.34500000	-6.30800000	98.65700000
C	14.28600000	-7.77500000	98.65000000
C	15.49100000	-8.27100000	97.79700000
O	15.55400000	-9.52600000	97.59200000
C	12.90600000	-8.24100000	98.13200000
C	12.57100000	-9.72000000	98.14800000
C	11.60600000	-10.18400000	97.23500000
C	13.14100000	-10.64200000	99.03600000
C	11.24000000	-11.53500000	97.19700000
C	12.76700000	-11.98600000	99.00500000
C	11.82200000	-12.44300000	98.08200000
N	14.33111027	-9.95839498	95.35040945
C	14.62991790	-8.63815138	94.68479632
C	13.45043721	-7.89655722	94.06558845
O	13.64798989	-6.70086633	93.79097922
C	15.80862837	-8.77053030	93.69027459
C	17.07622030	-9.24508299	94.41766407
C	15.46824551	-9.63864416	92.47137023
N	12.28005484	-8.51652611	93.84840954
C	11.11659372	-7.83132955	93.29385216
C	10.09407906	-7.53925915	94.40622334
O	9.76219817	-8.41500295	95.20649197
C	10.44410449	-8.71324701	92.20781171
C	9.11532557	-8.14638575	91.82571238
N	8.98200051	-6.92706038	91.18026985
C	7.87121438	-8.57249762	92.21863224
C	7.68462861	-6.64136585	91.19499273

N	6.97363699	-7.61852928	91.79689268
N	9.54700687	-6.30676722	94.36517883
C	8.51251047	-5.86928796	95.30371764
C	7.64350836	-4.86300601	94.54067508
O	8.00057858	-3.66699458	94.42853446
C	9.17653101	-5.27096935	96.55817711
C	8.28194284	-5.04426480	97.75714001
C	7.03966169	-4.39696342	97.66156558
C	8.74147803	-5.42841708	99.02701734
C	6.28915109	-4.13424454	98.81115621
C	7.99588844	-5.15763576	100.17572429
C	6.76461333	-4.50414442	100.07230439
N	6.54569953	-5.29875903	93.93576095
C	5.71151696	-4.36810895	93.17336416
C	5.02734346	-3.29251939	94.08775561
O	4.87509305	-3.56387975	95.30305860
C	4.65531682	-5.10925670	92.32009791
C	3.96686759	-6.26452178	93.07363443
O	4.69981700	-7.31406060	93.24000579
O	2.80628066	-6.11588960	93.49241892
C	13.81813774	-4.24404910	94.57700954
C	8.01363527	-2.69886299	91.38571577
C	6.74561909	-2.23652408	90.73043285
Pt	10.15239623	-4.56272424	92.21742052
As	10.19767579	-2.43815119	93.16386877
O	11.99214149	-2.80287059	94.07430865
O	8.59023664	-1.82559293	92.16022223
O	9.45288992	-2.04167004	94.69600117
O	10.96000069	-0.96509705	92.59777579
N	11.85310211	-4.94205792	93.27002038
N	8.49195547	-3.90894848	91.21157822
C	12.47238850	-4.00437665	93.94644939
O	12.45300000	-5.57400000	96.21600000
H	15.10200000	1.89400000	100.07800000
H	13.77800000	1.51900000	101.18800000
H	15.07400000	-0.48600000	99.24600000
H	13.75100000	-0.83600000	100.35500000
H	13.88200000	3.37100000	98.25400000
O	11.54400000	0.20200000	100.24000000
H	13.73900000	-2.27400000	98.55600000
H	13.24400000	-3.72900000	97.70500000
H	11.80300000	-2.49600000	100.10900000
H	11.27800000	-3.95500000	99.26700000
H	11.68100000	-1.22700000	97.93600000
H	11.00500000	-2.63900000	97.15200000
H	9.72700000	-1.31200000	99.60100000
H	8.95500000	-2.49100000	98.49700000
H	8.23700000	-0.32200000	97.89700000
H	9.79600000	0.32000000	97.96700000
H	9.33500000	-0.76900000	96.72700000
H	13.83200000	-3.77300000	100.69800000
H	15.88700000	-2.35200000	100.49200000
H	16.12500000	-4.44800000	99.88400000
H	15.75200000	-3.67100000	98.48800000
H	12.76000000	-7.84900000	97.12000000

H	12.16400000	-7.71300000	98.74300000
H	11.11300000	-9.48100000	96.56600000
H	13.90300000	-10.32000000	99.73600000
H	10.49800000	-11.86800000	96.47500000
H	13.22800000	-12.68400000	99.69900000
H	11.54400000	-13.49300000	98.05500000
H	14.41500000	-8.13400000	99.67900000
H	15.20500000	-5.97800000	98.21800000
H	15.97267430	-7.74460141	93.34484755
H	17.26146959	-8.64486063	95.31375157
H	17.93661429	-9.16476118	93.74615786
H	16.99880545	-10.29572850	94.72465749
H	14.57802022	-9.27374781	91.94646017
H	16.30468890	-9.62730320	91.76703094
H	15.29516639	-10.68900464	92.74049196
H	14.96240981	-7.98252002	95.49070924
H	13.38133925	-10.04222173	95.74004785
H	14.93300000	-10.00900000	96.23500000
H	14.49249140	-10.76638218	94.74365760
H	10.00031383	-5.93339954	96.82982046
H	9.65009930	-4.33007312	96.25981768
H	6.63407703	-4.08695617	96.70254650
H	9.70023850	-5.93477071	99.11826074
H	5.33049090	-3.63284602	98.71486933
H	8.37622771	-5.45704007	101.14860871
H	6.18004509	-4.29134078	100.96294214
H	7.92857330	-6.75313682	95.57357609
H	9.95789785	-5.59992370	93.76414651
H	5.12208305	-5.51122169	91.41270691
H	3.92100955	-4.36348367	92.00800591
H	6.20710153	-6.25547699	94.04019334
O	4.68263175	-2.24777597	93.47084068
H	11.47777302	-6.90830343	92.84029810
H	11.12788944	-8.79328556	91.35525910
H	10.29525331	-9.71856333	92.61656522
H	7.56156067	-9.44484046	92.77311358
H	7.24148724	-5.74161752	90.79491190
H	5.98929599	-7.54053533	92.16565104
H	14.57299297	-3.72038311	93.98007275
H	13.83301761	-3.81985207	95.58143199
H	14.06754276	-5.30555771	94.61606849
H	6.91276823	-1.26020817	90.26692622
H	5.96924909	-2.12550653	91.50305096
H	6.39611665	-2.94491436	89.97478727
H	8.65484109	-2.66763506	94.75140751
H	11.86607799	-0.95890082	92.96330352
H	7.96204304	-4.49458396	90.57422674
H	13.14400000	-5.75600000	96.87900000
H	12.87000000	-5.78800000	95.36400000
O	11.21872035	-5.63269219	90.45659076
H	10.42522859	-6.22645292	90.46997966
H	11.32434493	-5.35403701	89.53009806
H	12.16480998	-9.49870361	94.05565818
H	12.27593485	-5.86835564	93.30187954
C	13.20800000	1.52700000	99.10500000

H	13.54500000	1.06800000	98.17000000
H	6.37987893	-3.82684726	92.50220096
N	16.24900000	-7.24700000	98.15000000
H	16.16800000	-6.35600000	97.67900000
C	17.22700000	-7.38800000	99.24600000
H	17.84600000	-8.25300000	99.04500000
C	18.13100000	-6.16900000	99.36900000
H	17.54000000	-5.28500000	99.61100000
H	18.82400000	-6.35500000	100.19000000
C	18.91400000	-5.89200000	98.11500000
H	19.43600000	-6.80700000	97.82800000
H	18.26000000	-5.63300000	97.30500000
C	19.95800000	-4.80900000	98.30400000
O	20.50500000	-4.66700000	99.42700000
O	20.24900000	-4.12600000	97.30200000
C	16.53500000	-7.58800000	100.56700000
O	16.93400000	-8.49500000	101.33900000
N	15.49700000	-6.77600000	100.81700000
H	15.24300000	-6.06000000	100.15400000
C	14.65800000	-6.98700000	101.98300000
H	15.29300000	-6.86200000	102.86100000
C	13.51400000	-5.96500000	102.09100000
H	13.94600000	-4.96300000	102.06200000
H	12.85600000	-6.07300000	101.22700000
C	12.70100000	-6.10800000	103.39000000
H	12.17300000	-7.06100000	103.38900000
H	13.38200000	-6.09300000	104.24200000
C	11.67800000	-4.98300000	103.57100000
H	12.06442906	-4.23942757	104.26761296
H	11.50666826	-4.49891512	102.60847038
N	10.41552381	-5.52720689	104.05417465
H	9.70081934	-5.69551205	103.35621127
C	10.13135770	-5.82001606	105.32505521
N	10.99279931	-5.61788606	106.31265474
H	11.87069135	-5.18938166	106.09063324
H	10.74012496	-5.81476899	107.26826490
N	8.96443070	-6.34129943	105.60954956
H	8.74014435	-6.62636444	106.54713080
H	8.28160956	-6.43453954	104.86564394
C	14.04700000	-8.37600000	102.05000000
O	14.08400000	-9.02100000	103.10600000
N	13.43000000	-8.82500000	100.96700000
H	13.77381392	-8.58353087	100.05194937
C	12.49729550	-9.71939673	101.32922692
H	12.04720124	-9.31927942	102.23929027
C	11.44639590	-9.66050136	100.26622627
H	11.88191677	-9.98300694	99.31981689
H	10.64848787	-10.35736879	100.53148571
C	10.83256285	-8.27965087	100.11134923
H	10.50793898	-7.94061949	101.07868116
H	11.54768513	-7.54789737	99.75485667
C	9.62296993	-8.27747086	99.19689081
O	8.62116791	-7.61902507	99.49061663
N	9.70096803	-9.02743215	98.09477631
H	8.91845155	-9.07762386	97.46192740

H	10.53527859	-9.56720858	97.91294862
C	12.82435555	-11.20654728	101.62681888
O	12.10798088	-12.10183767	101.16722292
N	13.86884581	-11.49134748	102.40094997
H	14.42720704	-10.74822380	102.79373990
C	14.40136457	-12.84551458	102.42735542
H	13.73509807	-13.51916721	101.89304083
C	15.72665530	-12.85808303	101.66471100
C	15.94383756	-14.14061669	100.93017354
N	16.02542437	-15.34760223	101.58137897
H	16.00218146	-15.46555954	102.58478360
C	16.18503141	-16.31979841	100.70065757
H	16.35122927	-17.36219614	100.92848040
N	16.18684689	-15.78553120	99.49481792
C	16.02228230	-14.42273157	99.60673612
H	15.96852912	-13.71125968	98.79708660
C	14.58462714	-13.50156562	103.76497714
O	14.84255180	-14.67543090	103.80574690
H	16.52791155	-12.71651713	102.35958329
H	15.72996477	-12.05120028	100.96198113
O	14.27849467	-12.75268266	105.02809124
O	10.92298677	1.28042662	98.35223432
H	16.29097603	-16.28903145	98.63712064

## P

N	11.03900000	4.88600000	101.11200000
H	10.74400000	3.99000000	101.47800000
H	11.72200000	5.28300000	101.74400000
H	10.23900000	5.50100000	101.04600000
C	11.65500000	4.69800000	99.76700000
H	10.88900000	4.40100000	99.05000000
C	12.36000000	5.97600000	99.29800000
H	11.76300000	6.83500000	99.60900000
H	13.31600000	6.03900000	99.82300000
C	12.57500000	6.10500000	97.79700000
H	12.94400000	5.16300000	97.39800000
H	11.62900000	6.33000000	97.30300000
C	13.62800000	7.17000000	97.49900000
H	14.22700000	7.38800000	98.38400000
H	14.31500000	6.78800000	96.74100000
C	13.02600000	8.47600000	96.99300000
H	12.16800000	8.74800000	97.60900000
H	13.77800000	9.26300000	97.07300000
N	12.61800000	8.34900000	95.56600000
H	11.91500000	7.63000000	95.47000000
H	12.24000000	9.22800000	95.23900000
H	13.42000000	8.10800000	95.00000000
C	12.69300000	3.60400000	99.90200000
O	13.14700000	3.33900000	101.02600000
N	13.08100000	2.99900000	98.77300000
C	11.74800000	1.02800000	99.33800000
C	14.21600000	1.26400000	100.21800000
C	14.63800000	-0.20900000	100.21500000

C	15.63700000	-0.54700000	101.30800000
O	15.94500000	0.23300000	102.19300000
O	16.16800000	-1.77400000	101.27600000
N	15.51600000	-3.73900000	99.47700000
C	14.08900000	-4.06700000	99.67500000
C	13.70700000	-5.55600000	99.59300000
O	12.81400000	-6.00200000	100.31700000
C	13.24400000	-3.24200000	98.68700000
C	11.80100000	-3.00200000	99.13600000
C	11.07800000	-2.12400000	98.11400000
C	9.69100000	-1.69000000	98.57900000
N	9.22200000	-0.53700000	97.72800000
N	14.34500000	-6.30800000	98.65700000
C	14.28600000	-7.77500000	98.65000000
C	15.49100000	-8.27100000	97.79700000
O	15.55400000	-9.52600000	97.59200000
C	12.90600000	-8.24100000	98.13200000
C	12.57100000	-9.72000000	98.14800000
C	11.60600000	-10.18400000	97.23500000
C	13.14100000	-10.64200000	99.03600000
C	11.24000000	-11.53500000	97.19700000
C	12.76700000	-11.98600000	99.00500000
C	11.82200000	-12.44300000	98.08200000
N	14.30100000	-9.97600000	95.36000000
C	14.60800000	-8.66500000	94.68000000
C	13.43100000	-7.91600000	94.06500000
O	13.63900000	-6.72500000	93.77800000
C	15.77600000	-8.81900000	93.67600000
C	17.04500000	-9.30100000	94.39600000
C	15.41500000	-9.69400000	92.46800000
N	12.25200000	-8.52500000	93.86400000
C	11.09100000	-7.83200000	93.31400000
C	10.08200000	-7.51900000	94.43300000
O	9.74800000	-8.38400000	95.24400000
C	10.39900000	-8.71600000	92.24200000
C	9.07300000	-8.13800000	91.86700000
N	8.82526577	-6.79543970	91.81671821
C	7.87074553	-8.78963901	91.74986423
C	7.53119048	-6.62413265	91.66497268
N	6.90600763	-7.81918742	91.60471079
N	9.54800000	-6.28100000	94.38600000
C	8.52700000	-5.82400000	95.33000000
C	7.66200000	-4.81500000	94.56600000
O	8.03100000	-3.62400000	94.44000000
C	9.20900000	-5.22200000	96.57300000
C	8.32800000	-4.97500000	97.77800000
C	7.09200000	-4.31500000	97.68800000
C	8.79500000	-5.35300000	99.04700000
C	6.35500000	-4.03400000	98.84200000
C	8.06300000	-5.06400000	100.20000000
C	6.83800000	-4.39800000	100.10200000
N	6.55400000	-5.24400000	93.97500000
C	5.72300000	-4.31100000	93.21200000
C	5.05900000	-3.22000000	94.12300000
O	4.91500000	-3.47900000	95.34200000



C	4.65100000	-5.04800000	92.37500000
C	3.95700000	-6.18900000	93.14500000
O	4.68000000	-7.24500000	93.31400000
O	2.80200000	-6.02400000	93.57300000
C	14.40773318	-5.46679504	92.33394944
C	8.27113266	-2.84015483	91.02738898
C	6.96960133	-2.14235764	90.76259986
Pt	10.15906967	-5.08785371	91.37889088
As	10.95409009	-2.95238459	91.93818256
O	12.83964134	-3.67851800	92.24957419
O	9.24166402	-2.07272459	91.43241950
O	10.88866141	-2.28422741	93.55330425
O	11.73960094	-1.74269466	90.94263649
N	12.01274764	-5.77786607	91.86057744
N	8.41214915	-4.13778965	90.88710639
C	13.00038061	-4.96497271	92.15058344
O	12.45300000	-5.57400000	96.21600000
H	15.10200000	1.89400000	100.07800000
H	13.77800000	1.51900000	101.18800000
H	15.07400000	-0.48600000	99.24600000
H	13.75100000	-0.83600000	100.35500000
H	13.88200000	3.37100000	98.25400000
O	11.54400000	0.20200000	100.24000000
H	13.73900000	-2.27400000	98.55600000
H	13.24400000	-3.72900000	97.70500000
H	11.80300000	-2.49600000	100.10900000
H	11.27800000	-3.95500000	99.26700000
H	11.68100000	-1.22700000	97.93600000
H	11.00500000	-2.63900000	97.15200000
H	9.72700000	-1.31200000	99.60100000
H	8.95500000	-2.49100000	98.49700000
H	8.23700000	-0.32200000	97.89700000
H	9.79600000	0.32000000	97.96700000
H	9.33500000	-0.76900000	96.72700000
H	13.83200000	-3.77300000	100.69800000
H	15.88700000	-2.35200000	100.49200000
H	16.12500000	-4.44800000	99.88400000
H	15.75200000	-3.67100000	98.48800000
H	12.76000000	-7.84900000	97.12000000
H	12.16400000	-7.71300000	98.74300000
H	11.11300000	-9.48100000	96.56600000
H	13.90300000	-10.32000000	99.73600000
H	10.49800000	-11.86800000	96.47500000
H	13.22800000	-12.68400000	99.69900000
H	11.54400000	-13.49300000	98.05500000
H	14.41500000	-8.13400000	99.67900000
H	15.20500000	-5.97800000	98.21800000
H	15.94800000	-7.79800000	93.32000000
H	17.24500000	-8.69500000	95.28500000
H	17.90000000	-9.23600000	93.71600000
H	16.95900000	-10.34800000	94.71300000
H	14.52400000	-9.32400000	91.94800000
H	16.24500000	-9.69800000	91.75600000
H	15.23300000	-10.74000000	92.74800000
H	14.95500000	-8.00600000	95.47700000

H	13.35400000	-10.04600000	95.75900000
H	14.93300000	-10.00900000	96.23500000
H	14.44800000	-10.79100000	94.75900000
H	10.02800000	-5.89100000	96.84300000
H	9.69000000	-4.28900000	96.26200000
H	6.68100000	-4.00900000	96.73000000
H	9.74900000	-5.86900000	99.13400000
H	5.40100000	-3.52300000	98.75000000
H	8.44900000	-5.35900000	101.17200000
H	6.26400000	-4.17100000	100.99600000
H	7.93600000	-6.69900000	95.61300000
H	9.96100000	-5.58400000	93.77500000
H	5.10500000	-5.46300000	91.46700000
H	3.92200000	-4.29700000	92.06300000
H	6.20600000	-6.19600000	94.09100000
O	4.72000000	-2.17700000	93.50000000
H	11.45800000	-6.91700000	92.84900000
H	11.07400000	-8.81100000	91.38400000
H	10.24300000	-9.71600000	92.66100000
H	7.62611980	-9.84008203	91.78218596
H	7.01412335	-5.67752351	91.61438895
H	5.87623424	-7.99511492	91.74474392
H	14.99069524	-5.17815559	91.45235514
H	14.85747021	-4.98707543	93.20379696
H	14.44313539	-6.55160176	92.44561886
H	7.14959309	-1.26788915	90.13103211
H	6.55235334	-1.79796720	91.72140815
H	6.24572966	-2.80163915	90.27646541
H	10.05563702	-2.70071461	93.95875805
H	12.69900820	-1.92230686	90.98625049
H	7.58864171	-4.62740921	90.55229126
H	13.14400000	-5.75600000	96.87900000
H	12.87000000	-5.78800000	95.36400000
O	11.16170726	-6.41986045	89.75391380
H	10.30970726	-7.05486045	89.78591380
H	11.27270726	-6.13186045	88.83091380
H	12.12800000	-9.50400000	94.08100000
H	12.23072230	-6.77261325	91.88877485
C	13.20800000	1.52700000	99.10500000
H	13.54500000	1.06800000	98.17000000
H	6.39100000	-3.78300000	92.53000000
N	16.24900000	-7.24700000	98.15000000
H	16.16800000	-6.35600000	97.67900000
C	17.22700000	-7.38800000	99.24600000
H	17.84600000	-8.25300000	99.04500000
C	18.13100000	-6.16900000	99.36900000
H	17.54000000	-5.28500000	99.61100000
H	18.82400000	-6.35500000	100.19000000
C	18.91400000	-5.89200000	98.11500000
H	19.43600000	-6.80700000	97.82800000
H	18.26000000	-5.63300000	97.30500000
C	19.95800000	-4.80900000	98.30400000
O	20.50500000	-4.66700000	99.42700000
O	20.24900000	-4.12600000	97.30200000
C	16.53500000	-7.58800000	100.56700000

O	16.93400000	-8.49500000	101.33900000
N	15.49700000	-6.77600000	100.81700000
H	15.24300000	-6.06000000	100.15400000
C	14.65800000	-6.98700000	101.98300000
H	15.29300000	-6.86200000	102.86100000
C	13.51400000	-5.96500000	102.09100000
H	13.94600000	-4.96300000	102.06200000
H	12.85600000	-6.07300000	101.22700000
C	12.70100000	-6.10800000	103.39000000
H	12.17300000	-7.06100000	103.38900000
H	13.38200000	-6.09300000	104.24200000
C	11.67800000	-4.98300000	103.57100000
H	12.06442906	-4.23942757	104.26761296
H	11.50666826	-4.49891512	102.60847038
N	10.41552381	-5.52720689	104.05417465
H	9.70081934	-5.69551205	103.35621127
C	10.13135770	-5.82001606	105.32505521
N	10.99279931	-5.61788606	106.31265474
H	11.87069135	-5.18938166	106.09063324
H	10.74012496	-5.81476899	107.26826490
N	8.96443070	-6.34129943	105.60954956
H	8.74014435	-6.62636444	106.54713080
H	8.28160956	-6.43453954	104.86564394
C	14.04700000	-8.37600000	102.05000000
O	14.08400000	-9.02100000	103.10600000
N	13.43000000	-8.82500000	100.96700000
H	13.77381392	-8.58353087	100.05194937
C	12.49729550	-9.71939673	101.32922692
H	12.04720124	-9.31927942	102.23929027
C	11.44639590	-9.66050136	100.26622627
H	11.88191677	-9.98300694	99.31981689
H	10.64848787	-10.35736879	100.53148571
C	10.83256285	-8.27965087	100.11134923
H	10.50793898	-7.94061949	101.07868116
H	11.54768513	-7.54789737	99.75485667
C	9.62296993	-8.27747086	99.19689081
O	8.62116791	-7.61902507	99.49061663
N	9.70096803	-9.02743215	98.09477631
H	8.91845155	-9.07762386	97.46192740
H	10.53527859	-9.56720858	97.91294862
C	12.82435555	-11.20654728	101.62681888
O	12.10798088	-12.10183767	101.16722292
N	13.86884581	-11.49134748	102.40094997
H	14.42720704	-10.74822380	102.79373990
C	14.40136457	-12.84551458	102.42735542
H	13.73509807	-13.51916721	101.89304083
C	15.72665530	-12.85808303	101.66471100
C	15.94383756	-14.14061669	100.93017354
N	16.02542437	-15.34760223	101.58137897
H	16.00218146	-15.46555954	102.58478360
C	16.18503141	-16.31979841	100.70065757
H	16.35122927	-17.36219614	100.92848040
N	16.18684689	-15.78553120	99.49481792
C	16.02228230	-14.42273157	99.60673612
H	15.96852912	-13.71125968	98.79708660

C	14.58462714	-13.50156562	103.76497714
O	14.84255180	-14.67543090	103.80574690
H	16.52791155	-12.71651713	102.35958329
H	15.72996477	-12.05120028	100.96198113
O	14.27849467	-12.75268266	105.02809124
O	10.92298677	1.28042662	98.35223432
H	16.29097603	-16.28903145	98.63712064

### Coordinate of optimized structure using B3LYP functional

#### AP1-MET:

#### INT

Pt	-0.61004100	-0.72892500	-0.71975600
As	-2.16448900	0.56536500	0.46456100
Cl	0.88670800	-2.07247300	-2.03999500
O	-2.03973500	2.07124700	-0.85815600
O	-2.52198100	-0.97367200	1.78277400
O	-1.83772400	1.64970300	1.81000100
H	-0.96219900	2.08312000	1.63351100
O	-3.90862900	0.66960100	0.27733300
H	-4.31834200	-0.03037500	0.82091600
N	-0.45623900	0.86150100	-1.98375800
H	0.19573100	0.86475100	-2.76314300
N	-0.95106800	-2.19985900	0.65906500
H	-0.46353800	-3.09052600	0.62341000
C	-1.16975500	1.94787700	-1.82424400
C	-1.03386800	3.14345600	-2.72317500
H	-2.01873700	3.41868700	-3.11288600
H	-0.35082300	2.95609200	-3.55429900
H	-0.66044800	3.98668600	-2.13169700
C	-1.82601400	-2.06344300	1.62783700
C	-2.08787500	-3.15250800	2.63142000
H	-1.87035800	-2.77210200	3.63494200
H	-1.48277100	-4.04178200	2.44214700
H	-3.14879100	-3.42123000	2.60010600
N	2.95025500	-1.63949700	0.39953100
C	3.46431300	-0.22029000	0.33617800
C	4.97518400	-0.30366700	-0.05423800
O	5.53998500	0.79339400	-0.26343600
C	3.24462700	0.51158400	1.66185400
C	1.78515900	0.90821000	1.91999700
S	1.19099300	2.28478500	0.85328800
C	1.92434700	3.71495800	1.72778900
H	3.59921600	-0.11885200	2.48938200
H	3.88282600	1.39946000	1.65580100
H	1.10057300	0.07609800	1.72435600
H	1.64614200	1.20641300	2.96405600
H	1.58895400	4.61466800	1.20518500
H	1.58133700	3.74937200	2.76579500
H	3.01641400	3.67349100	1.69638200
H	2.92298100	0.28528900	-0.46754700
H	2.48447800	-1.87996600	1.27749100

H	2.30111800	-1.85717800	-0.37710400
H	3.81209800	-2.21798000	0.29216000
O	5.46410800	-1.47048200	-0.11421600

## TS

Pt	0.76299700	-0.28009000	0.36331000
As	2.81610000	0.61695800	-0.30824500
Cl	-0.53666400	-1.14089100	2.48867400
O	2.15978200	2.53017700	-0.05059200
O	3.60820900	-1.21146300	-0.66289000
O	3.26955900	1.03657300	-1.97584500
H	3.03415700	1.97046500	-2.13341600
O	4.26367100	1.06495100	0.62368500
H	4.87468100	0.30367300	0.63368000
N	0.18006000	1.65091700	0.68932100
H	-0.71613000	1.87249100	1.11198200
N	1.60587600	-2.10541000	-0.00370500
H	1.09565200	-2.96746400	0.16448100
C	0.95691400	2.67514700	0.43191500
C	0.51942400	4.09374500	0.67157400
H	1.21624000	4.57217100	1.36732600
H	-0.49399100	4.14976700	1.07493600
H	0.56287400	4.64374900	-0.27427500
C	2.83881000	-2.23993200	-0.42375500
C	3.46601300	-3.58578100	-0.65918600
H	3.80555000	-3.64673800	-1.69817000
H	2.77100200	-4.40335000	-0.45595900
H	4.34529900	-3.68804600	-0.01461600
N	-5.32612000	-0.15946500	1.45897400
C	-5.21922900	0.32783700	0.02706100
C	-6.63218200	0.11105200	-0.61745200
O	-6.72657500	0.39471400	-1.83011800
C	-4.09147400	-0.37269400	-0.73211400
C	-2.70797300	-0.02533100	-0.17498700
S	-1.34269500	-0.87523300	-1.05578500
C	-1.12538900	0.22864300	-2.49664100
H	-4.24942400	-1.45881300	-0.71311700
H	-4.17755900	-0.05628000	-1.77502500
H	-2.52843700	1.05382900	-0.22810000
H	-2.58597800	-0.33776000	0.86661400
H	-0.90241500	1.24180600	-2.15345500
H	-0.27703700	-0.15804600	-3.06569100
H	-2.02124600	0.22130400	-3.12287400
H	-5.03485000	1.40453700	0.07538500
H	-4.83422100	-1.04198900	1.62456800
H	-5.00423900	0.51644100	2.15516700
H	-6.35372200	-0.33556400	1.56874200
O	-7.52627600	-0.30849200	0.17578800

## PROD

Pt	0.51489100	-0.01769000	-0.09017100
As	2.83836800	-0.05691300	-0.54483800
Cl	1.12551000	-0.06272200	3.73055500
O	2.93241600	1.94790900	-0.70477000
O	2.92507500	-2.05949500	-0.41328700
O	3.63117700	-0.20640700	-2.11371900
H	3.79863700	0.68716100	-2.47011600
O	4.13273900	0.06588500	0.64533900
H	4.37005400	-0.83193500	0.94712300
N	0.67662200	2.01978800	-0.28783200
H	-0.10383800	2.65983900	-0.18385300
N	0.67222100	-2.05632000	0.01913400
H	-0.12701200	-2.65314200	0.21308000
C	1.82013700	2.61089000	-0.54696200
C	1.94065700	4.10274800	-0.68028000
H	2.66511300	4.46655700	0.05555800
H	0.98524800	4.60966200	-0.53014600
H	2.32811100	4.34225700	-1.67585700
C	1.80798400	-2.68746100	-0.15798100
C	1.91606900	-4.18427800	-0.08737900
H	2.30289400	-4.55767300	-1.04121400
H	0.95462200	-4.65656900	0.12405800
H	2.63481700	-4.45484900	0.69279500
N	-4.97662800	-1.20508200	-1.63799000
C	-4.97253700	-0.35182800	-0.38722300
C	-6.47487800	-0.10074600	-0.01605800
O	-6.67219600	0.63719800	0.97188000
C	-4.18584100	0.94467600	-0.57280200
C	-2.68350200	0.78330200	-0.82934600
S	-1.79180200	-0.08186600	0.53469600
C	-1.93915500	1.11136700	1.91203400
H	-4.60414600	1.51364600	-1.41400400
H	-4.35907100	1.55099200	0.32005900
H	-2.46693800	0.18339300	-1.71796300
H	-2.22113300	1.76263700	-0.97460100
H	-1.20138800	0.80303300	2.65787200
H	-1.72341100	2.12501000	1.56811200
H	-2.94969000	1.05561600	2.32269700
H	-4.53372100	-0.96223600	0.40584400
H	-4.66439600	-0.70813700	-2.47693700
H	-4.42519500	-2.06252400	-1.55582000
H	-5.99305400	-1.43710200	-1.75520400
O	-7.31647100	-0.69536800	-0.75172500

## AP1-HIS:

## INT

Pt	1.37690800	-0.05925900	-0.67362600
As	0.40512200	1.56604900	0.71112300
Cl	2.18869300	-1.90089500	-1.99863900
O	-0.65680200	2.34391700	-0.90712000
O	1.55486800	1.05185500	2.27094700
O	0.85042300	3.23234300	1.05916400



H	0.45021400	3.80240100	0.37530900
O	-1.07146800	1.34444100	1.60034300
H	-1.19315700	0.34608500	1.74970300
N	0.20153100	0.65379000	-2.18241700
H	0.16861600	0.19977200	-3.09069700
N	2.44579300	-0.61035900	0.97282600
H	3.07548300	-1.40756600	0.95965600
C	-0.59490000	1.68835400	-2.02491200
C	-1.50899100	2.16295500	-3.12080200
H	-2.54202700	1.98848600	-2.80021600
H	-1.33179700	1.63968600	-4.06323700
H	-1.37601300	3.23894300	-3.26835300
C	2.33771100	0.01902400	2.11719600
C	3.10385800	-0.39684700	3.34141500
H	3.70069300	0.44903500	3.69747200
H	3.75994900	-1.24766100	3.14564100
H	2.39392900	-0.66035200	4.13246900
N	-0.81835800	-3.09924000	-1.00299100
C	-1.50105500	-2.05092700	-0.16315300
C	-0.92506500	-2.20116200	1.27121700
O	-0.41840300	-3.31635600	1.54636600
C	-3.03705700	-2.22758000	-0.18900100
C	-3.72753100	-0.94343800	0.15437700
N	-3.73088700	0.11614300	-0.73846900
C	-4.37603200	-0.58298800	1.31218300
C	-4.37122700	1.09710000	-0.12634900
N	-4.78109800	0.72076100	1.11687200
H	-3.33979500	-2.52586200	-1.20131300
H	-3.34054100	-3.02988500	0.49184200
H	-4.56489200	-1.11518300	2.23233500
H	-5.29263700	1.29299200	1.77638500
H	-4.56439000	2.08193800	-0.52920600
H	-1.24867700	-1.08550900	-0.59770000
H	0.10920900	-2.77310500	-1.32794200
H	-1.35919000	-3.38253300	-1.82433600
H	-0.65016700	-3.91140500	-0.38742900
O	-1.06456900	-1.21833800	2.05385900

## TS

Pt	-0.42309900	-0.01057600	-0.45920400
As	-2.52768700	0.52670300	0.41191800
Cl	0.85420300	0.03909500	-2.75078800
O	-3.28124900	-1.30338800	-0.04010900
O	-1.93621900	2.38554800	0.96960300
O	-3.95713100	1.28325000	-0.32671500
H	-4.53647400	0.57817900	-0.67330300
O	-3.04952200	0.22035100	2.08574600
H	-2.86433500	1.01623600	2.61928000
N	-1.22298400	-1.83828700	-0.88574400
H	-0.67065800	-2.56086100	-1.33792300
N	0.09689900	1.91193700	0.03005600
H	0.99682200	2.30862400	-0.21997800
C	-2.46675100	-2.14528900	-0.61685800

C	-3.05239300	-3.48971900	-0.94938900
H	-3.39994300	-3.96693200	-0.02722900
H	-2.32902200	-4.14142800	-1.44407000
H	-3.92127700	-3.35056500	-1.60083600
C	-0.72300100	2.73346300	0.63715100
C	-0.33074400	4.13936400	0.99874100
H	-1.03366100	4.83952300	0.53610700
H	0.68570300	4.37786600	0.67823100
H	-0.40468400	4.26146000	2.08448600
N	3.78061800	0.82317900	-1.74948700
C	3.62390000	0.33821900	-0.33197600
C	4.93774300	0.62049100	0.44595700
O	5.97070900	0.77249000	-0.26434000
C	3.26670300	-1.16730900	-0.35813800
C	2.43841900	-1.57012000	0.81671300
N	1.14105200	-1.10862300	0.97498000
C	2.75072900	-2.38761500	1.87488500
C	0.68168600	-1.63415100	2.10310900
N	1.62760600	-2.41518500	2.67688500
H	2.69377100	-1.37798000	-1.26795600
H	4.18213000	-1.76666500	-0.39979500
H	3.65181500	-2.92704300	2.12385600
H	1.52602400	-2.92844300	3.54355600
H	-0.30522100	-1.48376100	2.51465400
H	2.79915700	0.90671100	0.09753600
H	3.95716700	1.83092300	-1.77563500
H	2.91646500	0.62306400	-2.29555600
H	4.59663700	0.37814300	-2.18248200
O	4.83259200	0.63261500	1.69669800

## PROD

Pt	-0.43394000	0.36373800	-0.57268500
As	-2.34567500	-0.80655600	0.10395500
Cl	1.30561200	0.52632300	4.28281700
O	-3.27628900	0.89186900	0.70464000
O	-1.58481000	-2.57291400	-0.43908700
O	-2.65756700	-1.50704500	1.69194300
H	-3.10454400	-0.83929700	2.24663000
O	-3.78100100	-1.13640800	-0.86520500
H	-3.66861400	-2.00054900	-1.30548100
N	-1.42355600	2.05737400	0.01704200
H	-1.01384300	2.98391000	-0.06042900
N	0.33048900	-1.46330300	-1.06156300
H	1.32409900	-1.56676200	-1.32189200
C	-2.62534700	2.01315500	0.54222900
C	-3.35190700	3.24806300	0.99688000
H	-4.29106100	3.33489900	0.44059500
H	-2.75836700	4.15282800	0.85005700
H	-3.60190700	3.14558500	2.05788200
C	-0.35119100	-2.56449600	-0.89174600
C	0.22606700	-3.91643400	-1.20200700
H	0.16329800	-4.54874600	-0.31046000
H	1.26574300	-3.84269900	-1.52742400

H	-0.37062100	-4.39052400	-1.98882500
N	3.23058000	-0.95113100	2.33034700
C	2.76277500	-0.59584100	0.94762300
C	3.71628200	-1.24164400	-0.09276800
O	4.91668000	-1.35861600	0.26254200
C	2.73425900	0.94718100	0.84533000
C	2.49702100	1.47436600	-0.53265400
N	1.28441300	1.36496900	-1.20457800
C	3.36774800	2.14633200	-1.35146400
C	1.42387500	1.94888700	-2.39282100
N	2.67566000	2.43238400	-2.50909800
H	1.96243500	1.31931800	1.52881900
H	3.69769800	1.33891800	1.19082100
H	4.39663300	2.43692400	-1.20696400
H	3.03811600	2.93156700	-3.31228100
H	0.66050800	2.03091100	-3.15047100
H	1.75818000	-1.00292000	0.85299300
H	3.13653000	-1.95612900	2.50168400
H	2.66769900	-0.44558300	3.06662800
H	4.22948300	-0.73090000	2.41908500
O	3.19387900	-1.55350200	-1.19940800

## APW-HIS

### INT

Pt	-1.78636500	-0.11481500	0.82893200
As	-0.81491600	-0.64657100	-1.22601000
O	0.08637300	-2.25710000	-0.49369600
O	-1.79835100	0.87077500	-2.14778200
O	0.65960000	0.02887100	-1.84139400
H	1.49503500	-0.52044300	-1.66112300
O	-1.48945600	-1.72318700	-2.45223900
H	-2.02925600	-1.18856300	-3.06573000
N	-0.77471700	-1.73410500	1.55909100
H	-0.80972400	-2.01370200	2.53521900
N	-2.72147100	1.41884800	-0.13140400
H	-3.34102700	2.06126100	0.35413700
C	-0.04810500	-2.49804000	0.78288100
C	0.68312800	-3.70837000	1.29152600
H	0.33481300	-4.59414500	0.75015400
H	0.52985400	-3.85607900	2.36234200
H	1.75323300	-3.59523800	1.08977500
C	-2.57159700	1.62499500	-1.41658400
C	-3.27461400	2.73951100	-2.14107700
H	-2.52948200	3.40658300	-2.58715800
H	-3.92176100	3.31457700	-1.47574000
H	-3.87373600	2.31929100	-2.95543800
O	-2.69179400	0.49619600	2.74368900
H	-2.01084000	0.72562000	3.40202800
H	-3.22196500	-0.21806500	3.14214600
N	5.15809300	-0.26444400	1.02975000
C	3.95805500	-0.01732000	0.13789800
C	3.87520600	-1.24774400	-0.81882200

O	2.92841500	-1.25584700	-1.65126400
C	4.06380800	1.31925600	-0.60085000
C	4.04161200	2.57140500	0.29163300
S	2.54207800	2.79470700	1.32858900
C	1.29314500	3.21296400	0.05700300
H	4.98642700	1.33637500	-1.19909100
H	3.23641900	1.35489900	-1.31468500
H	4.87577900	2.57598800	1.00196900
H	4.16296400	3.45979300	-0.33722300
H	0.37703600	3.47334700	0.59399000
H	1.61907500	4.07727600	-0.52979000
H	1.08794800	2.36478900	-0.60021700
H	3.07791100	-0.01511200	0.78621600
H	5.92722700	0.39203800	0.86945300
H	4.93736700	-0.25086500	2.02839600
H	5.48201000	-1.21838100	0.74777900
O	4.76183100	-2.12672400	-0.66942800

## TS

Pt	-1.32237600	-0.26656400	-0.76697600
As	-1.46184100	0.99187100	1.18194500
O	-0.77063100	2.64516500	0.27319300
O	-2.13158900	-0.55732900	2.25609900
O	-0.14004500	1.11248800	2.34865200
H	0.41187100	1.88414800	2.11848600
O	-2.76771800	1.99207600	1.82096300
H	-3.32050400	1.44827900	2.41467300
N	-0.68538600	1.44809700	-1.67237900
H	-0.47221300	1.48272700	-2.66557100
N	-2.01996500	-1.82323900	0.35449600
H	-2.18771400	-2.74506000	-0.03833100
C	-0.51989400	2.56359800	-1.00756900
C	-0.02827000	3.82858800	-1.65049100
H	0.15489700	3.70088300	-2.71915900
H	0.89800900	4.14133000	-1.15755000
H	-0.77000900	4.61929200	-1.49842000
C	-2.29806200	-1.69155000	1.62699400
C	-2.82321900	-2.82182000	2.46440500
H	-2.12461800	-3.01143100	3.28579800
H	-2.95550400	-3.73562200	1.88162200
H	-3.78185700	-2.52671500	2.90303600
N	4.43541700	1.55664400	0.07365100
C	4.37940000	0.13129200	0.58608400
C	5.86769100	-0.31697800	0.79654500
O	6.03663000	-1.50831500	1.13082000
C	3.59096800	-0.78395500	-0.35005400
C	2.10365300	-0.42772000	-0.40586300
S	1.12092800	-1.56506500	-1.46408100
C	1.02262000	-3.03143500	-0.37880400
H	4.02966600	-0.75821100	-1.35590500
H	3.72728500	-1.79916200	0.03118300
H	1.65995700	-0.43049800	0.59522400
H	1.93223500	0.56389400	-0.83678200

H	0.37226300	-3.75557800	-0.87613500
H	2.00984100	-3.47728100	-0.23633400
H	0.58580800	-2.75296900	0.58316400
H	3.90303600	0.17227600	1.56971300
H	4.22630300	1.63284300	-0.92580100
H	3.82739500	2.20751300	0.57536900
H	5.44240200	1.81336100	0.21129400
O	6.73121600	0.59462500	0.62880400
O	-1.91282400	-1.17475700	-2.84323800
H	-1.48692100	-2.03206800	-3.02118400
H	-2.87049200	-1.34292600	-2.89333500

## PROD

Pt	-0.48467100	0.06919800	0.05362200
As	-2.79386500	0.25638200	-0.42440800
O	-2.94934900	-1.71553000	-0.84126400
O	-2.83489100	2.20166600	-0.02674400
O	-3.66206900	0.61641000	-1.90688100
H	-3.86841200	-0.21996100	-2.36755100
O	-4.01805500	-0.03134700	0.82830000
H	-4.25776500	0.81828700	1.24714500
N	-0.71976300	-1.92778100	-0.34719800
H	0.03350900	-2.60538900	-0.29185600
N	-0.58350900	2.08409600	0.40450600
H	0.23094200	2.62977000	0.67270200
C	-1.87231100	-2.43842900	-0.72052900
C	-2.03848800	-3.89850100	-1.02964200
H	-2.79343300	-4.31859500	-0.35717000
H	-1.10434300	-4.45288000	-0.91948200
H	-2.40781900	-4.00534900	-2.05463300
C	-1.70106800	2.76243300	0.31301700
C	-1.77737600	4.23717900	0.58289900
H	-2.13526400	4.74503100	-0.31870200
H	-0.81039700	4.65140800	0.87452000
H	-2.50700800	4.41730200	1.37889500
N	4.93407900	1.22967000	-1.47694400
C	4.96955900	0.32472800	-0.26342800
C	6.48411800	0.08847600	0.06837400
O	6.71717500	-0.70380900	1.00492900
C	4.19770000	-0.97539800	-0.48142100
C	2.68731500	-0.82732300	-0.69969100
S	1.81925800	-0.01290800	0.71061400
C	1.90317500	-1.29792300	2.00741600
H	4.60385300	-1.50659800	-1.35254000
H	4.40201000	-1.61296100	0.38245400
H	2.44330300	-0.20167000	-1.56300400
H	2.23430500	-1.80763600	-0.86660800
H	1.29797400	-0.93729200	2.84086600
H	1.50661400	-2.24633700	1.64086600
H	2.94093000	-1.41131400	2.32726900
H	4.53935600	0.89640000	0.56263000
H	4.64688800	0.75365800	-2.33693700
H	4.34243100	2.05594000	-1.36046000

H	5.93886200	1.51172800	-1.58256700
O	7.29711700	0.75227800	-0.63973300
O	-2.55949800	-1.80111300	2.67536400
H	-3.15234900	-1.23010600	2.15025500
H	-1.68525000	-1.67670500	2.27076000

**APW-HIS:****INT**

Pt	1.17431400	-0.76901900	-0.71653100
As	0.87055400	0.20524700	1.37810300
O	0.15802600	-1.49040500	2.15148000
O	1.82838200	1.90661500	0.78657200
O	-0.51311000	1.12194600	1.86560600
H	-1.35136700	0.61229400	1.60086000
O	1.98347100	0.20150200	2.73810600
H	2.56464500	0.98197400	2.66020700
N	0.18009100	-2.38569600	0.05044300
H	-0.09870000	-3.18526000	-0.51135500
N	2.14058800	0.94002200	-1.25853300
H	2.53132500	1.06395500	-2.18848400
C	-0.14320000	-2.44723500	1.32040600
C	-0.90508300	-3.59986200	1.90641200
H	-0.41526300	-3.93225700	2.82609100
H	-0.98613700	-4.43618300	1.20877100
H	-1.90982900	-3.24576600	2.16228500
C	2.28867600	1.94570900	-0.43075900
C	3.00846200	3.20736600	-0.81761700
H	2.34703800	4.06470100	-0.65877400
H	3.33611100	3.19053500	-1.85911700
H	3.87838600	3.33552500	-0.16497300
O	1.57543200	-1.63377500	-2.70418700
H	0.76069300	-1.78583500	-3.21688400
H	2.01751500	-2.50018400	-2.64042700
N	-3.03090000	-1.38262600	-2.13106400
C	-2.60997000	-0.36324400	-1.09539500
C	-3.02738800	-0.95788700	0.27966300
O	-3.81855100	-1.93477400	0.24351300
C	-3.29310400	0.99439000	-1.37562500
C	-2.61868400	2.12552100	-0.66517800
N	-1.37189600	2.57447700	-1.06774600
C	-3.05402600	2.85493300	0.41640600
C	-1.06316100	3.55610200	-0.23916100
N	-2.05204100	3.76672400	0.67130800
H	-3.24400200	1.18703600	-2.45528400
H	-4.35215100	0.94122100	-1.09992400
H	-3.95182100	2.79921900	1.01346700
H	-2.04300400	4.45316000	1.41459800
H	-0.15366400	4.13905700	-0.25155500
H	-1.52691000	-0.25168800	-1.15991200
H	-2.25097700	-1.96747600	-2.44187500
H	-3.46449600	-0.97266700	-2.96296500
H	-3.71312700	-1.99411300	-1.63476300



O	-2.57053700	-0.39260900	1.30969400
---	-------------	-------------	------------

## TS

Pt	0.64350700	-0.25287700	-0.79975600
As	2.06323100	0.80075600	0.69669000
O	1.35398400	2.62355800	0.23820300
O	2.90102700	-0.90784600	1.28526300
O	3.69908600	1.40723400	0.43815300
H	3.64566300	2.34386600	0.16717900
O	1.65516900	1.17636400	2.37336900
H	1.95772600	0.44464000	2.94454900
N	-0.04165800	1.61346400	-1.26215100
H	-0.70999000	1.76750200	-2.01232300
N	1.54589700	-1.98887000	-0.20560000
H	1.28258300	-2.89516400	-0.58126000
C	0.44738600	2.69044800	-0.69995600
C	0.00048600	4.07547000	-1.07052500
H	-0.42537600	4.56044800	-0.18563400
H	-0.74256600	4.06574200	-1.87042000
H	0.86990300	4.66127800	-1.38538700
C	2.47387100	-2.00683500	0.71554000
C	3.11910300	-3.27374600	1.19806500
H	4.20156400	-3.20405600	1.05064300
H	2.73566600	-4.15206100	0.67491200
H	2.93464300	-3.38063500	2.27207400
N	-3.08646900	2.38827500	1.14317700
C	-3.09398900	1.03320700	0.47253300
C	-4.59755300	0.64553600	0.30923600
O	-5.37527300	1.18537800	1.15133500
C	-2.32350800	0.04291900	1.36789000
C	-2.23968900	-1.32955200	0.78310900
N	-1.53201200	-1.60057100	-0.37792500
C	-2.80215200	-2.49145800	1.25312100
C	-1.66661900	-2.90290900	-0.59802400
N	-2.42904500	-3.47730000	0.36395100
H	-1.30837000	0.42942500	1.52452800
H	-2.81153800	-0.01107100	2.34775400
H	-3.42280700	-2.69344400	2.11286700
H	-2.68005800	-4.45674500	0.41621800
H	-1.23509500	-3.46384100	-1.41499400
H	-2.60580700	1.12848100	-0.49586300
H	-3.95704000	2.40236700	1.71179100
H	-3.15219100	3.15224300	0.46488300
H	-2.26661800	2.56116400	1.73232800
O	-4.86688600	-0.16478200	-0.60347100
O	-0.22519700	-0.88931700	-2.88888200
H	0.27636400	-1.58219800	-3.35296500
H	-1.10312800	-1.26661700	-2.69851300

PROD

Pt	0.32787500	-0.56497700	-0.48057300
As	2.20089900	0.79197100	-0.13686200
O	3.23293500	-0.68614300	0.75960400
O	1.32962000	2.38308700	-0.99331300
O	2.29583500	1.73322300	1.37162300
H	2.96361000	1.32227300	1.95501300
O	3.66992200	1.13359700	-1.03410300
H	3.52231900	1.92178300	-1.59169400
N	1.41411700	-2.04484000	0.42470000
H	1.04311500	-2.97814700	0.57776000
N	-0.53822200	1.08416300	-1.31621000
H	-1.54596300	1.10037700	-1.52298900
C	2.62724200	-1.83964800	0.87877900
C	3.42186800	-2.91080500	1.57040900
H	4.36479200	-3.05939900	1.03440000
H	2.87900300	-3.85705200	1.62069000
H	3.66463400	-2.57624100	2.58444600
C	0.08749300	2.23018800	-1.38855800
C	-0.57324100	3.46776300	-1.92521400
H	-0.52392700	4.25763600	-1.16866200
H	-1.61502200	3.28218000	-2.19412900
H	-0.02388400	3.81561500	-2.80656700
N	-2.62475300	1.62107300	2.47539800
C	-2.53505100	0.85676500	1.18117200
C	-3.69248600	1.26566400	0.23755800
O	-3.44618800	1.14340700	-0.99466000
C	-2.50952600	-0.65206000	1.51312700
C	-2.51649000	-1.54691700	0.31398700
N	-1.44686800	-1.65949700	-0.56786800
C	-3.50478600	-2.39094200	-0.12228100
C	-1.78359300	-2.53799300	-1.50867700
N	-3.02545800	-2.99840700	-1.26381700
H	-1.61665800	-0.85406000	2.11715400
H	-3.38779100	-0.89933800	2.12039300
H	-4.48529900	-2.60184800	0.27508600
H	-3.51495000	-3.68619600	-1.82336000
H	-1.16577400	-2.84337900	-2.33846300
H	-1.58653000	1.15928900	0.74666500
H	-2.77850100	2.61679800	2.29323200
H	-1.72209700	1.51964300	2.98713400
H	-3.41264300	1.30863100	3.05010900
O	-4.74918700	1.66061000	0.79238700
O	0.04594100	1.18293200	3.19273400
H	0.41726600	1.46663500	4.04388200
H	0.70906900	1.44269100	2.51960200

### Coordinate of optimized structure using M062X functional

#### AP1-MET:

#### INT

Pt	0.85326000	0.20084800	-0.66502400
As	2.60222600	-0.36023900	0.71043300
Cl	-1.03082300	0.80825800	-2.10319400
O	2.05810500	-2.23531000	0.84506000

O	3.32597400	1.46377000	0.72182200
O	2.61328700	-0.26994300	2.45214200
H	2.33916800	-1.12972300	2.82062700
O	4.17096600	-0.94138300	0.22064600
H	4.77750100	-0.18445400	0.12402900
N	0.32150600	-1.75855400	-0.54829400
H	-0.51878000	-2.12841200	-0.99336000
N	1.59780300	2.09463400	-0.61746000
H	1.16204800	2.87090200	-1.10678800
C	0.98504600	-2.59556800	0.19072900
C	0.57617700	-4.02637700	0.35841100
H	1.39112400	-4.67348800	0.02357100
H	-0.32909100	-4.25183400	-0.20728900
H	0.40428300	-4.22302200	1.41998600
C	2.67822400	2.37648900	0.04979500
C	3.26619100	3.75260700	0.10809600
H	3.30148300	4.07714000	1.15131900
H	2.68483300	4.46451000	-0.47797100
H	4.29167000	3.71658600	-0.26832700
N	-2.67750400	1.86914300	0.49437400
C	-3.92076600	1.02897800	0.45322600
C	-5.11004800	1.93474500	0.88774300
O	-6.19608700	1.35150600	1.06323700
C	-3.76610600	-0.23080700	1.29025300
C	-2.67698900	-1.16950200	0.77290000
S	-3.06596500	-1.94537300	-0.83303200
C	-4.09257200	-3.32469500	-0.24771400
H	-3.54023700	0.04694100	2.32919600
H	-4.73785700	-0.73220200	1.30460500
H	-1.72937500	-0.63627800	0.63768700
H	-2.48370500	-1.97103100	1.49332400
H	-4.40362400	-3.89102300	-1.12758200
H	-3.51630500	-3.97669900	0.41315100
H	-4.98310200	-2.96142600	0.27008600
H	-4.07238300	0.76022200	-0.59863100
H	-2.14722400	1.74471500	1.36253300
H	-2.04538200	1.68277200	-0.30248200
H	-3.01442800	2.85010400	0.48069700
O	-4.84805600	3.16151700	0.98222700

## TS

Pt	-0.79917500	0.23675900	0.30600600
As	-2.83548300	-0.60545700	-0.30933600
Cl	0.61420800	0.99388300	2.49899300
O	-2.23079400	-2.47199200	-0.20017400
O	-3.64997000	1.16984300	-0.49766700
O	-3.38027700	-0.91881000	-1.94175300
H	-3.17618100	-1.84241000	-2.17602700
O	-4.17090800	-1.07613000	0.71747900
H	-4.76951800	-0.31584300	0.83251700
N	-0.20413200	-1.70332600	0.48289300
H	0.70755700	-1.97016700	0.84198000
N	-1.65117100	2.07138000	0.09795400

H	-1.14484600	2.93217800	0.28278500
C	-1.01384200	-2.68551500	0.21580000
C	-0.62688000	-4.12496600	0.36422600
H	-1.31530900	-4.60484100	1.06450800
H	0.39683700	-4.23427600	0.72271100
H	-0.73143900	-4.62099600	-0.60420100
C	-2.89645700	2.20527200	-0.24913800
C	-3.56519500	3.53838700	-0.38680900
H	-3.97289400	3.62842900	-1.39670800
H	-2.87236400	4.35806600	-0.19584200
H	-4.39867200	3.59073500	0.31869800
N	5.34786800	-0.16029700	1.48636100
C	5.26185700	-0.39607800	0.00378100
C	6.68322000	-0.10479100	-0.57593600
O	6.76702300	-0.09506800	-1.81581300
C	4.16366000	0.44081200	-0.63706500
C	2.76844100	0.03210900	-0.17038800
S	1.45690200	1.07922100	-0.86955300
C	1.33204600	0.37709800	-2.53814600
H	4.33800700	1.50352200	-0.42582900
H	4.26423300	0.30852700	-1.71775500
H	2.56069000	-1.00975500	-0.44089900
H	2.64319400	0.13773400	0.91258700
H	0.50496800	0.88567900	-3.03686900
H	2.25137100	0.55111500	-3.10072200
H	1.11822100	-0.69269700	-2.47936800
H	5.06202000	-1.46310900	-0.13231400
H	4.96543100	0.74924100	1.76469100
H	4.89475500	-0.88047000	2.05350100
H	6.37250900	-0.13969200	1.67955600
O	7.58579600	0.05842000	0.28470300

## PROD

Pt	0.60832200	0.08611200	0.02264500
As	2.83538400	-0.38772700	-0.43918700
Cl	-0.12535100	0.82450500	3.65557000
O	3.26740600	1.47058000	-0.87200100
O	2.62252000	-2.30068200	-0.05104500
O	3.48803800	-0.87120600	-1.98188200
H	3.80664300	-0.08384700	-2.46022500
O	4.17014100	-0.33956100	0.68092000
H	4.30479300	-1.22510900	1.06550500
N	1.10519100	2.02406300	-0.40531500
H	0.45743700	2.80382300	-0.35961500
N	0.42385900	-1.91040300	0.39671300
H	-0.44353300	-2.34487300	0.69776000
C	2.31237500	2.35192600	-0.76967800
C	2.70628400	3.75912800	-1.09756900
H	3.50028300	4.07115600	-0.41403700
H	1.86344100	4.44572900	-1.01819000
H	3.10754000	3.78554000	-2.11386200
C	1.44224600	-2.71757900	0.30926700
C	1.34497100	-4.18027600	0.61628800

H	1.65341900	-4.74771600	-0.26529000
H	0.33325100	-4.46791800	0.90266000
H	2.03783700	-4.41608900	1.42819400
N	-4.69282600	-1.79496300	-0.93618700
C	-4.80178600	-0.40082500	-0.38612900
C	-6.32499500	-0.05642700	-0.34242800
O	-6.60051500	1.11983300	-0.05173600
C	-3.99224100	0.60042700	-1.19647900
C	-2.48076500	0.37441300	-1.19990300
S	-1.74584400	0.42413700	0.47401400
C	-1.99869400	2.16660700	0.90423400
H	-4.33705400	0.59485700	-2.23862000
H	-4.23135100	1.59286300	-0.80428600
H	-2.20433200	-0.60489600	-1.60346700
H	-1.98650100	1.13295800	-1.81275800
H	-1.31698800	2.37754500	1.73043300
H	-1.79697900	2.80555700	0.04176000
H	-3.02908600	2.30574100	1.23626600
H	-4.44664600	-0.44492300	0.64774000
H	-4.43084700	-1.81401300	-1.92647200
H	-4.04492700	-2.39996000	-0.42589600
H	-5.66507900	-2.16750700	-0.86588200
O	-7.09528700	-1.02130000	-0.58123400

**AP1-HIS:****INT**

Pt	0.85340500	0.60695100	0.74172600
As	1.66456100	-0.99860500	-0.69388400
Cl	0.18147500	2.29270500	2.38075000
O	0.98115300	-2.48107900	0.36058500
O	2.45030000	0.34890600	-1.91164700
O	3.29316000	-1.60508600	-0.47554600
H	3.26505400	-2.41012700	0.07206100
O	1.07096400	-1.82455000	-2.10701700
H	1.32613000	-1.31729900	-2.89895100
N	0.11887700	-0.95772900	1.82001500
H	-0.43705900	-0.83837900	2.66186000
N	1.69112400	1.97411300	-0.51269300
H	1.65268700	2.97441000	-0.34181000
C	0.27216100	-2.18390600	1.41723000
C	-0.37070400	-3.35112800	2.09870100
H	-1.13630600	-3.75062800	1.42416500
H	-0.83542900	-3.06446800	3.04252000
H	0.37243500	-4.13312600	2.26919700
C	2.32480600	1.60458000	-1.58807000
C	2.95494600	2.58017300	-2.53500700
H	4.01991700	2.35048200	-2.62304300
H	2.82818300	3.60979000	-2.19978900
H	2.50191100	2.45741000	-3.52225900
N	-2.50634300	2.56175700	0.54735500
C	-2.70470000	1.32524300	-0.27350600
C	-3.79495000	1.65604000	-1.33291100

O	-3.95676700	2.88477000	-1.55479300
C	-2.98473500	0.13870100	0.64561900
C	-2.74621600	-1.17737200	-0.02785700
N	-1.64468000	-1.38742700	-0.82607700
C	-3.49515600	-2.32454000	0.04895800
C	-1.73014800	-2.63706000	-1.23352400
N	-2.83094600	-3.24531300	-0.72787000
H	-2.31313900	0.22142500	1.51258800
H	-4.01289100	0.18357000	1.02250200
H	-4.41644600	-2.55696500	0.56131000
H	-3.11491800	-4.20115200	-0.89880700
H	-1.01476500	-3.13932300	-1.86923200
H	-1.75838400	1.15714700	-0.79890600
H	-1.58918500	2.58296300	1.02554100
H	-3.23119000	2.65187300	1.26636800
H	-2.62237600	3.36444200	-0.08872100
O	-4.36549600	0.68847800	-1.87017100

## TS

Pt	-0.46368400	-0.04919300	-0.34395200
As	-2.54728100	0.57523800	0.35354000
Cl	0.97758500	-0.09384200	-2.62527300
O	-3.34856000	-1.16072500	-0.09304200
O	-1.97235800	2.37306500	0.88586600
O	-3.80824500	1.35632500	-0.57054000
H	-4.38658600	0.67704000	-0.96252900
O	-3.21602700	0.35413400	1.95500500
H	-3.04243000	1.15099500	2.48831400
N	-1.29556100	-1.84836400	-0.77599700
H	-0.75456800	-2.61413500	-1.16623300
N	0.11158900	1.85095100	0.14489300
H	1.04336900	2.21878500	-0.01878900
C	-2.55981600	-2.07643300	-0.58438900
C	-3.21051000	-3.38695500	-0.90494100
H	-3.66514600	-3.78832500	0.00432300
H	-2.49445600	-4.10534300	-1.30447700
H	-4.00679100	-3.21852100	-1.63453700
C	-0.73034800	2.69825400	0.65736600
C	-0.35623400	4.10185200	1.02321200
H	-0.98703200	4.79408200	0.45960900
H	0.69348900	4.30814600	0.81375800
H	-0.55462400	4.25543500	2.08702500
N	3.93348200	0.69091600	-1.88563200
C	3.67437900	0.27364300	-0.47377600
C	4.77970200	0.90943600	0.41619200
O	5.36358500	1.89854800	-0.09589000
C	3.52717200	-1.24608200	-0.39636800
C	2.64306400	-1.67660600	0.72941000
N	1.33013500	-1.26211200	0.83981600
C	2.93695500	-2.51993300	1.76750600
C	0.84903900	-1.84397600	1.92488200
N	1.78707300	-2.61331800	2.51264000
H	3.08501100	-1.59783500	-1.33655600

H	4.51055500	-1.71420900	-0.28720200
H	3.84461900	-3.03823900	2.03568400
H	1.66268700	-3.15839100	3.35624200
H	-0.15668400	-1.73683900	2.30594900
H	2.71797100	0.73378900	-0.20701700
H	4.29492900	1.65451400	-1.86130500
H	3.06027500	0.61300700	-2.43914800
H	4.65590600	0.11253600	-2.32504100
O	4.94420000	0.40081900	1.54089300

## PROD

Pt	-0.53795300	0.03187200	-0.60285800
As	-2.42545100	-0.44437900	0.63292700
Cl	2.74939800	2.90272500	3.27213700
O	-3.20863900	1.33956300	0.32278900
O	-1.83761200	-2.26096100	1.05330700
O	-2.65700400	-0.20809000	2.34421400
H	-3.03415200	0.67694000	2.50080700
O	-3.91689800	-1.09640400	0.00875000
H	-3.89734000	-2.06823900	0.07999400
N	-1.39006900	1.85669000	-0.95039300
H	-0.95752200	2.57998500	-1.51668800
N	0.08641100	-1.83047600	-0.09844000
H	1.03795300	-2.15473200	-0.32714700
C	-2.54066200	2.17115800	-0.42732200
C	-3.19078300	3.50449800	-0.63918100
H	-4.17349400	3.35131200	-1.09244300
H	-2.58905300	4.15166700	-1.27755800
H	-3.33980600	3.98283400	0.33234700
C	-0.64880200	-2.61290900	0.62979300
C	-0.19850400	-3.97996400	1.04569800
H	-0.21892400	-4.04657900	2.13656300
H	0.80680400	-4.19254600	0.68010400
H	-0.89830400	-4.72131900	0.65060700
N	3.49993500	0.07542100	2.20260700
C	2.80140100	-0.26144300	0.92768900
C	3.57033600	-1.43010200	0.25869300
O	4.80331900	-1.44332700	0.46980700
C	2.81917000	1.00027500	0.04447100
C	2.44750900	0.74495300	-1.37874700
N	1.17402100	0.39706100	-1.78406200
C	3.24601300	0.79623000	-2.48835800
C	1.20295800	0.23827500	-3.09799300
N	2.44288300	0.47343200	-3.55433000
H	2.15004600	1.74929800	0.48412000
H	3.83227300	1.41920100	0.05608200
H	4.29298400	1.03058000	-2.60217600
H	2.72692000	0.43161200	-4.52564700
H	0.36475300	-0.03385700	-3.72139300
H	1.77866600	-0.55171700	1.17367300
H	3.28755300	-0.60332200	2.93905600
H	3.24875800	1.03268600	2.56115200
H	4.51192900	0.01404400	2.01923300



O	2.89012400	-2.21157700	-0.44847000
---	------------	-------------	-------------

**APW-MET:****INT**

Pt	-0.76869800	0.00489800	-1.04841500
As	-0.95625900	1.24268400	0.86869100
O	0.48072300	2.46169100	0.36155700
O	-2.51728900	0.20872400	1.48608300
O	-0.16126200	0.79515700	2.32756900
H	0.84338700	0.90796800	2.28017300
O	-1.90196400	2.67426400	1.14950100
H	-2.78296700	2.41668000	1.47764100
N	0.72040100	1.31361300	-1.58204200
H	1.16548600	1.32081900	-2.49562200
N	-2.33207000	-1.08887500	-0.36488600
H	-2.71416700	-1.88063700	-0.87475700
C	1.01793100	2.32701100	-0.81632800
C	1.96533000	3.41176400	-1.22342700
H	1.40900600	4.35183900	-1.27950400
H	2.42619300	3.20201100	-2.18926700
H	2.73823000	3.50371800	-0.45798700
C	-2.92353500	-0.78743400	0.75413800
C	-4.09256500	-1.55335300	1.29062400
H	-3.82254200	-1.96735700	2.26620800
H	-4.39061400	-2.35996600	0.62055400
H	-4.92990400	-0.86634000	1.43696500
O	-0.57732800	-1.42331400	-2.74566100
H	-0.00822900	-2.14494300	-2.41543600
H	-0.15419900	-1.08978400	-3.55438500
N	3.37332600	-0.70065900	-0.85404700
C	2.60576400	-0.61217200	0.43298100
C	2.98247500	0.73473400	1.10067000
O	2.44604300	0.95269000	2.21869300
C	2.84957100	-1.82278200	1.32051000
C	2.51222600	-3.16677400	0.66938700
S	0.84050300	-3.28703800	-0.04845400
C	-0.14713800	-2.87519900	1.41790600
H	3.89802600	-1.84478000	1.64531800
H	2.24985000	-1.67556400	2.22366900
H	3.19619600	-3.40500300	-0.15192900
H	2.62572800	-3.96065700	1.41335900
H	-1.18103800	-3.14265300	1.19427700
H	0.19515600	-3.46613500	2.27190400
H	-0.09800100	-1.80860000	1.65856700
H	1.54945500	-0.56974800	0.14623800
H	4.22882700	-1.25992800	-0.76529300
H	2.82336300	-1.08627800	-1.62776800
H	3.67920000	0.26458600	-1.07559500
O	3.75214900	1.47571500	0.46045300

## TS

Pt	-1.18145900	-0.39034200	-0.69420500
As	-1.73154600	0.89951100	1.10480000
O	-1.44708800	2.57388900	0.13041900
O	-2.04851800	-0.64079400	2.26422900
O	-0.60411800	1.45427200	2.31512800
H	-0.27141400	2.33598800	2.06640200
O	-3.29271500	1.53087000	1.55158200
H	-3.73001700	0.91345100	2.16622700
N	-0.92087100	1.34970300	-1.70714800
H	-0.65663600	1.37780400	-2.68790600
N	-1.53412800	-1.99333900	0.51330500
H	-1.46381300	-2.95729200	0.20055500
C	-1.09724500	2.49798800	-1.12395600
C	-0.91498300	3.80874000	-1.82412900
H	-0.64116500	3.67241400	-2.87034900
H	-0.13448100	4.37700400	-1.31155700
H	-1.84543600	4.37800000	-1.75772300
C	-1.90196000	-1.83008700	1.74976700
C	-2.18467800	-2.97116900	2.67645400
H	-1.52325100	-2.89416400	3.54314500
H	-2.03779200	-3.93433400	2.18767600
H	-3.21538600	-2.89170400	3.03154400
N	4.30540000	1.84946400	0.42449900
C	4.43616900	0.37400200	0.68219700
C	5.96777200	0.07451800	0.76700300
O	6.28016100	-1.12731700	0.80956900
C	3.71345600	-0.45982300	-0.36616600
C	2.19957500	-0.27296700	-0.32829200
S	1.32085800	-1.31601500	-1.53976900
C	1.50704400	-2.93970000	-0.75177700
H	4.10149900	-0.22529300	-1.36561500
H	3.97754000	-1.49981200	-0.15632500
H	1.79667100	-0.49642400	0.66641900
H	1.90561100	0.75222300	-0.57981000
H	0.89956100	-3.64860900	-1.31840500
H	2.54762000	-3.26690500	-0.78156400
H	1.15420000	-2.89799900	0.28176800
H	4.00945500	0.19390600	1.67330000
H	4.13052300	2.06672700	-0.56146300
H	3.58441700	2.31126400	0.98336600
H	5.24703000	2.22940500	0.66383800
O	6.70610500	1.09193500	0.81134300
O	-1.70105400	-1.43621900	-2.80783500
H	-1.45947400	-2.37294900	-2.89000800
H	-2.66007700	-1.40654600	-2.95432600

## PROD

Pt	-0.60598000	-0.02160600	0.08127700
As	-2.90081100	-0.03908900	-0.31849400
O	-3.01133900	-1.99264600	-0.27536800
O	-3.00894200	1.91766800	-0.41000300

O	-3.67011900	-0.05751100	-1.88216600
H	-3.84274400	-0.97874200	-2.15037400
O	-4.16867700	-0.02974800	0.87622400
H	-4.44668800	0.88804300	1.05142200
N	-0.77089000	-2.06228700	0.14345600
H	0.00047800	-2.69830900	0.31780300
N	-0.76699700	2.00859200	-0.02060200
H	0.02597100	2.63214300	0.10257900
C	-1.91466500	-2.65778400	-0.04796000
C	-2.06624600	-4.14756000	-0.02454400
H	-2.78779400	-4.41648900	0.75114200
H	-1.11753900	-4.64956200	0.16527300
H	-2.46876300	-4.47540100	-0.98645300
C	-1.90942600	2.59502700	-0.24335700
C	-2.05360300	4.08322900	-0.32461100
H	-2.45437900	4.34321800	-1.30790500
H	-1.10193500	4.59140900	-0.16982400
H	-2.77396900	4.41061700	0.42942600
N	4.86391300	1.11261600	-1.65108100
C	4.87498100	0.24288700	-0.42617800
C	6.37442600	-0.05367000	-0.10072200
O	6.57099500	-0.84485100	0.83666200
C	4.05995900	-1.02913100	-0.61284100
C	2.55933200	-0.83089800	-0.83156100
S	1.76998200	0.14972000	0.49203500
C	2.09330900	-0.87886700	1.94947400
H	4.45142900	-1.59530900	-1.46790100
H	4.23896200	-1.65106200	0.26851900
H	2.33791900	-0.28192200	-1.75217900
H	2.06353800	-1.80244000	-0.90335500
H	1.47136400	-0.46706400	2.74449200
H	1.83641500	-1.92173900	1.75566600
H	3.14717400	-0.79473400	2.22204500
H	4.47718100	0.84942300	0.39270700
H	4.56206100	0.62121300	-2.49785500
H	4.29388500	1.95686000	-1.55670800
H	5.86924100	1.37253900	-1.77366600
O	7.20579200	0.56356800	-0.81435500
O	0.31347600	1.66911200	2.80484900
H	-0.60521100	1.60829700	2.49952400
H	0.25679900	1.88215900	3.74839200

**APW-HIS:****INT**

Pt	0.87232300	0.58235800	-0.26563800
As	2.54102400	-0.96144900	-0.00786300
O	1.33043600	-2.46710900	-0.00864700
O	3.94933500	0.41000500	0.02830500
O	3.19060600	-1.50901200	1.51418300
H	2.71265000	-2.31044700	1.79634300
O	3.42039700	-1.76467000	-1.27922200
H	4.22039900	-1.25027600	-1.49283000

N	-0.39751600	-1.01395100	-0.35115500
H	-1.41576300	-0.94069800	-0.48052100
N	2.33287100	1.99802700	-0.16095200
H	2.14306500	2.99469000	-0.20848800
C	0.05318500	-2.21914000	-0.19041400
C	-0.82877200	-3.42922600	-0.18444700
H	-0.46979200	-4.13414700	-0.93905600
H	-1.86570600	-3.15753200	-0.38166700
H	-0.74610200	-3.91604600	0.79178200
C	3.58363500	1.66041400	-0.04547700
C	4.69825700	2.65888500	0.01382300
H	5.23232000	2.53586000	0.95957600
H	4.32953100	3.68128400	-0.06885300
H	5.40053300	2.45447100	-0.79836500
O	-0.67763200	2.10842800	-0.39913300
H	-1.42952000	1.77096300	0.21406300
H	-1.02496200	2.16972700	-1.30345000
N	-5.44785300	1.19640800	-1.92230900
C	-4.34896100	0.74273000	-1.00827900
C	-4.33026900	-0.80719500	-1.07571000
O	-5.43233400	-1.33706000	-1.34616500
C	-4.66352800	1.24211900	0.40914500
C	-3.67752700	0.74065900	1.41312700
N	-2.33384800	1.02898100	1.32293600
C	-3.89449300	-0.05804500	2.50383300
C	-1.74929700	0.41444700	2.33463200
N	-2.65869900	-0.25147700	3.07317200
H	-4.66854800	2.33923700	0.40363700
H	-5.66656400	0.90492100	0.69508300
H	-4.79442500	-0.49358300	2.90982000
H	-2.46261500	-0.79361600	3.90520200
H	-0.69022800	0.42310200	2.55378900
H	-3.40940500	1.15656800	-1.37435800
H	-5.13636600	1.26213000	-2.89627200
H	-5.85274700	2.10111800	-1.66216900
H	-6.17219200	0.45744700	-1.87955300
O	-3.23663900	-1.35696400	-0.82403500

## TS

Pt	0.56935800	-0.00771200	-0.39576500
As	2.71300400	0.41390300	0.25836600
O	2.32172700	2.23262600	0.82221100
O	3.33943600	-1.36507800	-0.23243100
O	3.98849900	1.03022300	-0.75470600
H	4.02073800	2.00187500	-0.68281100
O	3.37752300	0.17344000	1.84978300
H	3.77642300	-0.71427000	1.90558400
N	0.16523300	1.90358100	0.19246300
H	-0.73664800	2.35486100	0.07800800
N	1.24497400	-1.84264200	-0.96521800
H	0.66599000	-2.54466400	-1.41772400
C	1.10357800	2.67064600	0.66347900
C	0.87860600	4.09816100	1.05260800

H	1.16257300	4.22864900	2.09978000
H	-0.16071100	4.39577400	0.91347800
H	1.52757100	4.73448200	0.44533300
C	2.48588200	-2.18491700	-0.77831000
C	3.03376900	-3.52389900	-1.16117900
H	3.84763000	-3.38138500	-1.87657300
H	2.26780700	-4.16415700	-1.59842300
H	3.44801600	-4.00211900	-0.26981500
N	-3.83782900	1.34207200	-1.52663300
C	-4.15263400	0.28396500	-0.51614400
C	-5.69903400	0.26398300	-0.32370500
O	-6.10647900	-0.29998600	0.70830100
C	-3.33584000	0.47846400	0.77098200
C	-2.65495700	-0.76999000	1.23206800
N	-1.55239500	-1.27839600	0.56863800
C	-2.96402800	-1.58375900	2.28846900
C	-1.20530000	-2.37402100	1.22328200
N	-2.03151400	-2.59280800	2.26447300
H	-2.56607300	1.24087700	0.60742600
H	-3.99579700	0.84752000	1.55963400
H	-3.74497000	-1.53259500	3.03151600
H	-1.97062000	-3.36469900	2.91660000
H	-0.36894100	-3.01703400	0.98819000
H	-3.89246400	-0.66383000	-0.99826600
H	-2.92590700	1.16578200	-1.97233200
H	-3.82464300	2.27738900	-1.10778900
H	-4.60603400	1.33879700	-2.21467600
O	-6.35251900	0.78557300	-1.26039400
O	-1.31898500	-0.05740000	-1.90143200
H	-1.03880800	-0.23543000	-2.81534000
H	-1.67733400	-0.89267100	-1.54420900

## PROD

Pt	0.32405100	-0.63839100	-0.39871700
As	2.18730300	0.68143500	-0.09941400
O	3.12815600	-0.68644100	0.94726600
O	1.42524100	2.19632300	-1.06764500
O	2.30064100	1.76599100	1.27866800
H	2.72016400	1.30730700	2.03017300
O	3.65371400	0.86110400	-1.01102800
H	3.55166700	1.58957000	-1.65106300
N	1.36486100	-2.08937000	0.59043600
H	1.00542600	-3.02393600	0.75958700
N	-0.49207400	0.98354700	-1.30736900
H	-1.49888200	1.02541000	-1.50422900
C	2.54478500	-1.84749400	1.08416200
C	3.33185300	-2.86705500	1.84866300
H	4.29879200	-3.00848700	1.35944300
H	2.80729600	-3.82065900	1.90980900
H	3.51773600	-2.48520500	2.85610700
C	0.17062200	2.09234300	-1.43156100
C	-0.43991000	3.33521800	-1.99899200
H	-0.35713400	4.13544200	-1.25904300

H	-1.48761200	3.17797000	-2.25813500
H	0.11938800	3.63651500	-2.88882400
N	-2.47487100	1.51749100	2.54931400
C	-2.41265200	0.74357500	1.27083700
C	-3.51395300	1.30527500	0.33246700
O	-4.46086800	1.89607600	0.90149500
C	-2.58083000	-0.75529600	1.56076500
C	-2.57435200	-1.61169800	0.33477900
N	-1.48755100	-1.72090400	-0.51332500
C	-3.56793900	-2.41877500	-0.14598300
C	-1.81936300	-2.56094800	-1.48072700
N	-3.07159000	-3.00202200	-1.28588800
H	-1.76650100	-1.07270800	2.22532900
H	-3.52613600	-0.92059100	2.08994700
H	-4.56281800	-2.61896500	0.22006200
H	-3.55872000	-3.66236100	-1.87949700
H	-1.19018500	-2.85732700	-2.30628600
H	-1.43329700	0.94419000	0.84200900
H	-3.42286500	1.92935600	2.59667200
H	-1.76251200	2.26889200	2.50978300
H	-2.30343400	0.95354600	3.38640300
O	-3.34551100	1.09421100	-0.89051900
O	-0.29145000	2.98521000	1.55422600
H	-0.21287200	3.95061000	1.57274600
H	0.62515300	2.64670500	1.59125800

## Supplementary Materials

# The Se-S bond formation in the covalent inhibition mechanism of SARS-CoV-2 main protease by ebselen-like inhibitors. A computational study

Angela Parise<sup>1,2</sup>, Isabella Romeo<sup>1</sup>, Nino Russo<sup>1</sup> and Tiziana Marino<sup>1\*</sup>

<sup>1</sup> Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy

<sup>2</sup> Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France

\* Correspondence: tiziana.marino65@unical.it

## Table of Contents

Molecular Dynamics procedure

Molecular Dynamics analysis

**Scheme S1.** Proposed mechanism for the covalent inhibition of M<sup>pro</sup>

**Figure S1.** Superposition of crystallographic structures

**Figure S2.** MD most populated structures

**Figure S3.** RMSD analysis

**Figure S4.** RMSF analysis

**Figure S5.** Pocket volume cavity

**Figure S6.** SASA

**Figure S7.** Catalytic dyad distance

**Figure S8.** Salt bridge

**Figure S9.** RDF water

**Figure S10.** Docking protocol setting

**Figure S11.** Best docked poses

**Figure S12.** Specific distances M<sup>pro</sup> –EBS

**Figure S13.** Specific distances M<sup>pro</sup> –EBS-OH

**Figure S14.** Optimized structures of HIE stationary points



**Figure S15.** Free energy profiles of M<sup>pro</sup> HIE

**Table S1.** pKa values

**Table S2.** Donor-Acceptor H-bond M<sup>pro</sup> HID

**Table S3.** Donor-Acceptor H-bond M<sup>pro</sup> HIE

**Table S4.** RMSD-based clustering

**Table S5.** Docking scores

**Table S6.** NBO charges of EBS and EBS-OH

**Table S7.** NBO charges by atoms (M<sup>pro</sup> HID – EBS)

**Table S8.** NBO charges by atoms (M<sup>pro</sup> HIE – EBS)

**Table S9.** NBO charges by atoms (M<sup>pro</sup> HID – EBS-OH)

**Table S10.** Energy values corrected including free energy (kcal/mol) in gas phase for the three used basis sets for the inhibition phase

**References**

## Molecular Dynamics procedure

The crystal structure of main protease bound to potent broad-spectrum non-covalent inhibitor X77 with PDB code 6W63<sup>1</sup> were used as starting point. The structure is characterized by one chain of 306 amino acids. The apo-enzyme was obtained by removing the inhibitor from the crystallographic structure. The structure thus obtained was superimposed on other crystallographic structures (PDB codes: 6LU7<sup>2</sup> and 6Y2F<sup>3</sup>) with an RMSD value lower than 1 Å. (Figure S1).

The absent hydrogen atoms were added using H++ server<sup>4</sup> to calculate the protonation states of titratable residues at pH 7.4 (Table S1). The protonation states of the catalytic residue neutral His41, with the hydrogen on the side chain N $\epsilon$  (HIE) and N $\delta$  (HID), have been considered. 300 ns MDs were performed for apo-M<sup>pro</sup>, in the two protonation states (HIE and HID).

Each system has been solvated in an orthorhombic box with a buffer of 10 Å, using TIP3P water model. Na<sup>+</sup> ions have been added to ensure zero total charge. The solvated structure was minimized in stages using AMBER 16.<sup>5</sup> In particular, water molecules and counterions were minimized over 2500 steps of steepest descent and 2500 steps of conjugate gradient minimization with a 50 kcal mol<sup>-1</sup> Å<sup>-2</sup> constraint on the solute. Subsequently, hydrogen atoms were minimized over 5000 steps of steepest descent and 5000 steps of conjugate gradient with a 50 kcal mol<sup>-1</sup> Å<sup>-2</sup> constraint on all heavy atoms. Next, 5000 steps of steepest descent and 5000 steps of conjugate gradient minimization was performed, with 50 kcal mol<sup>-1</sup> Å<sup>-2</sup> constraint on protein backbone. Finally, the whole system has been released with 3000 and 6000 steps of unconstrained steepest descent and conjugate gradient respectively. Each stage of minimization was run until the root-mean-square of the Cartesian gradient was low than 10<sup>-4</sup> kcal mol<sup>-1</sup> Å<sup>-1</sup>. Then a progressive heating phase was carried out from 0 to 310 K for 0.5 ns using the Langevin thermostat (  $g = 2.0$  ) in NVT ensemble. The production phase consist of 300 ns of MDs using FF14SB force field<sup>6</sup> under the following conditions: integration step of 2 fs coupling SHAKE algorithm; NPT ensemble at 1 bar pressure using the Berendsen barostat<sup>7</sup> with a time constant  $\tau_p = 2.0$  ps. The Particle mesh Ewald summation method<sup>8</sup> has been employed for the electrostatic potential long-range interactions with a 12 Å cut-off distance.

In order to select different representative conformations of the system, root-mean square deviation (RMSD) based clustering of the whole trajectory has been performed using the an agglomerative bottom-up approach (hieragglo algorithm) available in Amber16 tools. After removing overall rotations and translations by RMS-fitting the C $\alpha$  atoms' positions of the trajectory, the average linkage clustering algorithm has been applied, identifying 10 representative conformations of the protein (Figure S2 and Table S2).

Afterwards, the representative structures with the highest percentage of population (Table S2) were considered for the further molecular docking studies, performed by using AutoDock (version 4.2).<sup>9</sup>

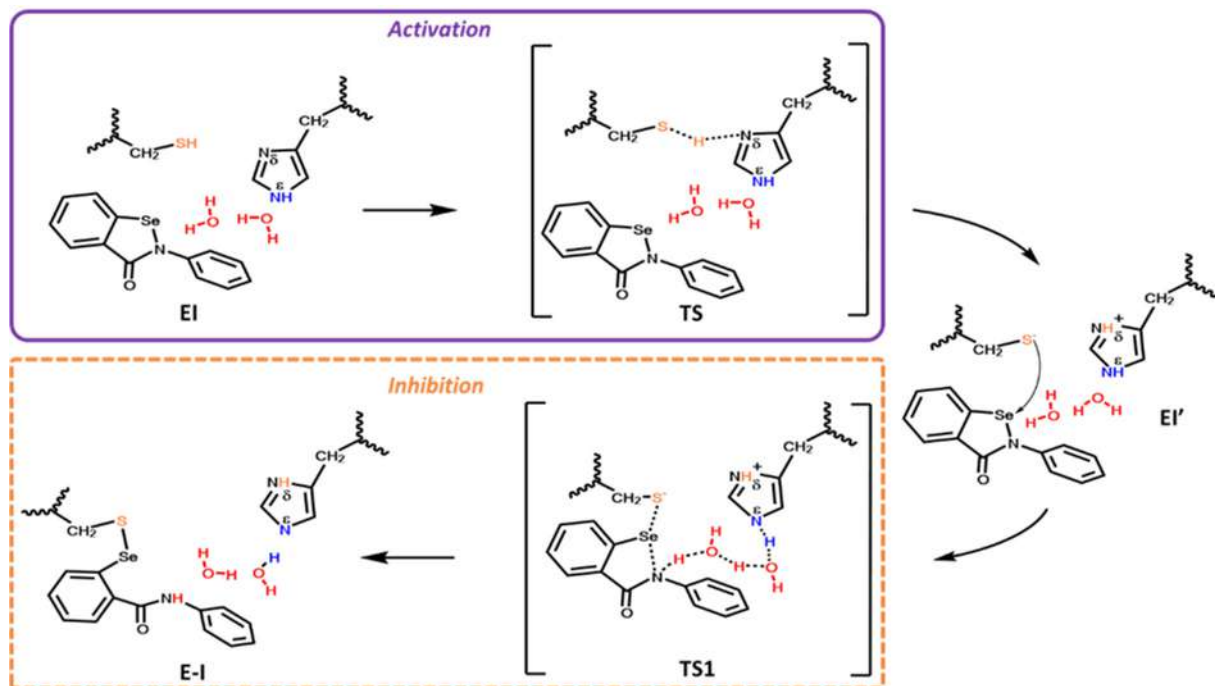
The ligands (EBS and EBS-OH) and the SARS-CoV-2 M<sup>pro</sup> models, apo-HID and apo-HIE, were processed using the AutoDock Tools<sup>9</sup> to obtain the PDBQT (Protein Data Bank, Partial Charge (Q), & Atom Type (T)) coordinate files containing the information needed by AutoGrid and AutoDock, namely polar hydrogen atoms, partial charges, correct atom types, and information on the articulation of flexible molecules. In order to verify if the adopted docking procedure is adequate, we have performed the docking of X77 inhibitor present in the used crystallographic structure. The results, reported in Figure S10 (as docking protocol setting) return a good superposition with a RMSD of 1.91 Å. In particular, Gasteiger-Marsili charges<sup>10</sup> were loaded in ADT (Auto Dock Tools). The center of the grid box locates on the sulfur atom of Cys145, and a dimension box of 70x70x70 Å has been chosen to abundant cover the active site. For ligand conformational searching the Lamarckian genetic algorithm (LGA)<sup>11</sup> has been used. The docking conditions were as follows: 10 independent runs, population size of 150, random starting position and conformation, local search rate of 0.6 and 2500000 energy evaluations. Final docked poses have been clustered using RMSD tolerance of 0.5 Å (Table S5).

The best docking pose was chosen according to the most negative docking score (Figure S11).

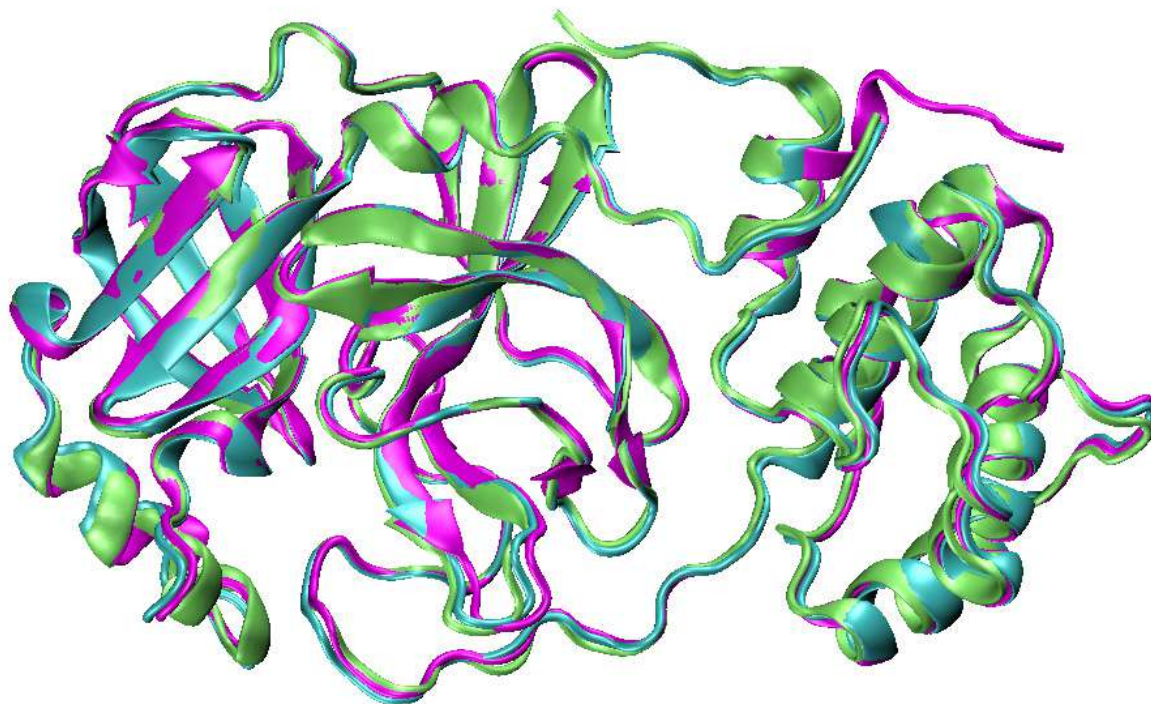
In order to obtain ebselen-like compound (EBS and EBS-OH) parameters gas phase geometry optimization has been carried out using B3LYP/6-31G\*. Atomic charges are derived by fitting the electrostatic potential according to the Merz–Singh–Kollman scheme,<sup>12</sup> using the RESP fitting procedure. Antechamber and parmchk modules of Amber16 has been used for generating preparatory files to perform MM relaxation of the complexes HID-EBS, HIE-EBS and HID-EBS-OH.

### **Molecular Dynamics analysis**

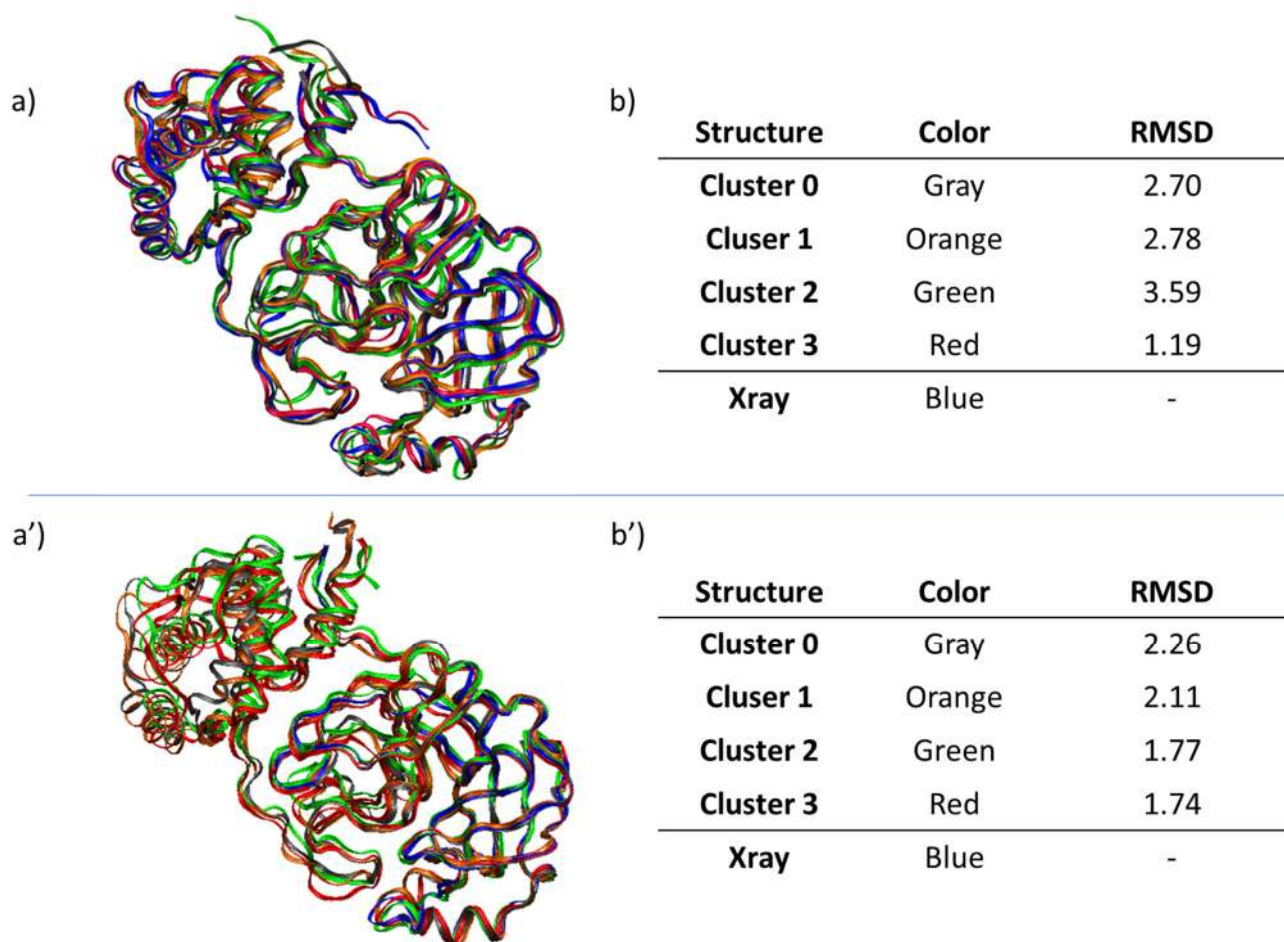
Root-mean-square deviation (RMSD) of the protein M<sup>pro</sup> HID and M<sup>pro</sup> HIE backbone, RMSD of the active site, root mean square fluctuation (RMSF), solvent accessible surface area (SASA), radial distribution function (RDF), hydrogen (H)-bond analysis and RMSD-based clustering MD simulation were performed using cpptraj module of AmberTools 16.<sup>5</sup> Electrostatic calculations were performed with the Adaptive Poisson-Boltzmann Solver (APBS).<sup>13</sup> Volume cavity of the binding pocket was performed with CAVER 2.0 Analyst<sup>14</sup>, a very rough estimation of the volume. The volume of the pocket, in cubic Å, is computed by random sampling the bounding box of filling balls with a radius of 5 Å.



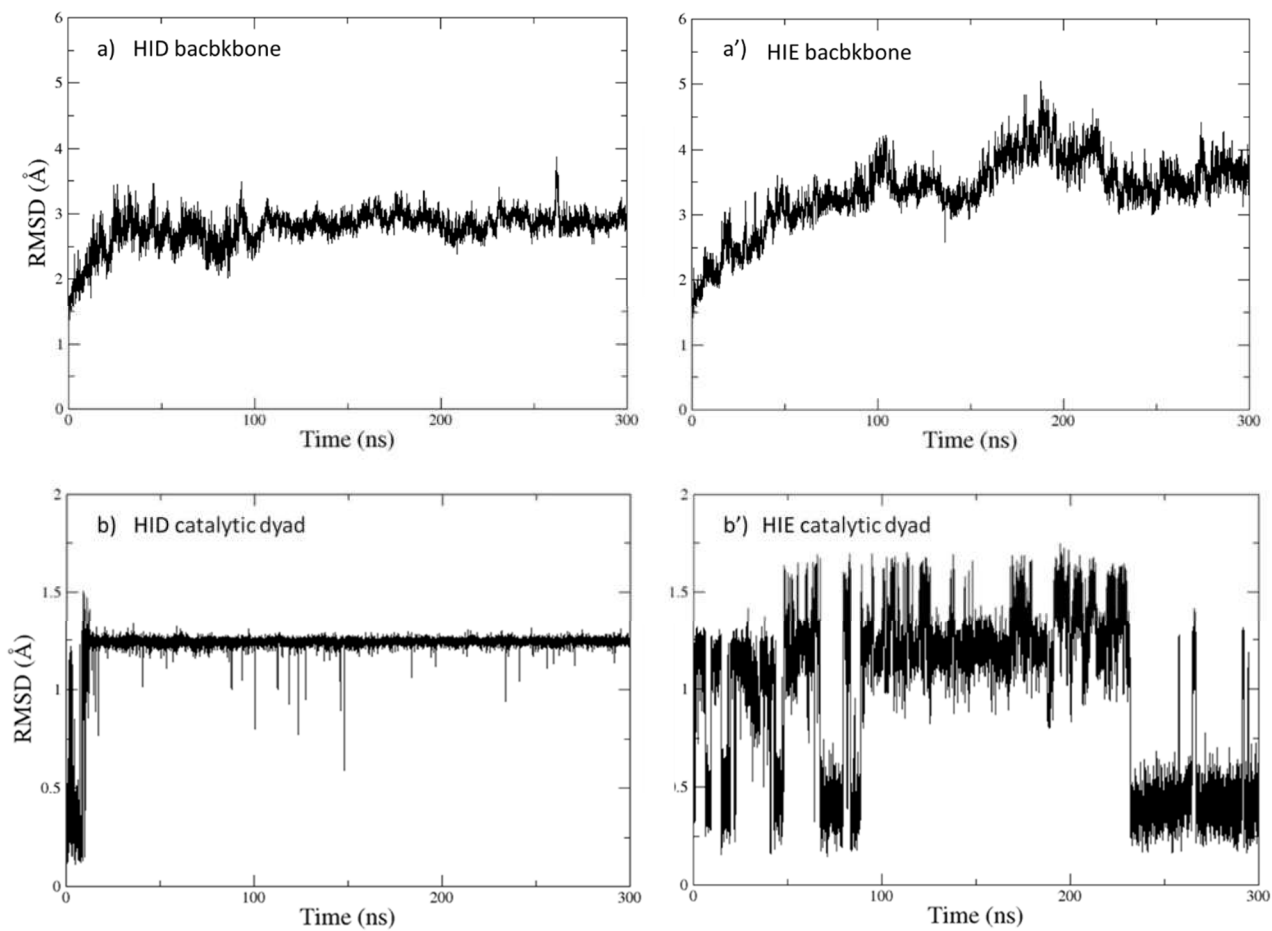
**Scheme S1.** Proposed mechanism for the covalent inhibition of M<sup>pro</sup> SARS-CoV-2 for HIE by EBS in two phases: activation (in violet) and inhibition through covalent bond formation (in dashed orange).



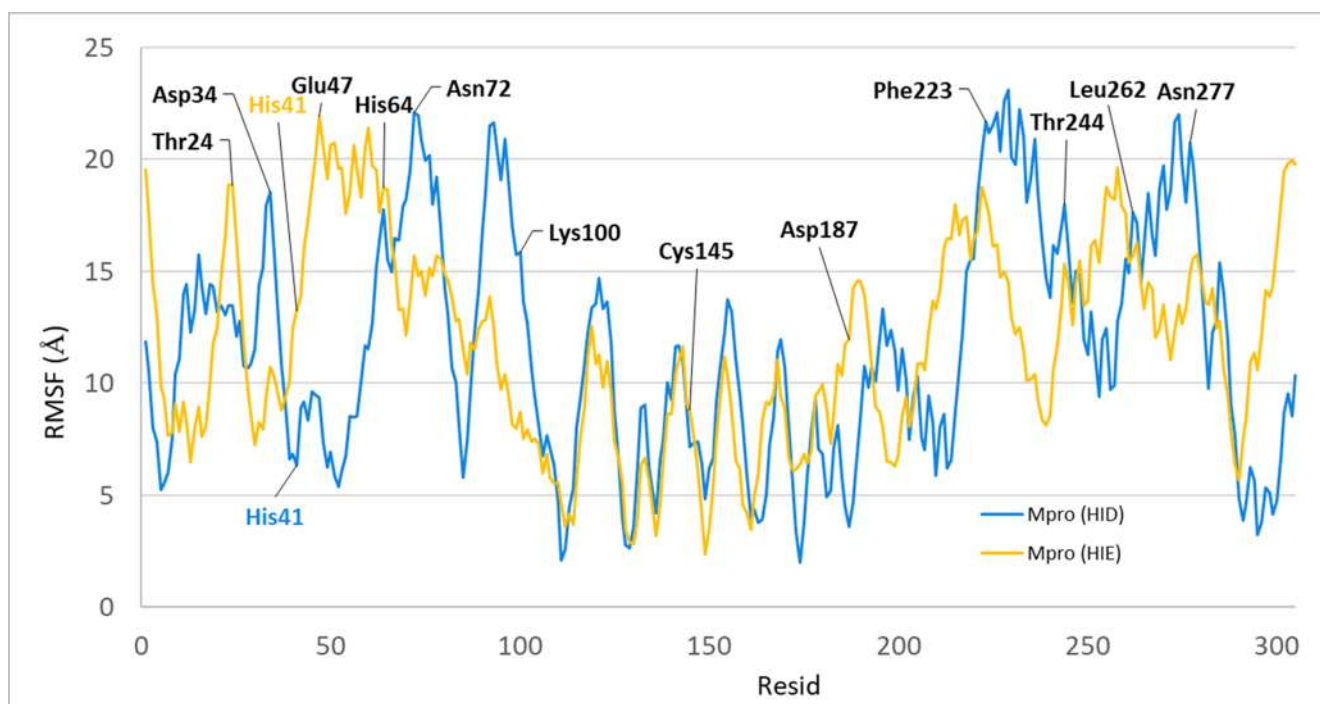
**Figure S1.** Superposition of M<sup>pro</sup> crystallographic structures: 6W63 in cyan, 6LU7 in magenta, and 6Y2F in lime.



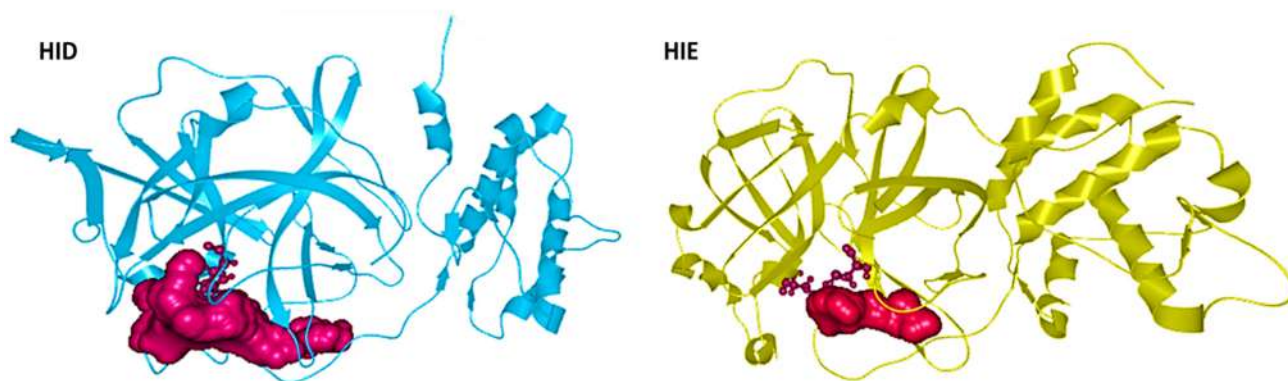
**Figure S2.** Superposition of 4 most populated structures obtained by RMSD-based clustering MD simulation of (a) M<sup>Pro</sup> HIE and (a') M<sup>Pro</sup> HID. RMSD values of HIE clusters (b) with respect to X-Ray 6W63, RMSD values of HID clusters (b') with respect to X-Ray 6W63.



**Figure S3.** Up: backbone RMSD trend of a)  $M^{\text{Pro}}$  (HID) and a')  $M^{\text{Pro}}$  (HIE) along MD simulation time, expressed in Å. Down: RMSD trend of residues His41 and Cys145 b)  $M^{\text{Pro}}$  (HID) and b')  $M^{\text{Pro}}$  (HIE).

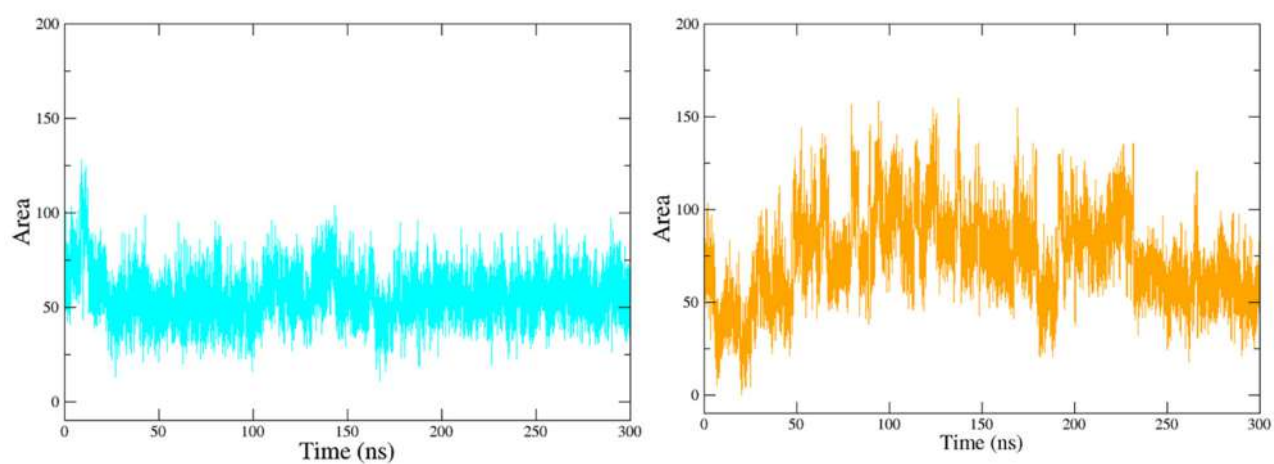


**Figure S4.** RMSF comparison between M<sup>pro</sup> (HID) in blue and M<sup>pro</sup> (HIE) in yellow; some residues with high DRMSF value between the two protonation states of His41 simulations are explicitly reported in the graph.

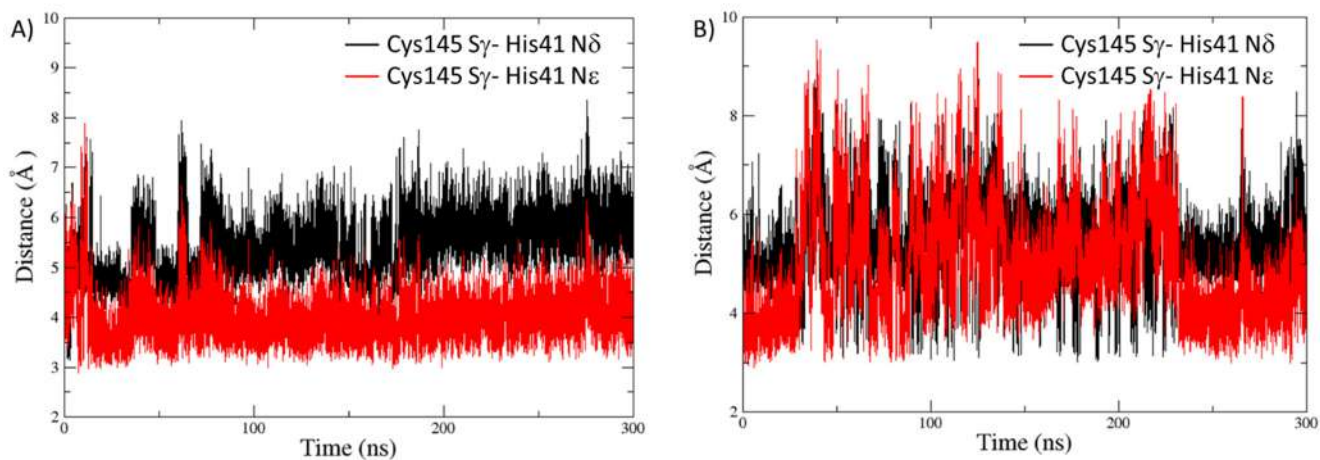


**Figure S5.** Pocket volume cavity of binding site: value of 598.1 Å<sup>3</sup> for M<sup>pro</sup> HID (in blue) and value of 227.7 Å<sup>3</sup> M<sup>pro</sup> HIE (in gold). The calculation was performed with CAVER Analyst 2.0. 5 Å sphere size used to define the pocket of each system for volume calculations.

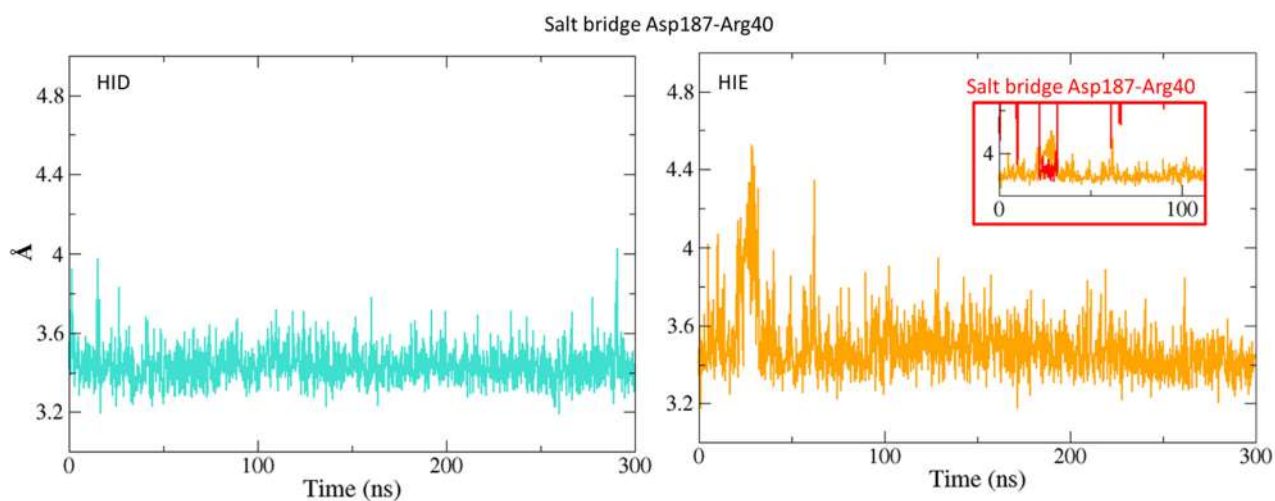




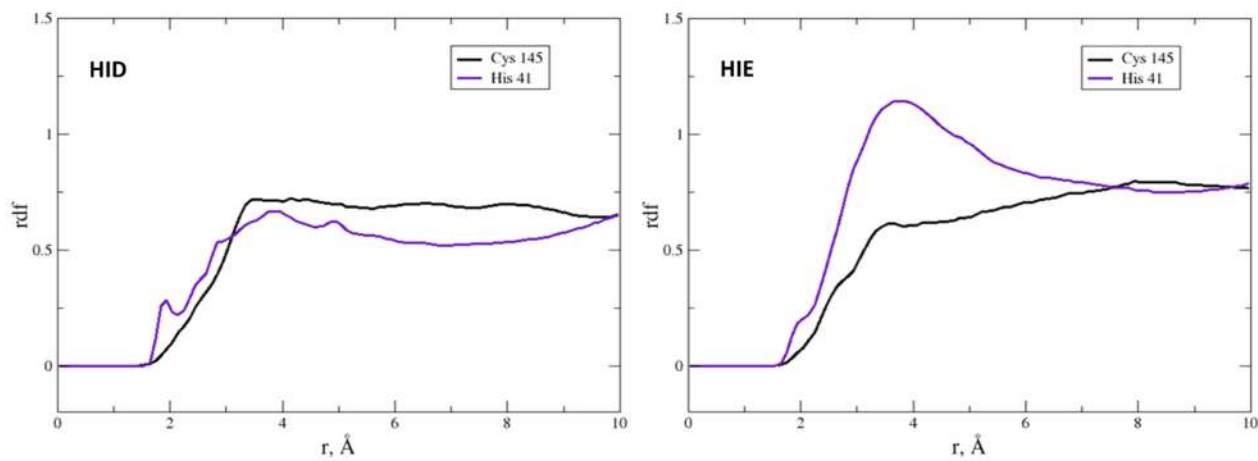
**Figure S6.** Solvent accessibility surface area (SASA) along the trajectory, expressed in  $\text{\AA}^2$  and calculated based on the residues Cys145 and His41 in different protonation states: HID in orange and HID in blue. SASA was calculated with CPPTRAJ using LCPO algorithm.<sup>15</sup>



**Figure S7.** Distance between N $\delta$  and S $\gamma$  (red) and N $\delta$  and S $\gamma$  (black) in the catalytic residues His41 and Cys145. Graph A) refers to HID and graph B) refers to HIE trajectories.



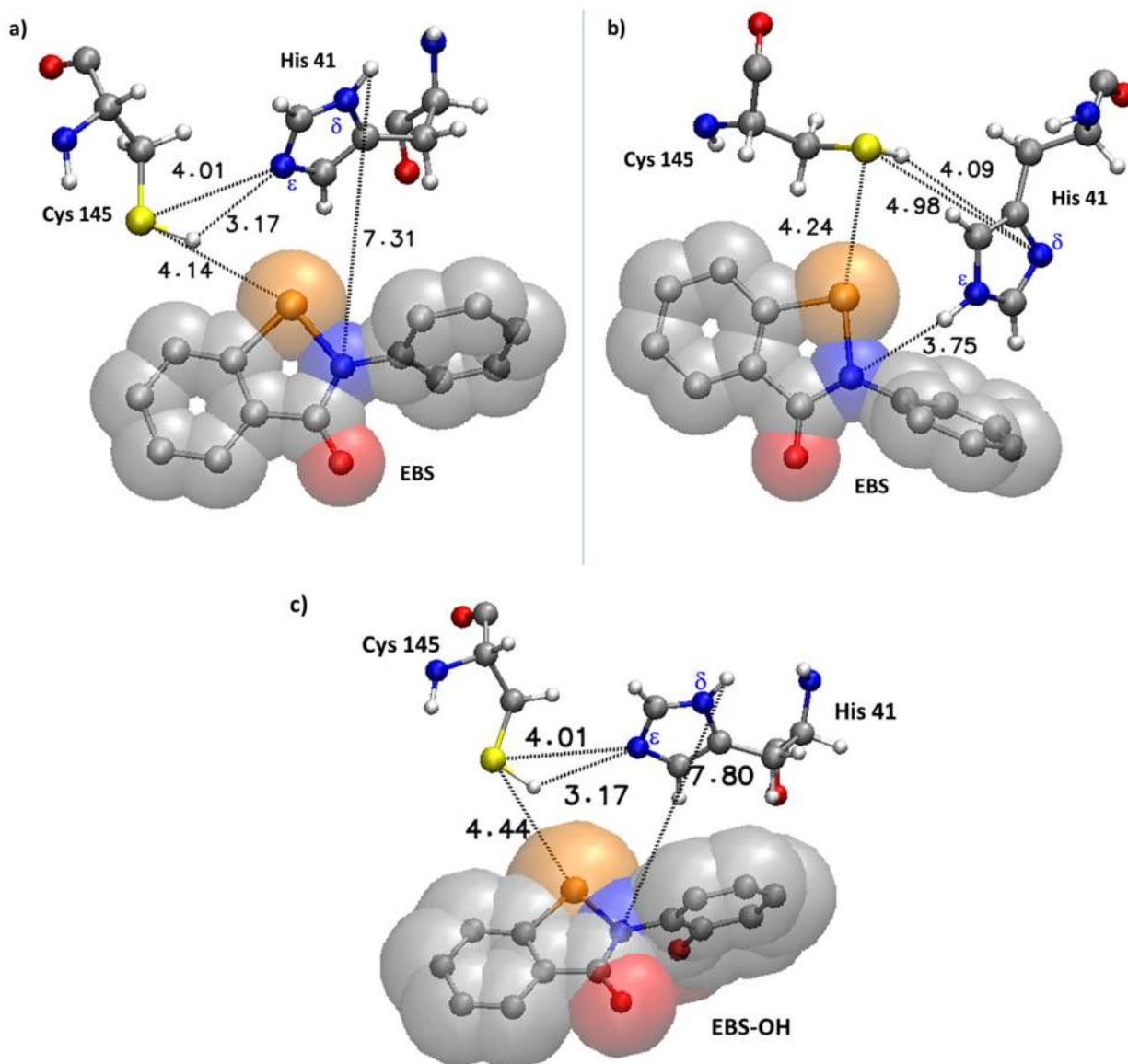
**Figure S8.** Salt bridges between residues Asp186 and Arg40 along the MD simulation trajectory, for HID and HIE, cyan and orange respectively. For the HIE trajectory, an additional data of salt bridge Asp187-Arg40 is zoom and shown in the red square. Analysis was performed by VMD plugin Version 1.1 with the cutoff distance of 3.2 Å.



**Figure S9.** Water radial distribution function (RDF) analysis obtained as function of the distance ( $r$ ) expressed in Å between water oxygen and side chains of Cys145 (black line) and His41 (purple line), for HID on left and for HIE on right.



**Figure S10.** Superposition of X77 structure obtained by the best docking pose (in pink) and X-Ray structure (colored by atom) with a RMSD of 1.91 Å.

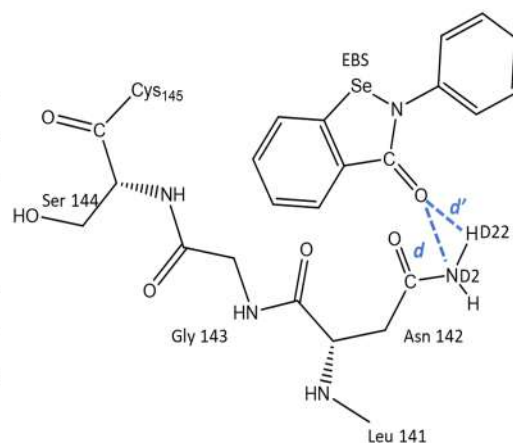


**Figure S11.** Structure and main distances (in Å) between atoms of  $M^{pro}$  active site and inhibitor best docked pose, a) refers to EBS docked in the  $M^{pro}$  HID most populated cluster, b) refers to EBS docked in the  $M^{pro}$  HIE most populated cluster and c) is the best docked pose of EBS-OH in the catalytic pocket of  $M^{pro}$  HID; van der Waals radii of EBS and EBS-OH atoms are shown in transparent.

HID-EBS	EI	TS	EI'	TS1	E-I
$d$ (O-ND2) Å	2.94	3.13	2.99	3.10	2.78
$d'$ (O-HD22) Å	1.95	2.25	2.12	2.22	1.75

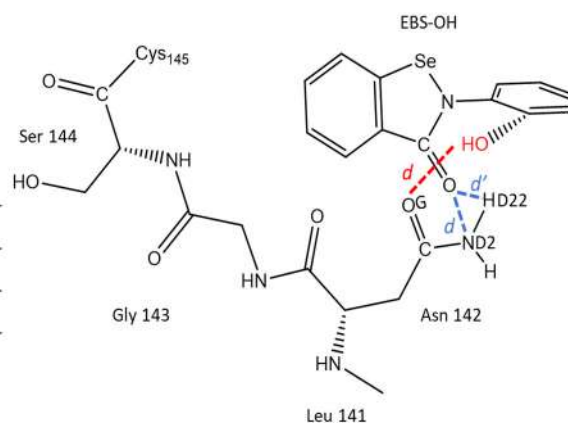
  

HIE-EBS	EI	TS	EI'	TS1	E-I
$d$ (O-ND2) Å	4.49	4.49	3.83	3.60	3.13
$d'$ (O-HD22) Å	3.93	3.12	3.33	3.11	2.46

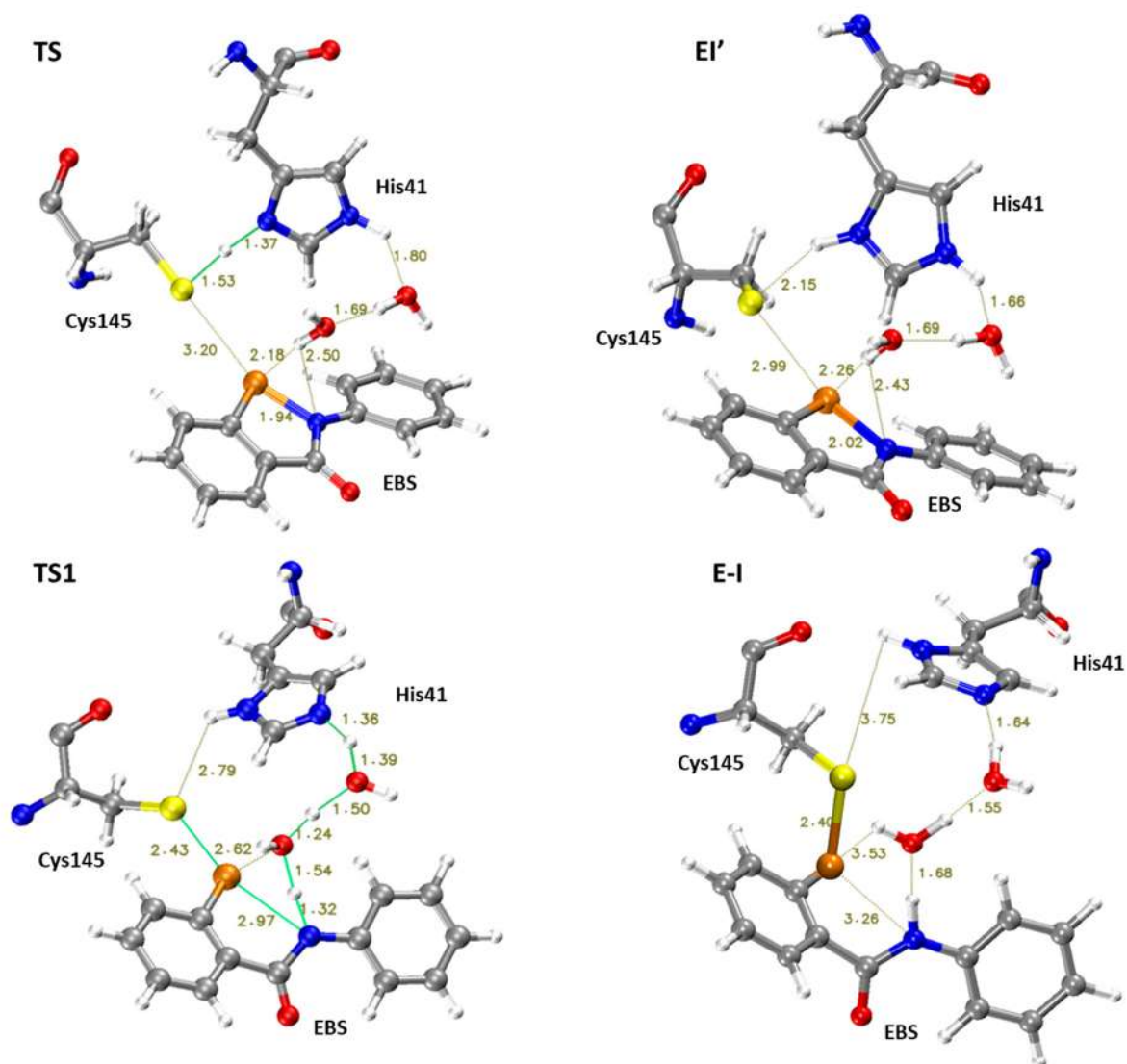


**Figure S12.** Tables containing distances between EBS oxygen atom and Asn 142 atoms ND2 ( $d$ ) and HDD2 ( $d'$ ) as depicted in the included scheme on the right. Table on top refers to HID; table down refers to HIE.

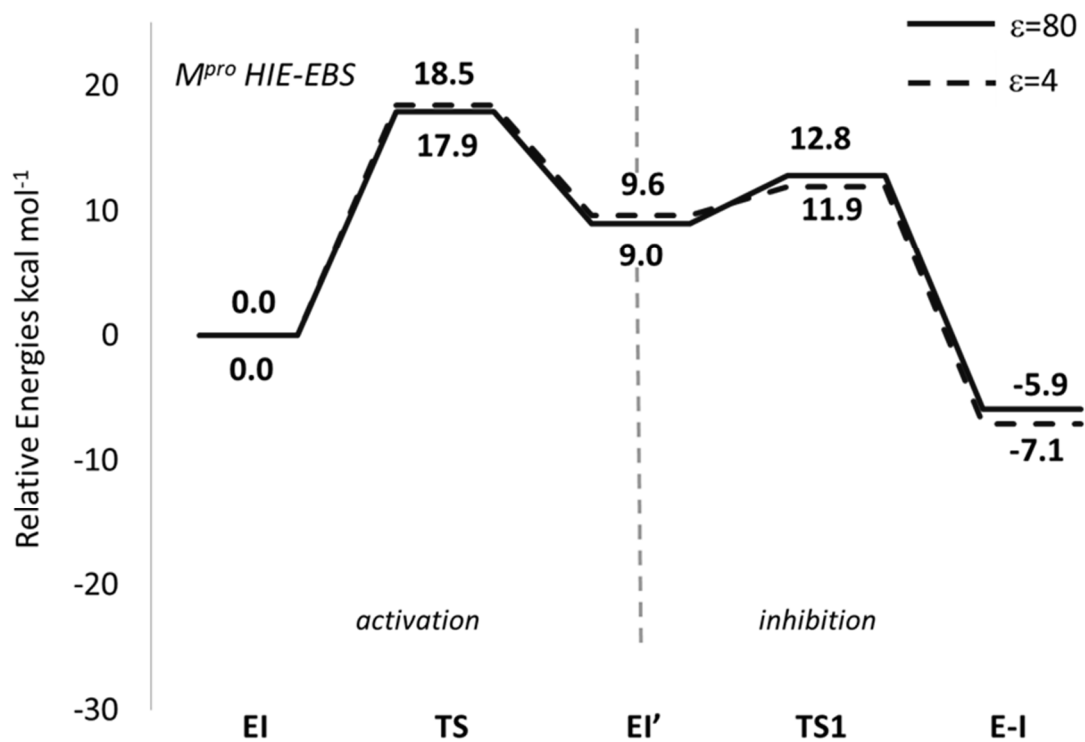
HID-EBS-OH	EI	TS	EI'	TS1	E-I
$d$ (O-ND2) Å	5.95	3.04	3.02	2.83	2.73
$d'$ (O-HD22) Å	5.12	2.66	2.42	1.83	1.74
$d$ (H-OG) Å	1.77	1.67	1.68	1.88	1.81



**Figure S13.** Tables containing distances between some selected atoms of EBS-OH and Asn 142 side chain. Distance between carbonyl oxygen O of EBS-OH and Asn 142 ND2 ( $d$ )/HDD2 ( $d'$ ) are in blue, distance between H of EBS-OH hydroxyl group and OG of Asn 142 is in red, as depicted in the included scheme on the right.



**Figure S14.** Optimized structures of the stationary points involved in the  $M^{Pro}$  HIE inhibition process promoted by EBS. For clarity, only residues of the model implicated in the chemical event are shown. Main distances are in Å. Imaginary frequencies are reported in  $\text{cm}^{-1}$  for all transition states.



**Figure S15.** B3LYP-D3/6-31+G (2d,2p) free energy profiles of  $M^{pro}$  HIE SARS-CoV-2 covalent inhibition mechanism by EBS in two phases: activation and inhibition through covalent bond formation, expressed in kcal mol<sup>-1</sup>. Solid line indicates calculations in water ( $\epsilon=80$ ) and dotted line in protein environment ( $\epsilon=4$ ).



**Table S1.** Calculated pKa for ionizable residues of M<sup>pro</sup>. Residues fully protonated (positively charged) and deprotonated (negatively charged) are colored in blue and red respectively. Residues belonging to the catalytic dyad are in bold.

	<i>pKa</i>	Residue	<i>pKa</i>
Arg4	11.81	Asp153	4.40
Lys5	8.45	Tyr154	9.67
Lys12	9.30	Asp155	4.37
Glu14	3.81	Cys156	9.62
Cys16	12.15	Cys160	10.25
Cys22	11.90	Tyr161	11.64
Asp33	3.87	His163	4.17
Asp34	4.50	His164	3.75
Tyr37	11.32	Glu166	4.67
Cys38	12.12	His172	4.08
Arg40	11.20	Asp176	2.86
<b>His41</b>	<b>4.42</b>	Glu178	3.92
Cys44	9.99	Tyr182	13.32
Glu47	4.10	Asp187	5.09
Asp48	3.37	Arg188	11.61
Tyr54	13.17	Asp197	4.16
Glu55	4.55	Tyr209	13.34
Asp56	4.46	Asp216	4.29
Arg60	12.16	Arg217	11.81
Lys61	10.19	Arg222	11.97
His64	6.37	Asp229	4.15
Arg76	11.77	Lys236	10.40
His80	7.39	Tyr237	10.96
Cys85	10.18	Tyr239	11.16
Lys88	10.65	Glu240	4.60
Lys90	10.18	Asp245	3.56
Asp92	3.81	His246	3.97
Lys97	10.43	Asp248	4.05
Lys100	10.04	Asp263	3.92
Tyr101	10.51	Cys265	11.78
Lys102	9.91	Lys269	9.18
Arg105	11.70	Glu270	4.58
Cys117	10.91	Arg279	12.06
Tyr118	11.68	Glu288	4.07
Tyr126	13.65	Asp289	3.38
Cys128	10.29	Glu290	4.52
Arg131	10.66	Asp295	4.14
Lys137	9.56	Arg298	10.30
<b>Cys145</b>	<b>8.94</b>	Cys300	10.05

**Table S2.** Donor-Acceptor hydrogen bond details of 300 ns M<sup>pro</sup> HID simulation. Only H-bonds involving the side chains atoms of amino acid residues have been reported It also shows the occupancy of each hydrogen bond.

A residue	A atom	D residue	D atom	Freq
GLU240	O	THR201	OG1	0.9789
ASN133	OD1	THR135	OG1	0.9790
CYS160	O	TYR182	OH	0.9791
ASP176	O	THR175	OG1	0.9793
ASP187	OD1	ARG40	NH2	0.9794
ASP263	O	SER267	OG	0.9797
<b>ASP187</b>	<b>OD2</b>	<b>ARG40</b>	<b>NE</b>	<b>0.9798</b>
ASN28	OD1	GLY120	N	0.9804
ILE259	O	SER254	OG	0.9807
ILE259	O	TYR209	OH	0.9808
PHE181	O	ARG105	NH2	0.9811
LEU253	O	THR257	OG1	0.9812
LEU141	O	SER144	OG	0.9815
GLN127	OE1	SER113	OG	0.9819
ASP289	O	ASN203	ND2	0.9821
TRP31	O	ASN95	ND2	0.9824
GLY195	O	ASN133	ND2	0.9836
ASN231	OD1	LEU242	N	0.9838
CYS145	O	ASN28	ND2	0.9841
ASP289	OD2	ARG131	NH2	0.9844
ASP295	OD2	THR111	OG1	0.9852
ASN180	O	ARG105	NH1	0.9863
ASN65	OD1	SER62	N	0.9868
SER113	OG	PHE8	N	0.9870
ASN95	OD1	LYS97	N	0.9878
THR280	O	TRP218	NE1	0.9883
ASP176	OD2	ARG105	NH1	0.9885
ALA210	O	ASN214	ND2	0.9886
LEU286	O	SER284	OG	0.9887

**Table S3.** Donor-Acceptor hydrogen bond details of 300 ns M<sup>pro</sup> HIE simulation. Only H-bonds involving the side chains atoms of amino acid residues have been reported. It also shows the occupancy of each hydrogen bond.

A residue	A atom	D residue	D atom	Freq
CYS160	O	TYR182	OH	0.9519
GLU240	O	THR201	OG1	0.9502
ASP263	O	SER267	OG	0.8845
ASN133	OD1	THR135	OG1	0.8843
ASP176	O	THR175	OG1	0.8626
ASP187	OD1	ARG40	NH2	0.8189
ILE259	O	TYR209	OH	0.7847
ASN28	OD1	GLY120	N	0.7741
ILE259	O	SER254	OG	0.7659
LEU253	O	THR257	OG1	0.7652
LEU141	O	SER144	OG	0.7566
ASN231	OD1	LEU242	N	0.7233
TRP31	O	ASN95	ND2	0.7071
THR175	O	MET162	N	0.7016
ASP187	OD2	ARG40	NE	0.6625
ASP289	O	ASN203	ND2	0.6545
PHE181	O	ARG105	NH2	0.6023
CYS145	O	ASN28	ND2	0.5805
SER144	O	SER147	OG	0.5753
ASP48	OD2	THR45	OG1	0.5590
LEU286	O	SER284	OG	0.5453
ASN95	OD1	LYS97	N	0.5163
ASP176	OD1	ASN180	N	0.5045

**Table S4.** Tables of 10 structures obtained by RMSD-based clustering of MD simulations. M<sup>Pro</sup> HID (top) and M<sup>Pro</sup> HIE (down). From first column: cluster name, number of frames included in the cluster, population respect to total processed frames in percentage, average distance between frames inside the cluster expressed in Å, standard deviation related to Avg d and the centroid frame.

<b>Name</b>	<b>Frames</b>	<b>Pop %</b>	<b>Avg d (Å)</b>	<b><math>\sigma</math></b>	<b>Centroid</b>
<b>Cluster 0</b>	6594	44.0	2.09	0.40	9055
<b>Cluster 1</b>	3467	23.1	2.06	0.41	10699
<b>Cluster 2</b>	1978	13.2	2.01	0.38	1017
<b>Cluster 3</b>	999	6.7	1.86	0.36	2320
<b>Cluster 4</b>	699	4.7	2.00	0.39	227
<b>Cluster 5</b>	675	4.5	1.90	0.39	469
<b>Cluster 6</b>	269	1.8	1.72	0.43	4677
<b>Cluster 7</b>	176	1.2	1.64	0.38	64
<b>Cluster 8</b>	108	0.7	1.37	0.19	13153
<b>Cluster 9</b>	35	0.2	0.00	0.00	16

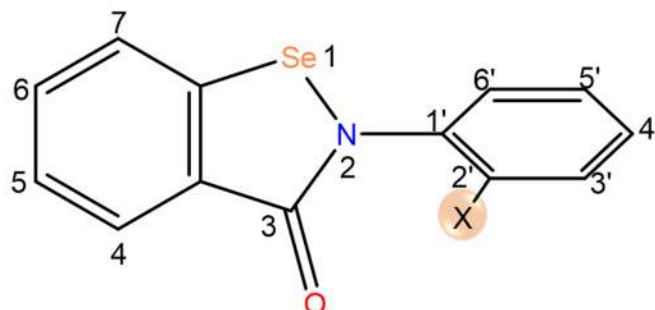
<b>Name</b>	<b>Frames</b>	<b>Pop %</b>	<b>Avg d (Å)</b>	<b><math>\sigma</math></b>	<b>Centroid</b>
<b>Cluster 0</b>	5012	25.1	1.71	0.34	15341
<b>Cluster 1</b>	4213	21.1	1.86	0.39	7693
<b>Cluster 2</b>	3456	17.3	1.74	0.38	18467
<b>Cluster 3</b>	2302	11.5	1.60	0.33	939
<b>Cluster 4</b>	1553	7.8	1.66	0.37	9760
<b>Cluster 5</b>	1551	7.8	1.60	0.29	2581
<b>Cluster 6</b>	1276	6.4	1.50	0.25	4517
<b>Cluster 7</b>	323	1.6	1.38	0.26	3995
<b>Cluster 8</b>	290	1.5	1.38	0.30	17856
<b>Cluster 9</b>	25	0.1	0.00	0.00	2349

**Table S5.** Docking scores values related to the free energy of protein-ligand binding affinity expressed in kcal mol<sup>-1</sup> and it's relative RMSD in Å. a) M<sup>pro</sup> HID most representative cluster and EBS, b) M<sup>pro</sup> HIE most representative cluster and EBS and c) M<sup>pro</sup> HID with EBS-OH. Results are reported for all 10 docking poses ranked by energy. Best docking poses values are shown in red.

Pose	a) M <sup>pro</sup> HID and EBS		b) M <sup>pro</sup> HIE and EBS		c) M <sup>pro</sup> HID and EBS-OH	
	E (kcal mol <sup>-1</sup> )	RMSD (Å)	E (kcal mol <sup>-1</sup> )	RMSD (Å)	E (kcal mol <sup>-1</sup> )	RMSD (Å)
1	<b>-5.85</b>	<b>0.00</b>	<b>-6.39</b>	<b>0.00</b>	<b>-5.67</b>	<b>0.00</b>
2	-5.85	0.02	-6.39	0.01	-5.48	1.59
3	-5.85	0.01	-6.39	0.02	-5.47	1.58
4	-5.85	0.06	-6.39	0.02	-5.27	1.75
5	-5.85	0.05	-6.38	0.11	-5.27	1.74
6	-5.85	0.05	-6.38	0.14	-5.26	1.73
7	-5.84	0.08	-6.38	0.12	-5.23	0.45
8	-5.84	0.07	-6.38	0.06	-5.21	1.71
9	-5.84	0.11	-6.38	0.09	-5.18	1.39
10	-5.55	0.00	-6.38	0.04	-5.04	1.37

**Table S6.** NBO charge values of heavy atoms in |e| for EBS and EBS-OH. Atoms numbering follows the sketch on the right.

	EBS	EBS-OH
Se 1	0.634	0.622
N 2	-0.602	-0.661
C 3	0.673	0.665
O 3	-0.618	-0.597
C 4	-0.177	-0.177
C 5	-0.252	-0.252
C 6	-0.213	-0.212
C 7	-0.257	-0.258
C 1'	0.152	0.053
C 2'	-0.254	-0.202
C 3'	-0.229	-0.270
C 4'	-0.250	-0.218
C 5'	-0.231	-0.285
C 6'	-0.273	0.322
O 6' (OH)	-	-0.702



**Table S7.** NBO charge values of heavy atoms in  $|e|$  for every stationary point along the potential energy surface of inhibition promoted by EBS on M<sup>Pro</sup> HID. Heavy atoms of the inhibitor, of the two catalytic waters, of the catalytic dyad and of the residues immediately close to the dyad are reported.

Residue	Atom	EI	TS	EI'	TS1	E-I
His 41	N	-0.100	-0.027	-0.002	-0.224	-0.066
His 41	CA	0.591	0.123	-0.015	0.181	0.337
His 41	CB	-0.912	-0.857	-0.727	0.124	-0.885
His 41	CG	0.515	0.599	0.375	0.145	-0.079
His 41	ND1	-0.219	-0.003	-0.007	0.152	-0.193
His 41	CE1	0.469	0.135	0.908	0.315	0.266
His 41	NE2	-0.223	0.274	-0.026	0.066	0.086
His 41	CD2	0.069	0.189	0.374	0.139	0.639
His 41	C	0.163	0.410	0.464	0.597	0.236
His 41	O	-0.375	-0.424	-0.427	-0.598	-0.364
Asn 142	N	0.044	-0.088	-0.097	-0.181	0.140
Asn 142	CA	0.068	0.683	0.691	0.252	-0.109
Asn 142	CB	-0.231	-0.312	-0.326	-0.025	-0.273
Asn 142	CG	0.813	0.550	0.541	0.949	0.853
Asn 142	OD1	-0.472	-0.454	-0.449	-0.788	-0.464
Asn 142	ND2	0.029	0.126	0.123	-0.180	0.134
Asn 142	C	0.073	-0.306	-0.290	0.791	-0.081
Asn 142	O	-0.475	-0.510	-0.516	-0.752	-0.435
Gly 143	N	0.259	0.170	0.174	-0.313	0.129
Gly 143	CA	-0.171	0.075	0.074	0.296	-0.267
Gly 143	C	0.345	0.181	0.173	0.836	0.680
Gly 143	O	-0.338	-0.402	-0.405	-0.704	-0.492
Ser 144	N	0.217	0.215	0.226	-0.405	0.311
Ser 144	CA	0.113	0.364	0.384	0.240	0.343
Ser 144	CB	-0.039	0.123	0.115	0.387	-0.174
Ser 144	OG	-0.107	-0.098	-0.099	-0.298	-0.093
Ser 144	C	0.479	0.141	0.115	1.043	0.178
Ser 144	O	-0.533	-0.572	-0.585	-0.925	-0.511
Cys 145	N	0.298	0.373	0.344	-0.509	0.407
Cys 145	CA	0.683	-0.135	-0.290	0.196	0.530
Cys 145	CB	-0.782	-0.528	-0.500	0.061	-0.495
Cys 145	SG	0.143	-0.074	0.542	-0.461	-0.077
Cys 145	C	-0.092	0.129	0.439	1.020	-0.447
Cys 145	O	-0.414	-0.387	-0.527	-0.960	-0.389
Gly 146	N	0.086	0.068	0.074	-0.473	0.099
Gly 146	CA	0.163	0.189	0.178	0.260	0.149
Gly 146	C	0.471	0.477	0.475	0.694	0.498
Gly 146	O	-0.415	-0.402	-0.407	-0.653	-0.397
EBS	C5	-0.154	-0.109	-0.119	-0.038	0.217
EBS	C6	0.257	-0.194	-0.238	0.058	-0.017
EBS	C4	-0.224	-0.518	-0.520	0.083	-0.742
EBS	C7	0.353	0.737	0.806	-0.128	0.981



EBS	C(-CO)	-0.797	0.068	0.028	-0.196	0.287
EBS	C(-Se)	0.586	-0.248	-0.287	0.153	-0.761
EBS	C3	0.004	-0.354	-0.346	0.929	1.042
EBS	N2	-0.265	-0.002	0.011	-0.378	0.162
EBS	O	-0.460	-0.530	-0.546	-0.860	-0.506
EBS	Se1	0.343	0.614	0.714	0.666	-0.258
EBS	C1'	-1.152	-0.077	-0.129	0.332	-0.748
EBS	C6'	1.461	0.307	0.329	-0.048	-0.230
EBS	C5'	-0.946	-0.416	-0.434	-0.004	0.279
EBS	C4'	0.420	0.176	0.160	-0.053	0.211
EBS	C3'	-0.525	0.052	0.052	0.011	-0.208
EBS	C2'	0.750	0.344	0.326	-0.033	0.550
W1	O	0.071	0.085	0.082	-0.846	-0.053
W2	O	0.085	0.053	0.103	0.058	0.070

**Table S8.** NBO charge values of heavy atoms in |e| for every stationary point along the potential energy surface of inhibition promoted by EBS on M<sup>pro</sup> H1E. Heavy atoms of the inhibitor, of the two catalytic waters, of the catalytic dyad and of the residues immediately close to the dyad are reported.

Residue	Atom	EI	TS	EI'	TS1	E-I
His 41	N	-0.609	-0.576	-0.583	-0.585	-0.584
His 41	CA	0.003	-0.029	-0.005	-0.005	-0.014
His 41	CB	-0.310	-0.327	-0.332	-0.333	-0.327
His 41	CG	0.128	0.182	0.080	0.302	0.266
His 41	ND1	-0.570	-0.580	-0.430	-0.697	-0.720
His 41	CE1	0.228	0.342	0.232	0.369	0.365
His 41	NE2	-0.704	-0.666	-0.667	-0.612	-0.578
His 41	CD2	0.087	0.095	0.111	0.016	0.002
His 41	C	0.377	0.388	0.352	0.347	0.381
His 41	O	-0.421	-0.425	-0.418	-0.403	-0.418
Asn 142	N	-0.538	-0.539	-0.536	-0.527	-0.537
Asn 142	CA	-0.040	-0.040	-0.050	-0.060	-0.038
Asn 142	CB	-0.317	-0.317	-0.312	-0.323	-0.328
Asn 142	CG	0.547	0.551	0.539	0.559	0.575
Asn 142	OD1	-0.508	-0.511	-0.498	-0.508	-0.536
Asn 142	ND2	-0.719	-0.718	-0.716	-0.715	-0.681
Asn 142	C	0.521	0.497	0.513	0.524	0.507
Asn 142	O	-0.503	-0.486	-0.491	-0.494	-0.493
Gly 143	N	-0.655	-0.656	-0.666	-0.648	-0.645
Gly 143	CA	-0.126	-0.176	-0.182	-0.166	-0.171
Gly 143	C	0.546	0.553	0.509	0.535	0.540
Gly 143	O	-0.497	-0.511	-0.456	-0.518	-0.521
Ser 144	N	-0.604	-0.608	-0.640	-0.616	-0.605
Ser 144	CA	-0.092	-0.052	-0.080	-0.066	-0.058
Ser 144	CB	-0.031	-0.042	-0.024	-0.034	-0.041
Ser 144	OG	-0.619	-0.624	-0.635	-0.616	-0.620
Ser 144	C	0.602	0.579	0.603	0.548	0.570
Ser 144	O	-0.567	-0.550	-0.495	-0.495	-0.539
Cys 145	N	-0.645	-0.622	-0.616	-0.641	-0.634
Cys 145	CA	-0.034	-0.030	-0.019	-0.043	-0.016
Cys 145	CB	-0.494	-0.473	-0.501	-0.486	-0.517
Cys 145	SG	0.009	-0.169	-0.430	-0.115	-0.110
Cys 145	C	0.546	0.555	0.551	0.551	0.557
Cys 145	O	-0.472	-0.487	-0.466	-0.484	-0.473
Gly 146	N	-0.654	-0.682	-0.633	-0.641	-0.683
Gly 146	CA	-0.178	-0.136	-0.182	-0.159	-0.140
Gly 146	C	0.360	0.346	0.368	0.352	0.351
Gly 146	O	-0.409	-0.411	-0.409	-0.394	-0.402
EBS	C5	-0.145	-0.142	-0.136	-0.132	-0.152
EBS	C6	-0.104	-0.111	-0.115	-0.112	-0.097
EBS	C4	-0.109	-0.117	-0.114	-0.120	-0.090
EBS	C7	-0.169	-0.146	-0.137	-0.157	-0.149

EBS	C(-CO)	0.115	0.127	0.093	0.039	0.047
EBS	C(-Se)	-0.381	-0.384	-0.395	-0.337	-0.248
EBS	C3	0.540	0.540	0.498	0.494	0.490
EBS	N2	-0.944	-0.924	-0.952	-0.789	-0.826
EBS	O	-0.467	-0.470	-0.430	-0.501	-0.424
EBS	Se1	0.705	0.633	0.619	0.389	0.321
EBS	C1'	0.334	0.341	0.286	0.304	0.458
EBS	C6'	-0.203	-0.193	-0.098	-0.167	-0.166
EBS	C5'	-0.158	-0.146	-0.137	-0.145	-0.141
EBS	C4'	-0.120	-0.116	-0.101	-0.114	-0.117
EBS	C3'	-0.169	-0.167	-0.147	-0.155	-0.115
EBS	C2'	-0.125	-0.120	-0.092	-0.099	-0.261
W1	O	-0.796	-0.811	-0.779	-0.818	-0.825
W2	O	-0.752	-0.759	-0.768	-0.769	-0.725

**Table S9.** NBO charge values of heavy atoms in |e| for every stationary point along the potential energy surface of inhibition promoted by EBS-OH on M<sup>pro</sup> HID. Heavy atoms of the inhibitor, of the two catalytic waters, of the catalytic dyad and of the residues immediately close to the dyad are reported.

Residue	Atom	EI	TS	EI'	TS1	E-I
His 41	N	-0.046	0.053	0.053	-0.006	0.009
His 41	CA	0.360	-0.235	-0.007	-0.034	-0.009
His 41	CB	-0.796	-0.308	-0.942	-0.586	-0.141
His 41	CG	0.150	0.823	0.513	0.360	0.333
His 41	ND1	0.214	0.035	0.133	0.211	-0.044
His 41	CE1	0.129	0.146	0.374	0.046	-0.183
His 41	NE2	-0.212	-0.363	0.141	0.215	-0.217
His 41	CD2	0.209	-0.255	0.771	0.615	0.416
His 41	C	0.414	0.557	0.226	0.433	0.224
His 41	O	-0.343	-0.401	-0.368	-0.403	-0.369
Asn 142	N	0.114	0.090	0.109	-0.013	0.131
Asn 142	CA	-0.185	0.076	-0.237	0.191	-0.237
Asn 142	CB	-0.624	-0.137	-0.504	-0.812	-0.558
Asn 142	CG	1.218	0.628	1.019	1.304	1.119
Asn 142	OD1	-0.533	-0.514	-0.474	-0.456	-0.511
Asn 142	ND2	0.060	0.088	0.044	0.120	0.062
Asn 142	C	0.180	0.029	0.079	-0.036	0.433
Asn 142	O	-0.431	-0.441	-0.448	-0.516	-0.422
Gly 143	N	0.181	0.321	0.340	0.199	0.207
Gly 143	CA	-0.461	-0.386	-0.139	-0.193	-0.181
Gly 143	C	0.985	0.581	0.594	0.530	0.393
Gly 143	O	-0.324	-0.499	-0.477	-0.400	-0.468
Ser 144	N	0.380	0.363	0.282	0.290	0.285
Ser 144	CA	-0.229	0.276	0.085	-0.183	0.356
Ser 144	CB	-0.190	-0.098	-0.189	0.062	-0.120
Ser 144	OG	-0.058	-0.042	-0.069	-0.099	-0.065
Ser 144	C	0.335	0.278	0.221	0.586	0.318
Ser 144	O	-0.470	-0.517	-0.521	-0.573	-0.517
Cys 145	N	0.403	0.553	0.270	0.236	0.447
Cys 145	CA	0.689	0.397	1.013	0.063	0.857
Cys 145	CB	-0.509	-0.542	-0.932	-0.665	-1.122
Cys 145	SG	0.321	0.421	-0.552	0.305	0.847
Cys 145	C	-0.496	-0.593	-0.174	-0.037	-0.759
Cys 145	O	-0.411	-0.451	-0.432	-0.490	-0.463
Gly 146	N	0.123	0.134	0.106	0.080	0.166
Gly 146	CA	0.141	0.089	0.104	0.201	0.072
Gly 146	C	0.500	0.550	0.531	0.440	0.572
Gly 146	O	-0.414	-0.413	-0.431	-0.395	-0.419
EBS	C5	-0.022	-0.201	-0.077	-0.228	-0.231
EBS	C6	0.827	0.092	0.025	-0.225	-0.062
EBS	C4	0.119	0.167	0.325	0.364	-0.618
EBS	C7	0.025	-0.052	0.236	0.621	1.537

EBS	C(-CO)	-0.374	0.167	-0.496	-0.029	0.041
EBS	C(-Se)	0.727	0.179	-0.702	-0.791	-0.141
EBS	C3	0.521	0.130	0.210	0.812	0.084
EBS	N2	-0.129	0.540	-0.415	-0.740	0.341
EBS	O	-0.389	-0.388	-0.071	-0.458	-0.307
EBS	Se1	0.036	-0.393	-0.472	-0.178	-0.628
EBS	C1'	0.005	-0.999	-0.399	0.112	-0.578
EBS	C6'	-0.116	0.692	0.252	0.050	0.323
EBS	C5'	-0.142	-0.302	0.028	-0.077	-0.348
EBS	C4'	0.148	0.188	-0.118	0.181	0.280
EBS	C3'	0.137	-0.110	0.793	0.295	-0.038
EBS	C2'	-0.033	0.363	-0.065	-0.412	0.012
EBS	O (2')	-0.094	-0.050	-0.497	0.039	0.000
W1	O	-0.058	0.048	0.371	0.039	0.037
W2	O	-0.099	-0.366	0.622	0.037	-0.147

**Table S10.** Energy values corrected including free energy (kcal/mol) in gas phase for the three used basis sets for the inhibition phase

EBS	6-31+G	6-311+G(2d,p)	AUG-cc-pVTZ
EI	0.0	0.0	0.0
TS1	11.5	10.6	11.1
<hr/>			
EBS-OH			
EI	0.0	0.0	0.0
EI'	-1.0	-1.2	-2.5
TS1	11.1	8.2	8.7

## References

- 1 [https://www.rcsb.org/pdb?id=pdb\\_00006w63](https://www.rcsb.org/pdb?id=pdb_00006w63), 2020, .
- 2 Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao and H. Yang, *Nature*, 2020, **582**, 289–293.
- 3 L. Zhang, D. Lin, X. Sun, U. Curth, C. Drostén, L. Sauerhering, S. Becker, K. Rox and R. Hilgenfeld, 2020, **368**, 409–412.
- 4 R. Anandakrishnan, B. Aguilar and A. V. Onufriev, *Nucleic Acids Res*, 2012, **40**, W537-541.
- 5 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.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017; University of California, San Francisco, 2017., .
- 6 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, ff14SB, <https://pubs.acs.org/doi/full/10.1021/acs.jctc.5b00255>, (accessed 13 January 2021).
- 7 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *The Journal of Chemical Physics*, 1998, **81**, 3684.
- 8 T. Darden, D. York and L. Pedersen, *The Journal of Chemical Physics*, 1998, **98**, 10089.
- 9 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *Journal of Computational Chemistry*, **30**, 7.
- 10 J. Gasteiger and M. Marsili, *Tetrahedron*, 1980, **36**, 3219–3228.
- 11 G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *JOURNAL OF COMPUTATIONAL CHEMISTRY*, **19**, 24.
- 12 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
- 13 N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proc Natl Acad Sci U S A*, 2001, **98**, 10037–10041.
- 14 E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor and J. Damborsky, *PLoS Comput Biol*, 2012, **8**, e1002708.
- 15 J. Weiser, P. S. Shenkin and W. C. Still, *Journal of Computational Chemistry*, 1999, **20**, 217–230.

## *Supporting Informations*

# **New Insights on the redox Properties of Viperin from Advanced Numerical Simulations**

Angela Parise<sup>1,2</sup> Nino Russo<sup>2</sup>, Tiziana Marino<sup>2</sup>, Aurélien de la Lande<sup>1,\*</sup>

1 Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France

2 Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy

\* Correspondence: aurelien.de-la-lande@universite-paris-saclay.fr;

### **Table of Contents**

Scheme S1. Catalytic mechanism of viperin

Figure S2. Root Mean Square Fluctuations (RMSF)

Figure S3. Ribbons structure colored by RMSF value

Figure S4. Ribbons CTP-viperin representation

Figure S5. Evolution of the secondary structure (DSSP)

Figure S6. Electrostatic potential surface (APBS)

Figure S6. Models of the QM-region used in QM/MM calculation

Figure S7. Cytidine triphosphate atom name

Figure S8. RMSD of two redox states MD

Table S1. Calculated pKa for ionizable residues

Table S2 Residues contact

Table S3. Hydrogen bonds

Table S4. RMSD-based clustering of viperin MD

Table S5. RMSD-based clustering of CTP-viperin MD

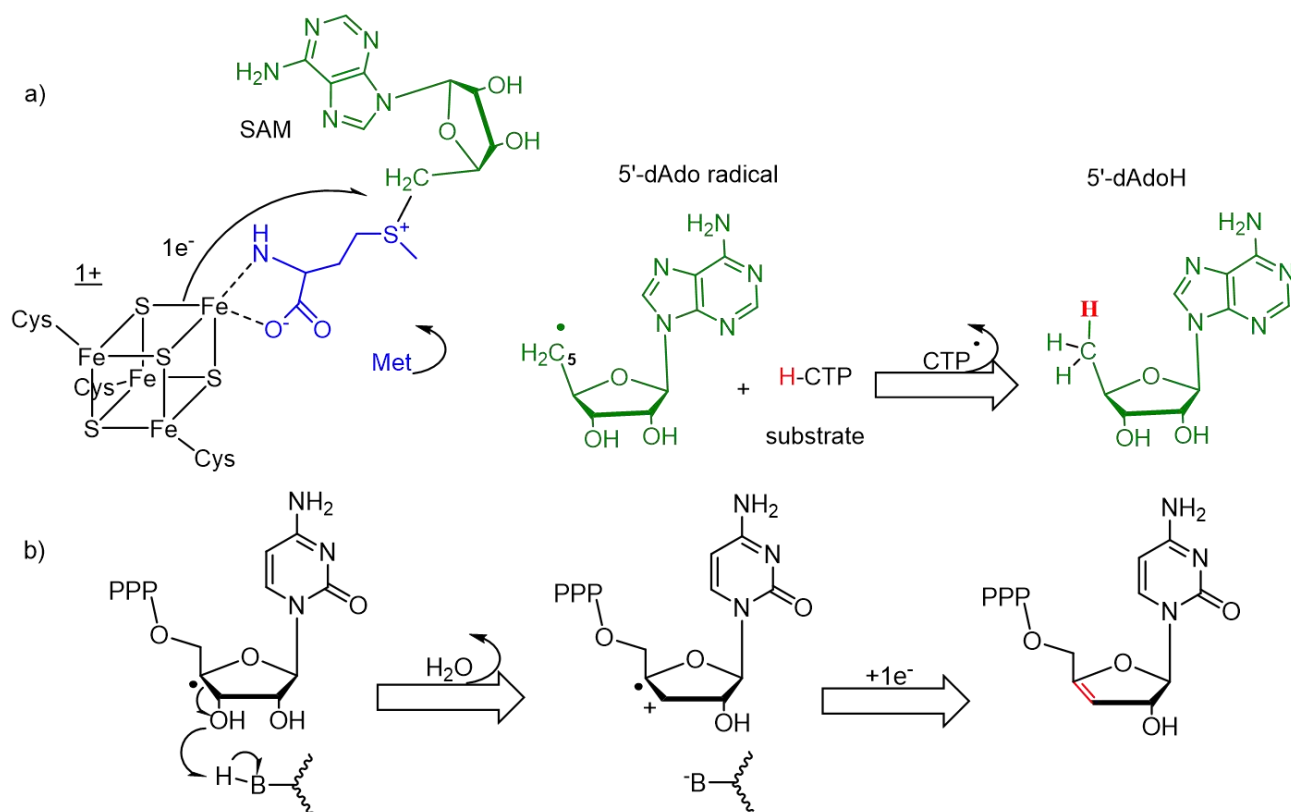
Table S6 Spin multiplicity ( $\Delta E$  in eV)

Table S7 Possible Spin configuration

Table S8. Hirshfeld population analysis

References.





Scheme S1. Catalytic mechanism of viperin. . a) Single electron transfer from Cuban to SAM with breaking of the S-C bond and formulation of a methionine, 5'-deoxyadenosyl radical abstract an hydrogen atom at the 4' position of CTP, releasing radical CTP. b) CTP radical general base-assisted loss of the 3' hydroxyl group leads to a carbocation/radical intermediate that is reduced by one electron to yield the ddh-CTP product.[1]

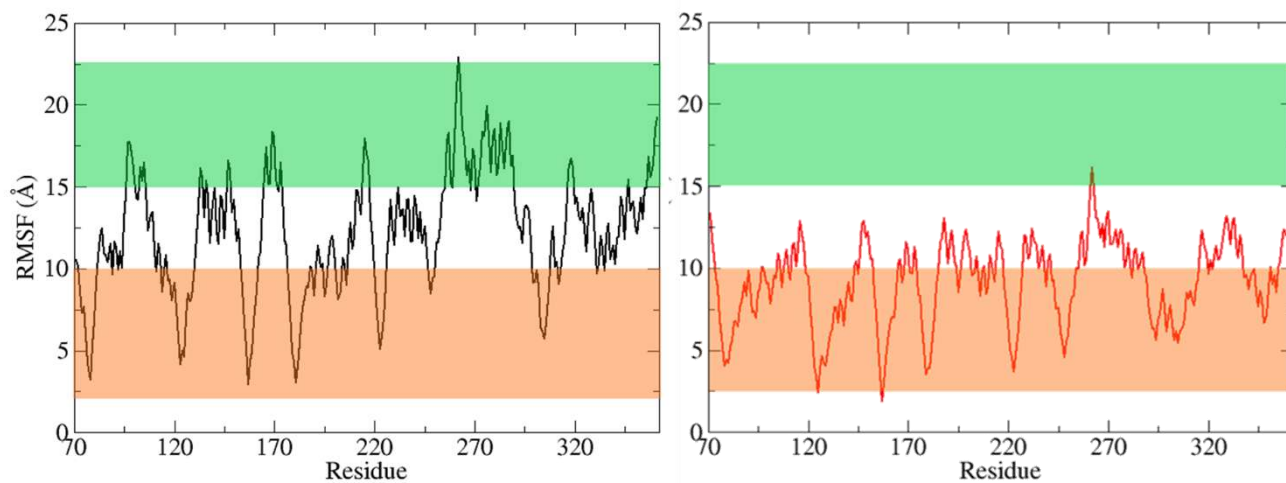


Figure S2. Root Mean Square Fluctuations (RMSF) of the structures ( $\text{\AA}$ ) with respect to residues of viperin (in black) and CTP-viperin (in red). We have arbitrarily divided the RMSF plot into 3 different colored sections green, white and orange for the RMSF value range corresponding to 2,5-10  $\text{\AA}$ , 10-15  $\text{\AA}$  and 15-22,5  $\text{\AA}$  respectively.

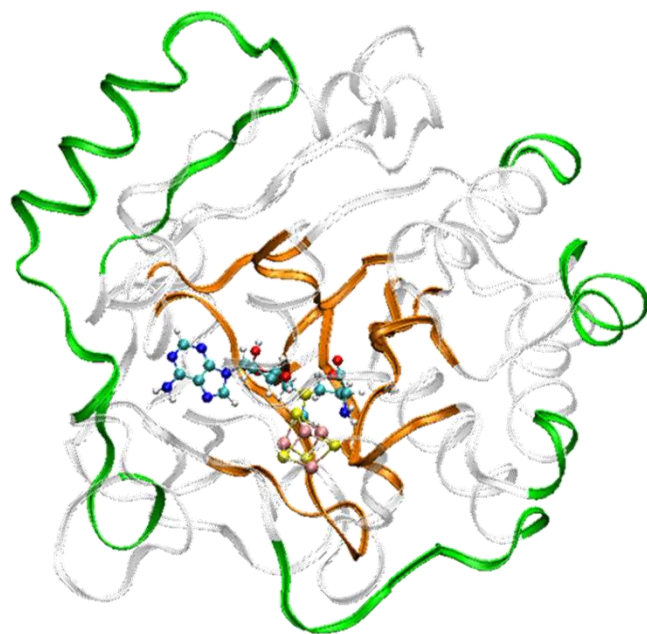


Figure S3. Ribbons structure representation of viperin colored by RMSF value as describes in Figure S2.

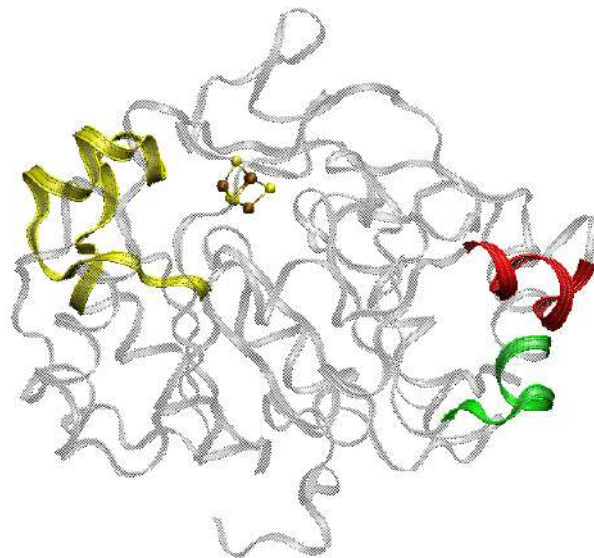


Figure S4. Ribbons CTP-viperin representation. The colored portions of the protein indicate the parts that have a  $\Delta\text{RMSF}$  value greater than zero, where  $\Delta\text{RMSF}$  is derived from the difference in RMSF of CTP-viperin minus RMSF of viperin.

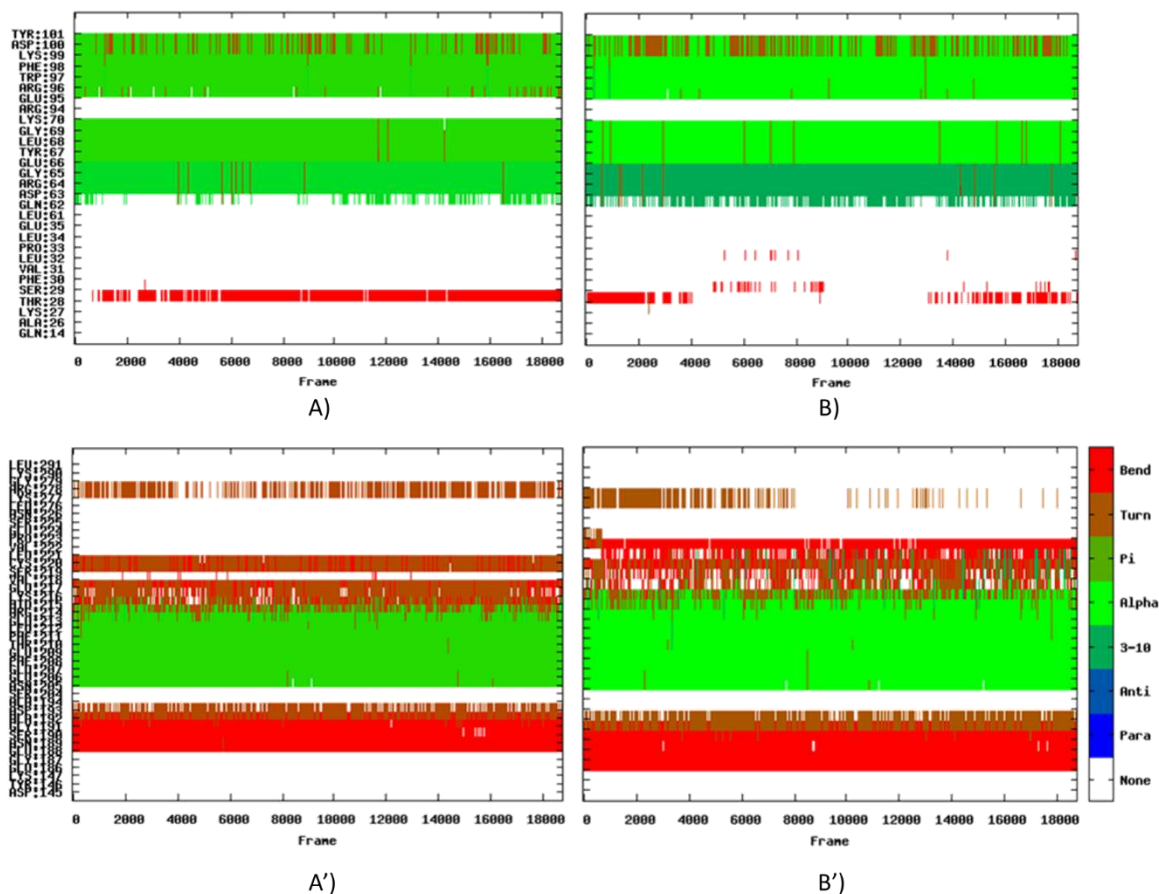


Figure S5. Time evolution of the secondary structural elements of amino acid residues respect to trajectory frames: on top, secondary structure of amino acid segment Asp 83 - Tyr 170 (index 14 to 101) is reported and down, secondary structure of segment Asp 214 – Leu 360 (index 145 to 291). A) and A') refers to apo-viperin; B) and B') refers to CTP-viperin complex.

The so-called “Dictionary of Secondary Structure of Proteins” (DSSP) by Kabsch and Sander makes its sheet and helix assignments solely on the basis of backbone-backbone hydrogen bonds.[2]

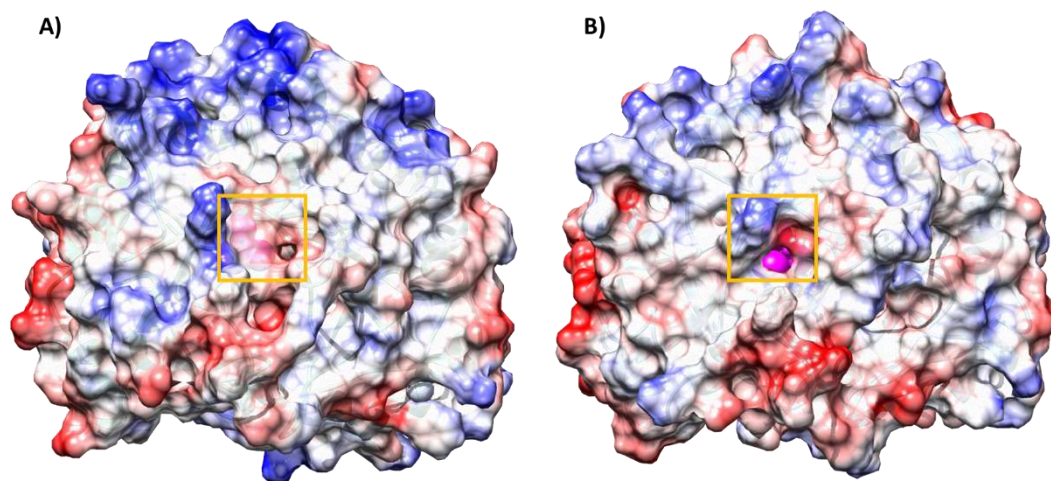


Figure S6. Electrostatic potential surface (APBS) in: A) viperin and B) CTP-viperin, in both cases the most representative frames of MD simulation is used. Yellow square indicates the portion containing the Fe-S cluster.

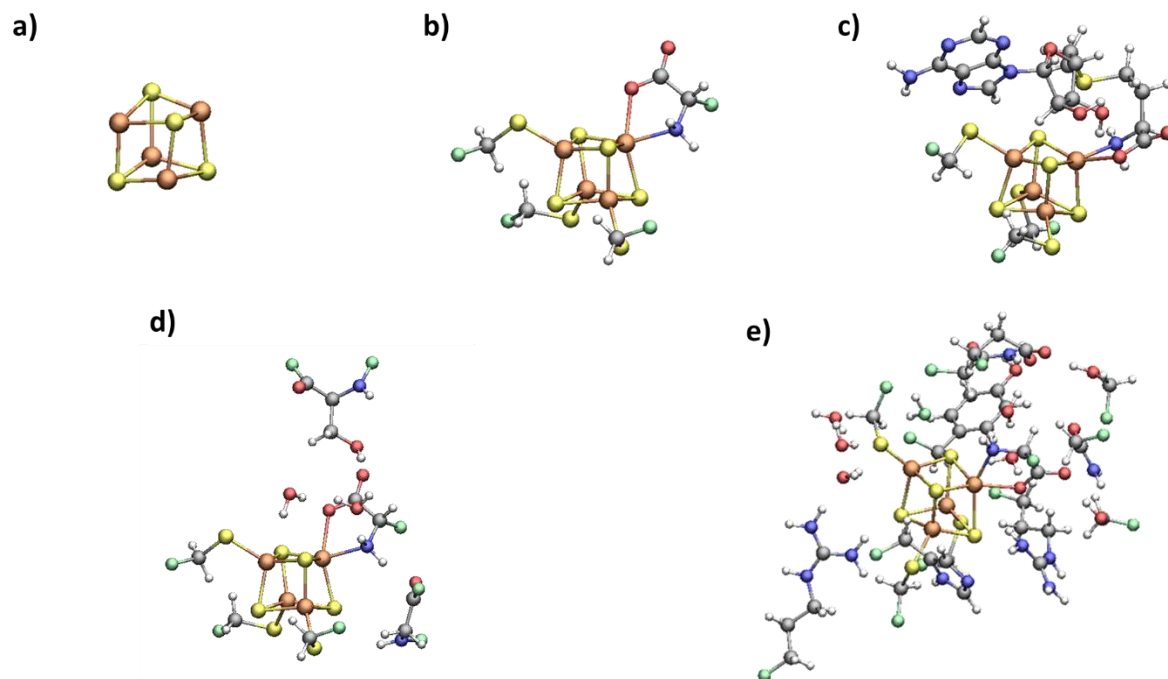


Figure S6. Models of the QM-region that were used in this work in QM/MM calculations: a) 8 atoms, b) 32 atoms (including 4 link-atoms (HL) ), c) 72 atoms (including 3 HL), d) 60 atoms (including 8 HL) , e) 156 atoms (including 14 HL).

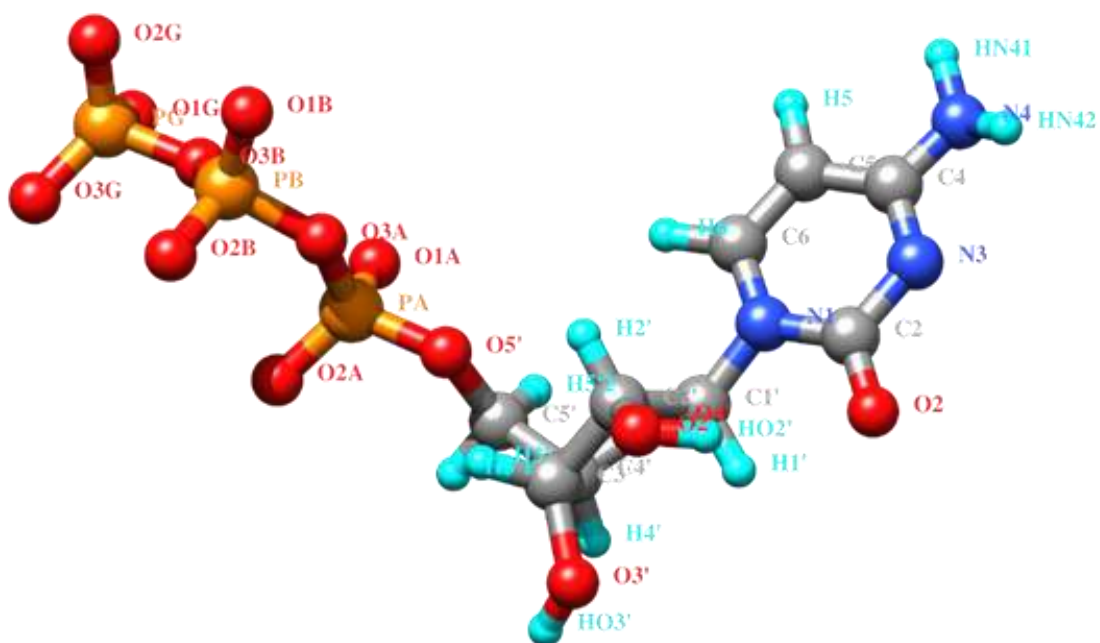


Figure S7. Cytidine triphosphate substrate structure and related color-defined atom names.



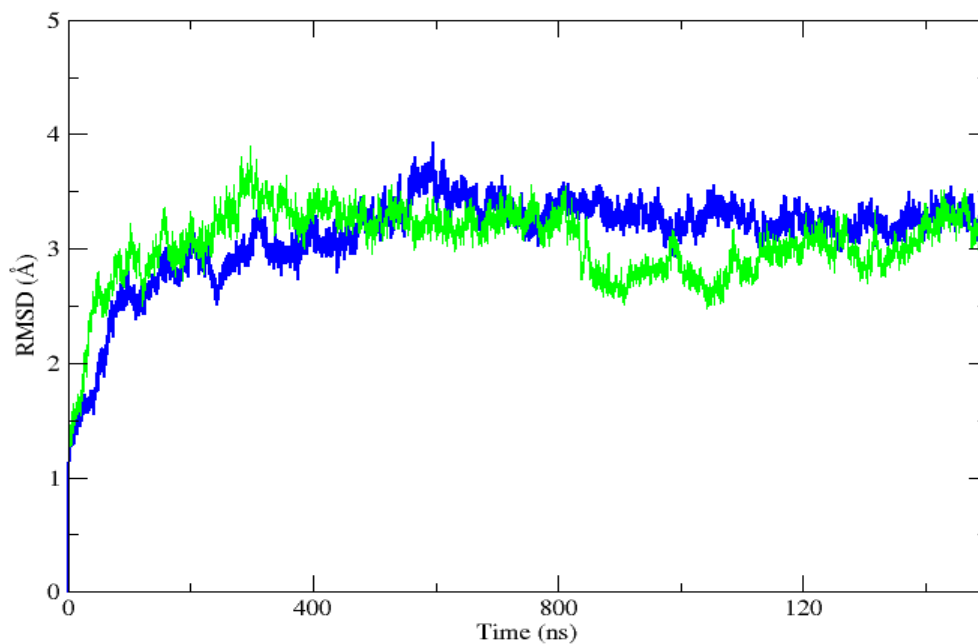


Figure S8. The root-mean-square deviation of protein along 150 ns of MD simulation for viperin with 4Fe-4S cluster in its reduced state (in blue) and ii oxidized state (in green).

Table S1. Calculated pKa for ionizable residues of viperin obtained by using H++ software.

<b>Residue</b>	<b>pKint</b>	<b>Residue</b>	<b>pKint</b>	<b>Residue</b>	<b>pKint</b>	<b>Residue</b>	<b>pKint</b>
TYR-78	19.494	LYS-220	5.503	TYR-78	19.359	LYS-220	5.700
HIS-79	2.869	ARG-227	10.754	HIS-79	2.590	ARG-227	10.955
ARG-82	9.530	ASP-231	4.190	ARG-82	9.151	ASP-231	4.225
CYS-84	5.927	GLU-232	2.938	CYS-84	5.612	GLU-232	3.103
TYR-86	10.670	ASP-233	2.114	TYR-86	13.142	ASP-233	2.280
LYS-87	9.750	GLU-236	3.994	LYS-87	10.013	GLU-236	5.029
CYS-88	9.948	HIS-237	4.898	CYS-88	10.306	HIS-237	4.798
CYS-91	8.163	LYS-239	9.037	CYS-91	8.078	LYS-239	9.444
HIS-93	0.885	ARG-245	13.287	HIS-93	0.657	ARG-245	13.448
LYS-96	9.663	LYS-247	7.458	LYS-96	9.676	LYS-247	7.800
GLU-104	3.979	CYS-251	12.995	GLU-104	5.257	CYS-251	12.897
GLU-105	5.581	GLU-255	4.820	GLU-105	5.305	GLU-255	4.616
LYS-107	8.500	GLU-257	7.060	LYS-107	8.600	GLU-257	6.684
ARG-108	11.949	ASP-262	3.752	ARG-108	11.927	ASP-262	3.683
LYS-114	10.160	ARG-265	10.701	LYS-114	9.885	ARG-265	10.599
GLU-119	5.684	GLU-268	4.348	GLU-119	5.581	GLU-268	4.334
LYS-120	8.741	ARG-269	10.962	LYS-120	8.641	ARG-269	10.849
GLU-127	0.454	GLU-275	3.720	GLU-127	-0.018	GLU-275	3.639
ASP-132	3.657	GLU-276	3.787	ASP-132	3.584	GLU-276	4.643
ARG-133	11.461	GLU-278	4.822	ARG-133	11.226	GLU-278	4.791
GLU-135	3.888	GLU-282	5.493	GLU-135	4.082	GLU-282	5.349
TYR-136	20.997	ARG-283	11.213	TYR-136	22.255	ARG-283	11.323
LYS-139	9.259	HID-284	-2.599	LYS-139	9.319	HID-284	-2.421
ARG-142	12.002	LYS-285	10.185	ARG-142	11.730	LYS-285	10.188
CYS-144	17.652	GLU-286	5.758	CYS-144	17.738	GLU-286	6.488
LYS-145	10.468	CYS-289	13.117	LYS-145	10.486	CYS-289	13.826
GLU-146	5.527	GLU-293	11.278	GLU-146	5.623	GLU-293	11.365
GLU-147	4.426	LYS-297	8.677	GLU-147	4.554	LYS-297	9.025
ARG-163	11.531	LYS-299	8.028	ARG-163	11.710	LYS-299	8.204
GLU-164	3.575	ASP-300	8.332	GLU-164	3.794	ASP-300	8.361
ARG-165	11.204	TYR-302	20.802	ARG-165	11.439	TYR-302	20.742
LYS-168	9.342	ASP-306	4.196	LYS-168	10.062	ASP-306	4.109
ASP-169	4.899	GLU-307	5.664	ASP-169	6.583	GLU-307	5.377
TYR-170	13.928	TYR-308	11.857	TYR-170	14.066	TYR-308	11.820
GLU-172	5.308	ARG-310	10.146	GLU-172	5.495	ARG-310	10.198
TYR-173	14.341	CYS-314	17.843	TYR-173	14.046	CYS-314	18.067
ASP-175	10.019	ARG-318	11.895	ASP-175	9.606	ARG-318	11.863
CYS-181	14.058	LYS-319	5.492	CYS-181	14.463	LYS-319	5.269
ASP-182	4.543	ASP-320	5.627	ASP-182	5.563	ASP-320	5.626
ASP-185	3.427	LYS-323	9.828	ASP-185	2.836	LYS-323	9.789
GLU-186	3.776	ASP-327	3.836	GLU-186	3.703	ASP-327	3.866
ARG-194	5.099	GLU-331	4.368	ARG-194	5.014	GLU-331	4.458
LYS-198	9.999	GLU-332	4.936	LYS-198	9.173	GLU-332	5.315

LYS-199	8.275	LYS-335	9.689	LYS-199	8.310	LYS-335	9.455
HIS-201	0.640	ASP-340	3.724	HIS-201	0.429	ASP-340	3.780
GLU-203	5.777	GLU-341	4.453	GLU-203	4.843	GLU-341	4.709
LYS-207	9.838	LYS-342	9.983	LYS-207	10.412	LYS-342	10.024
ARG-209	13.730	LYS-346	9.716	ARG-209	15.995	LYS-346	9.771
ARG-210	11.597	ARG-347	13.454	ARG-210	11.748	ARG-347	13.204
CYS-212	19.955	LYS-350	10.196	CYS-212	20.291	LYS-350	10.002
ARG-213	12.075	TYR-351	21.431	ARG-213	11.729	TYR-351	21.283
ASP-214	5.933	LYS-355	13.220	ASP-214	5.890	LYS-355	12.848
TYR-215	13.540	ASP-357	4.324	TYR-215	12.792	ASP-357	4.469
LYS-216	10.676	LYS-359	10.317	LYS-216	10.679	LYS-359	9.890

Table S2. For each atom of CTP (**CTP atoms** as in Figure S7) contacts are calculated by cpptraj modules. **Resid** refers to viperin amino acid residue, **Atype** is the atom type of respective resid, **N frames** is the number of frames in which CTP-atom and Resid have a distance below the cutoff value of 3.5 Å, **Freq** is the frequency with which the contact is maintained during the 150 ns trajectory (Frequencies of 100% along the trajectory correspond to Freq=1.0), **Avg** is the average value of the distance between the atoms of the substrate and the relative resid involved in the contact.

<b>CTP atoms</b>	<b>Resid</b>	<b>Atype</b>	<b>N frames</b>	<b>Freq</b>	<b>Avg</b>
O1A	247	NZ	18748	1.0	2.7
O1A	302	OH	18747	1.0	2.6
O2A	220	NZ	18486	1.0	2.8
O2B	120	NZ	17924	1.0	2.9
O2B	220	NZ	17420	0.9	2.9
O1G	247	NZ	16915	0.9	2.8
O3B	247	NZ	16478	0.9	3.2
O1B	120	NZ	16255	0.9	3.0
O3G	220	NZ	16163	0.9	2.8
O1G	347	NH1	15925	0.8	2.8
O1A	302	CZ	13928	0.7	3.4
O1A	302	CE1	13225	0.7	3.3
O2G	120	NZ	12436	0.7	3.0
PB	120	NZ	12263	0.7	3.4
O2'	79	NE2	12152	0.6	3.1
O3G	120	NZ	11365	0.6	3.1
O3B	220	NZ	10435	0.6	3.3
O1G	245	NH2	8117	0.4	3.2
O1A	247	CE	8069	0.4	3.4
O1A	247	CD	6869	0.4	3.4
O2B	122	ND2	6413	0.3	3.1
PG	120	NZ	4750	0.3	3.4
O2A	222	ND2	4214	0.2	2.9
O2A	220	CE	4202	0.2	3.4
O2	319	NZ	3964	0.2	2.9
N3	319	NZ	3805	0.2	3.0
O4'	298	CE	3778	0.2	3.3
PG	245	NH2	3389	0.2	3.4
O3G	245	NE	2419	0.1	2.9
C3'	77	ND2	1984	0.1	3.3
O3'	124	OG	1852	0.1	3.2
O3A	77	ND2	1834	0.1	3.1
O2'	77	OD1	1609	0.1	3.2
O3'	77	OD1	1607	0.1	2.9
N4	314	SG	1485	0.1	3.3
O3'	77	ND2	1404	0.1	3.3
O1G	347	NH2	1361	0.1	3.4

O3'	79	NE2	1307	0.1	3.0
O3'	77	CG	1200	0.1	3.2
O2'	77	CG	1187	0.1	3.3
C3'	77	CG	1021	0.1	3.4

Table S3. Hydrogen bonds over the course of a trajectory. Hydrogen bonds are determined using simple geometric criteria: the donor to acceptor heavy atom distance and angle in cpptraj module. **A-RES** and **A-atom** (as in Figure S7) refer to H-bond acceptor residue and atom, respectively; **D-RES** and **D-atom** refer to donor residue and atom, Freq is the frequency with which hydrogen bond is maintained during the 150 ns of CTP-viperin MD simulation.

<b>A-RES</b>	<b>A-atom</b>	<b>D-RES</b>	<b>D-atom</b>	<b>Freq</b>
CTP	O1A	Tyr302	OH	1.0
CTP	O1G	Arg347	NH1	0.8
CTP	O1G	Arg245	NH1	0.7
CTP	O3G	Arg245	NH2	0.7
CTP	O2G	Arg347	NH2	0.6
CTP	O1A	Lys247	NZ	0.4
CTP	O1G	Lys247	NZ	0.4
CTP	O2A	Lys220	NZ	0.3
CTP	O1A	Lys247	NZ	0.3
CTP	O2A	Lys220	NZ	0.3
CTP	O3G	Lys220	NZ	0.3
CTP	O3G	Lys220	NZ	0.3
CTP	O1A	Lys247	NZ	0.3
CTP	O1G	Lys247	NZ	0.3
CTP	O1G	Lys247	NZ	0.2
CTP	O2A	Lys220	NZ	0.2
CTP	O2B	Lys120	NZ	0.2
CTP	O3G	Lys220	NZ	0.2
CTP	O2B	Lys120	NZ	0.2
CTP	O2A	Asn222	ND2	0.2
CTP	O2B	Lys120	NZ	0.2
CTP	O2'	Asn77	ND2	0.1
CTP	O1G	Arg245	NH2	0.1
CTP	O2'	Ser124	OG	0.1
CTP	O2B	Asn122	ND2	0.1
CTP	O3G	Arg245	NE	0.1
CTP	O2G	Arg347	NH1	0.1
CTP	O2G	Lys120	NZ	0.1
CTP	O2G	Lys247	NZ	0.1
CTP	O2G	Lys120	NZ	0.1
CTP	HO2'	Ser124	OG	0.1
CTP	O1B	Arg347	NH2	0.1
CTP	O1B	Arg347	NH1	0.1

Table S4. Table of 10 structures obtained by RMSD-based clustering of viperin MD simulations. From first column: cluster name, number of frames included in the cluster, population respect to total processed frames in percentage, average distance between frames inside the cluster expressed in Å, standard deviation related to Avg d and the centroid frame.

Cluster	N° frames	Pop %	Avg d (Å)	$\sigma$	Centroid
0	24048	32.1	1.360	0.193	28044
1	13992	18.7	1.261	0.190	51636
2	12416	16.6	1.255	0.190	69168
3	7288	9.7	1.184	0.162	20720
4	6100	8.1	1.082	0.152	58564
5	4304	5.7	1.257	0.219	3596
6	2680	3.6	1.205	0.189	41444
7	2204	2.9	1.090	0.151	14736
8	1844	2.5	1.134	0.203	608
9	124	0.2	0.000	0.000	19732

Table S5. Table of 10 structures obtained by RMSD-based clustering of CTP-viperin MD simulations. From first column: cluster name, number of frames included in the cluster, population respect to total processed frames in percentage, average distance between frames inside the cluster expressed in Å, standard deviation related to Avg d and the centroid frame.

Cluster	N° frames	Pop %	Avg d (Å)	$\sigma$	Centroid
0	17655	23,5	1,010	0,113	69134
1	13761	18,3	1,020	0,117	27365
2	12948	17,3	1,063	0,129	54788
3	9132	12,2	1,020	0,115	42175
4	8452	11,3	1,029	0,143	4774
5	6897	9,2	1,029	0,130	11501
6	2314	3,1	0,975	0,119	33372
7	1699	2,3	0,967	0,129	17358
8	1390	1,9	0,948	0,125	16182
9	752	1,0	0,858	0,108	34832



Table S6  $\Delta E$  in eV (respect to the lower in Energy) with QM region model of 72 atoms testing different spin multiplicity ( $2S+1$ ) by different exchange and correlation functionals

( $2S+1$ )	PBE	PBE0	B3LYP
2	0.00	0.00	0.00
4	0.43	2.18	1.66
6	0.31	1.70	1.21

Table S7 Possible spin configuration expressed in terms of Spin charge.

Configuration	Fe <sub>1</sub>	Fe <sub>2</sub>	Fe <sub>3</sub>	Fe <sub>4</sub>
	d <sup>6</sup>	d <sup>6</sup>	d <sup>5</sup>	d <sup>6</sup>
$\alpha\beta\alpha\beta$	4	-4	5	-4
$\alpha\alpha\beta\beta$	4	4	-5	-4
$\beta\alpha\alpha\beta$	-4	4	5	-4
$\alpha\beta\beta\alpha$	4	-4	-5	4
$\beta\beta\alpha\alpha$	-4	-4	5	4
$\beta\alpha\beta\alpha$	-4	4	-5	4

Table S8. Hirshfeld population analysis of Kohn-Sham Density.

PDB index	atom	Atomic charge	Spin charge
243	C	-0,128	0,013
244	H	0,026	0,004
245	H	0,026	0,002
246	S	-0,378	0,124
310	C	-0,118	-0,016
311	H	0,029	-0,003
312	H	0,030	-0,003
313	S	-0,350	-0,195
347	C	-0,120	-0,005
348	H	0,032	0,001
349	H	0,028	0,001
350	S	-0,395	0,048
4720	<b>Fe</b>	0,036	3,174
4721	<b>S</b>	-0,365	-0,098
4722	<b>Fe</b>	-0,038	1,497
4723	<b>S</b>	-0,249	-0,011

4724	Fe	0,014	0,029
4725	S	-0,346	-0,173
4726	Fe	0,049	-3,337
4727	S	-0,360	-0,047
4728	N	-0,121	-0,009
4729	C	-0,020	0,001
4733	C	0,120	0,001
4734	O	-0,452	0,002
4735	O	-0,354	0,011
4755	H	0,152	-0,003
4756	H	0,157	-0,001
4761	H	0,052	0,000
	<u>HL</u>	0,012	0,003
	<u>HL</u>	0,013	-0,003
	<u>HL</u>	0,018	-0,004
	<u>HL</u>	0,000	0,000

## References

1. Gizzi, A.S.; Grove, T.L.; Arnold, J.J.; Jose, J.; Jangra, R.K.; Garforth, S.J.; Du, Q.; Cahill, S.M.; Dulyaninova, N.G.; Love, J.D.; et al. A Naturally Occurring Antiviral Ribonucleotide Encoded by the Human Genome. *Nature* **2018**, *558*, 610–614, doi:10.1038/s41586-018-0238-4.
2. Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22*, 2577–2637, doi:10.1002/bip.360221211.

## *Supporting Information*

**The catalytic product of the natural antiviral viperin, ddhCTP as possible inhibitor of SARS-CoV-2 RNA-polymerase-RNA-dependent: a computational insight**

A. Parise, G. Ciardullo, M. Prejanò, A. de la Lande, T. Marino

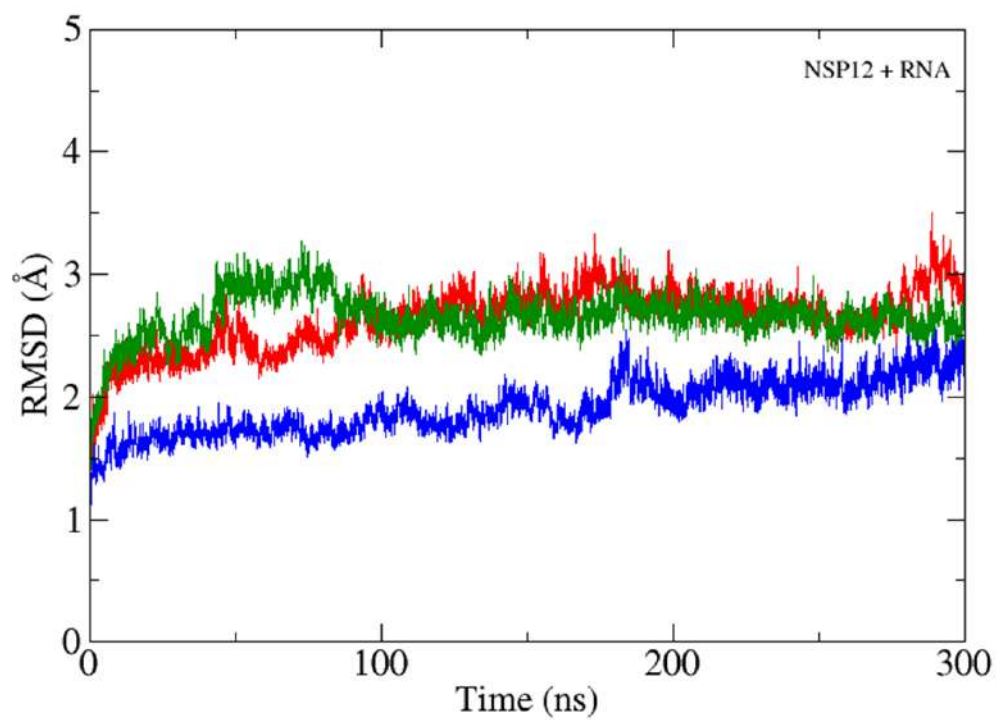


Figure S1. Root Mean Square Deviation (RMSD) of the protein part and RNA heavy atoms in Å with respect to time of systems: RdRp-RNA in blue, RdRp-RNA-ddhCTP in green, RdRp-RNA-CTP in red.

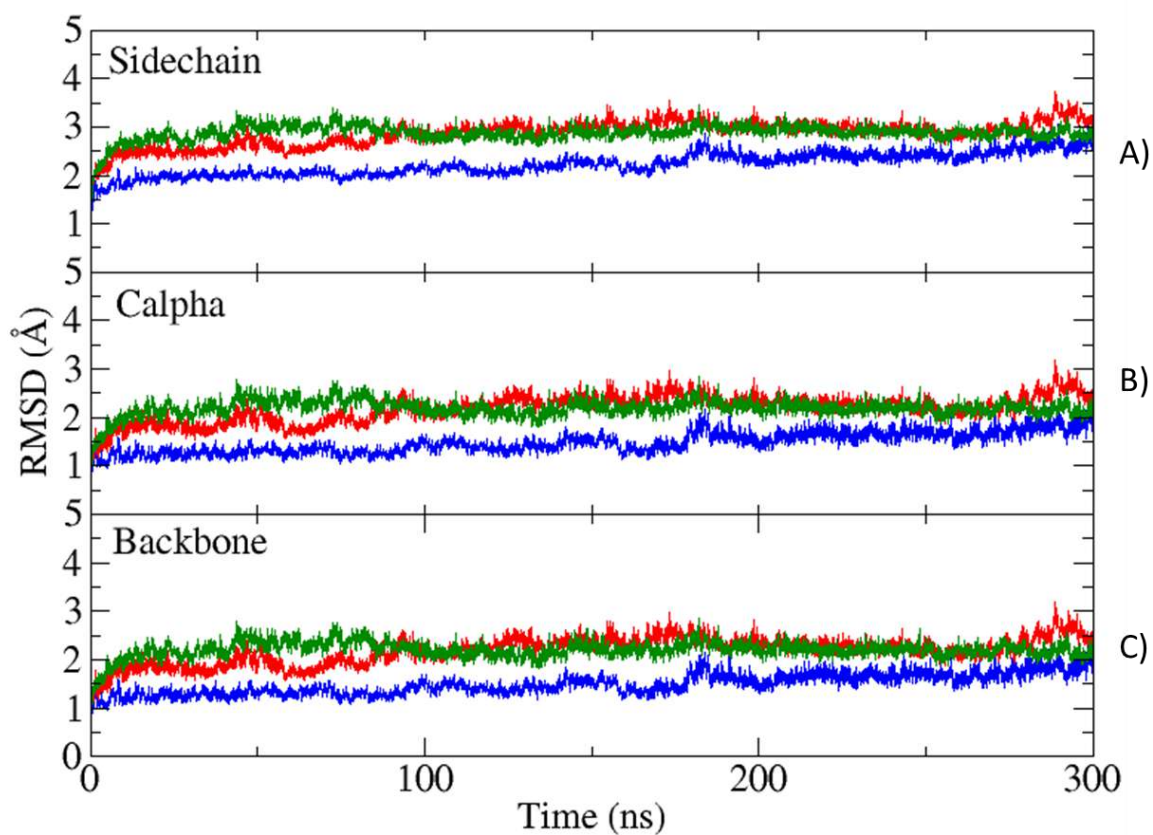


Figure S2. Root Mean Square Deviation (RMSD) of the protein part in Å with respect to time of systems: RdRp-RNA in blue, RdRp-RNA-ddhCTP in green, RdRp-RNA-CTP in red. Panels A) B) and C) refer to RMSD calculated for sidechain, alpha-carbon and backbone, respectively.

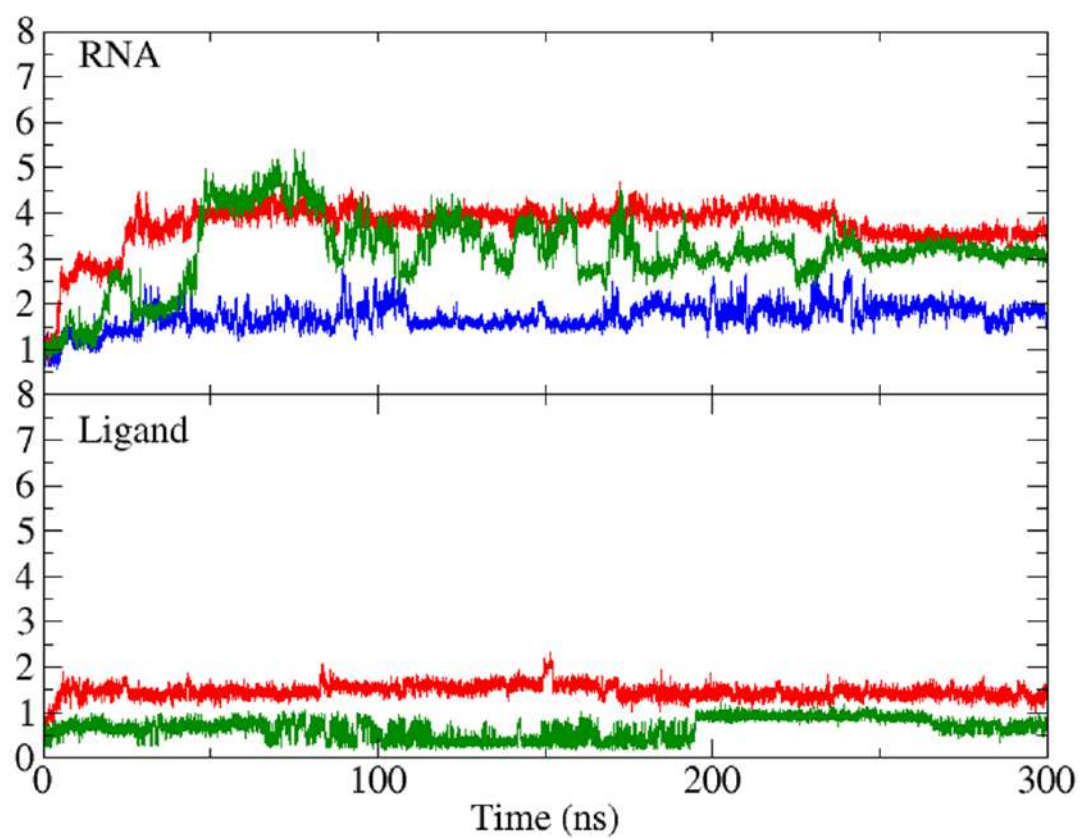


Figure S3. Root Mean Square Deviation (RMSD) of RNA in Å with respect to time of systems: RdRp-RNA in blue, RdRp-RNA-ddhCTP in green, RdRp-RNA-CTP in red, on top. RMSD trends of ddhCTP (in green) and CTP (in red) are reported down.

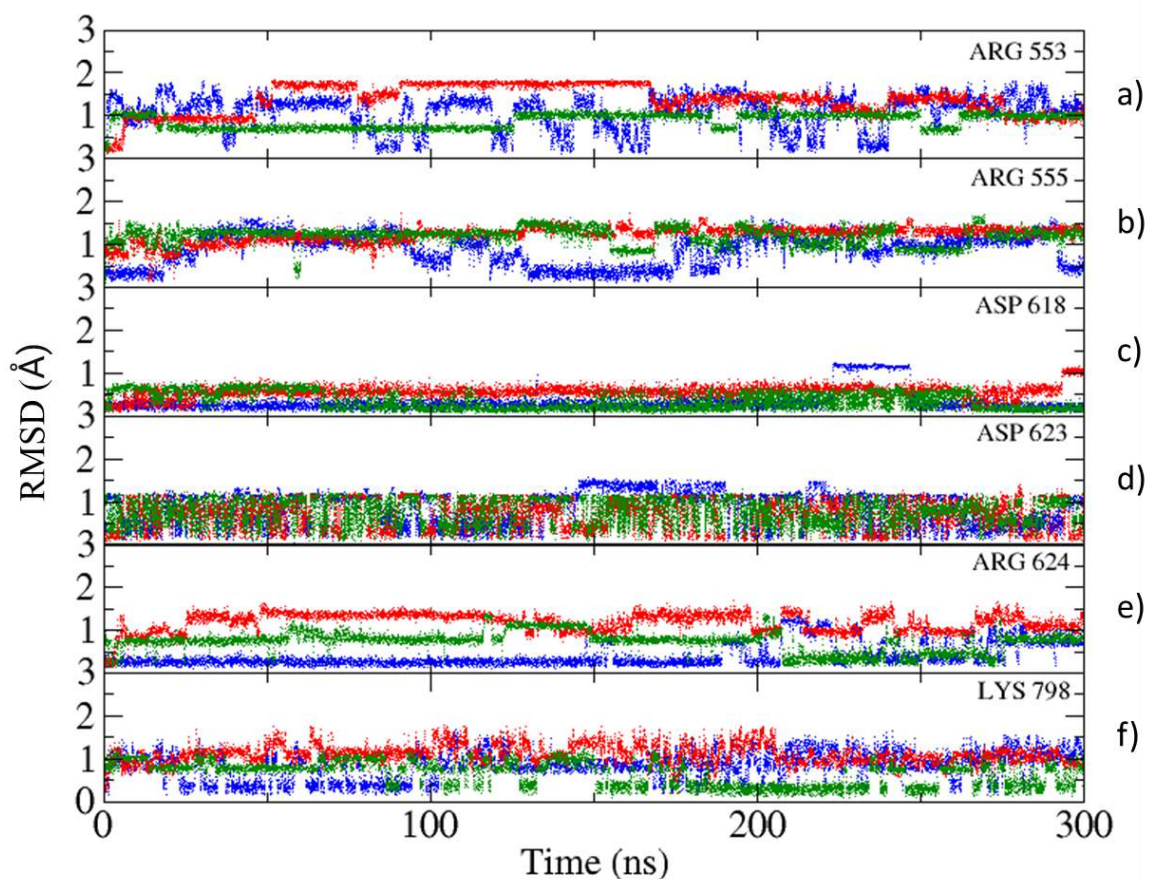


Figure S4. Root Mean Square Deviation (RMSD) of residues present in the nucleotide triphosphate (NTP) channel part in Å with respect to time of systems: RdRp-RNA in blue, RdRp-RNA-ddhCTP in green, RdRp-RNA-CTP in red. Each panel refers to a specific residue, as follows: a) Arg553, b) Arg555, c) Asp618, d) Asp623, e) Arg624 and f) Lys798.



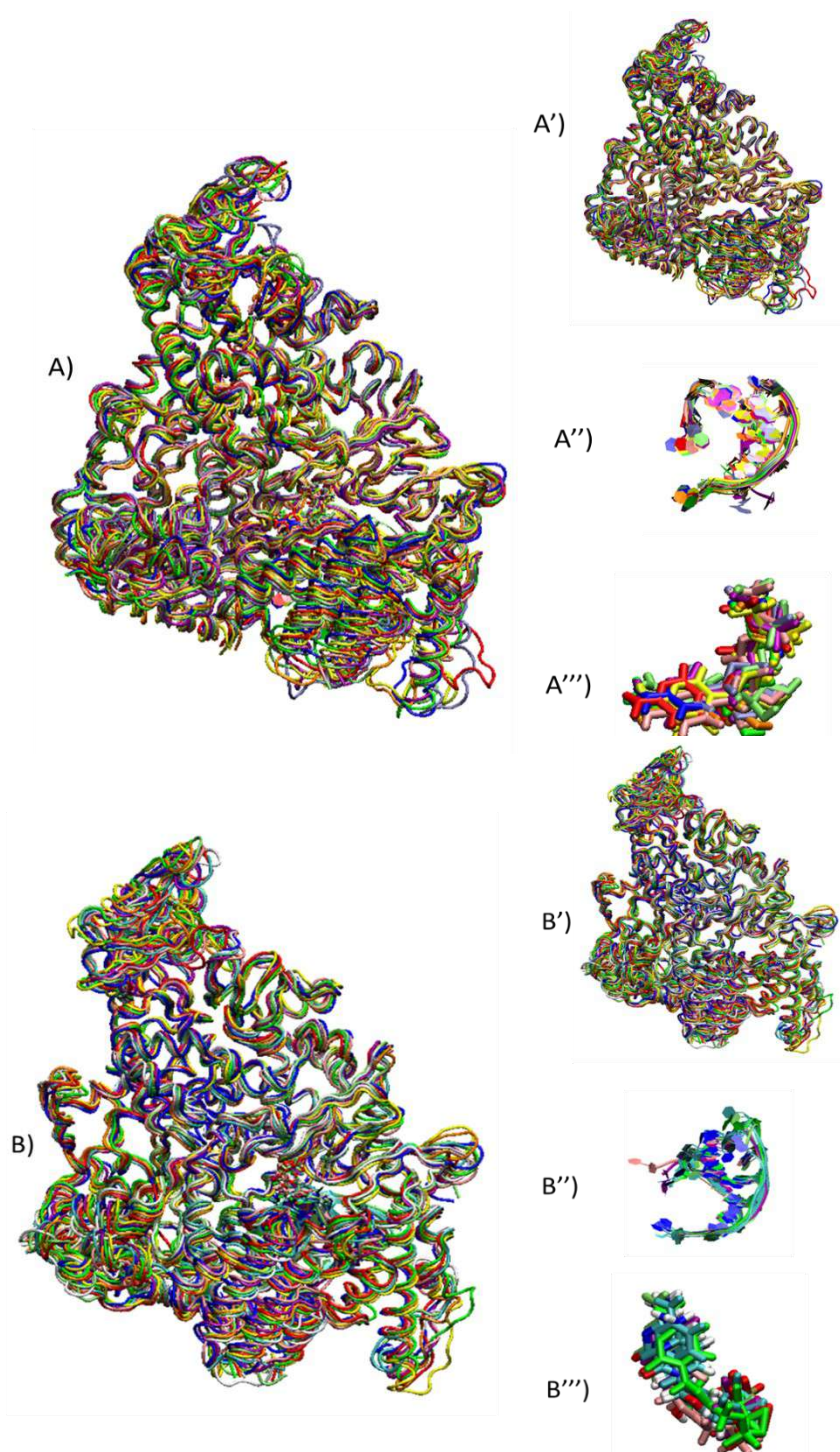


Figure S5. Superposition of 10 structures obtained by RMSD-based clustering of RdRp-RNA-CTP (on top) and RdRp-RNA-ddhCTP (down), each cluster has a different color. A) includes nsp7-nsp8-ns12, RNA primer:template and CTP, better represented in A') nsp12, A'') RNA part and A''') CTP. B) includes nsp7-nsp8-ns12, RNA primer:template and ddhCTP, better represented in B') nsp12, B'') RNA part and B''') ddhCTP.

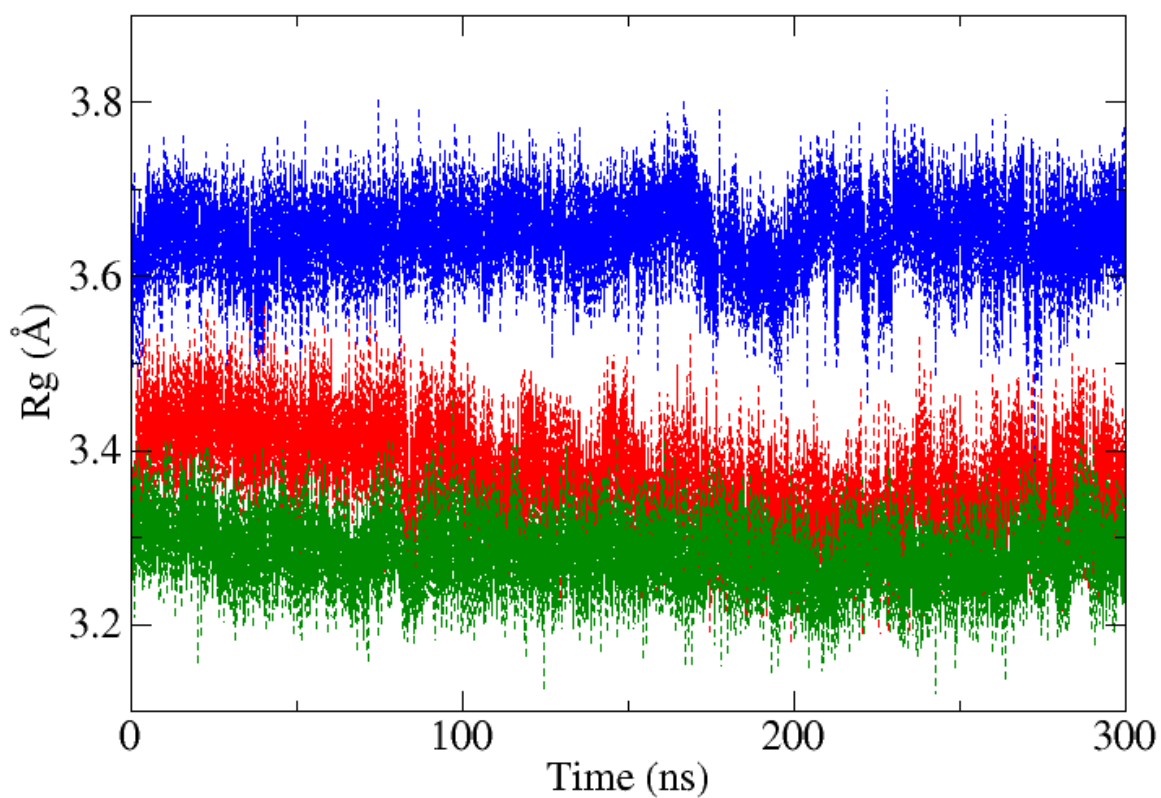


Figure S6. Radius of gyration (Rg) calculate for catalytic triad (Ser759, Asp760 and Asp761) and the two magnesium ions 1003 and 1004 along 300 ns of MD simulation of systems: RdRp-RNA in blue, RdRp-RNA-ddhCTP in green, RdRp-RNA-CTP in red.

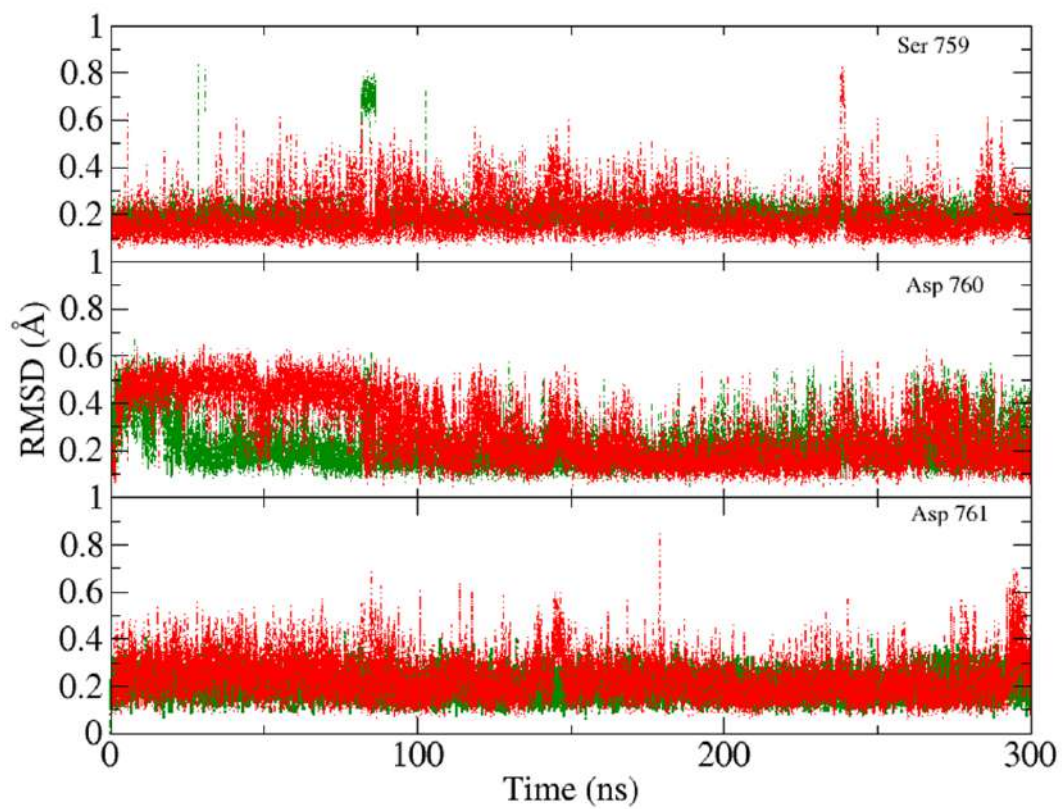


Figure S7. RMSD calculate separately for catalytic triad (Ser759, Asp760 and Asp761) along 300 ns of MD simulation of systems: RdRp-RNA-ddhCTP in green, RdRp-RNA-CTP in red.

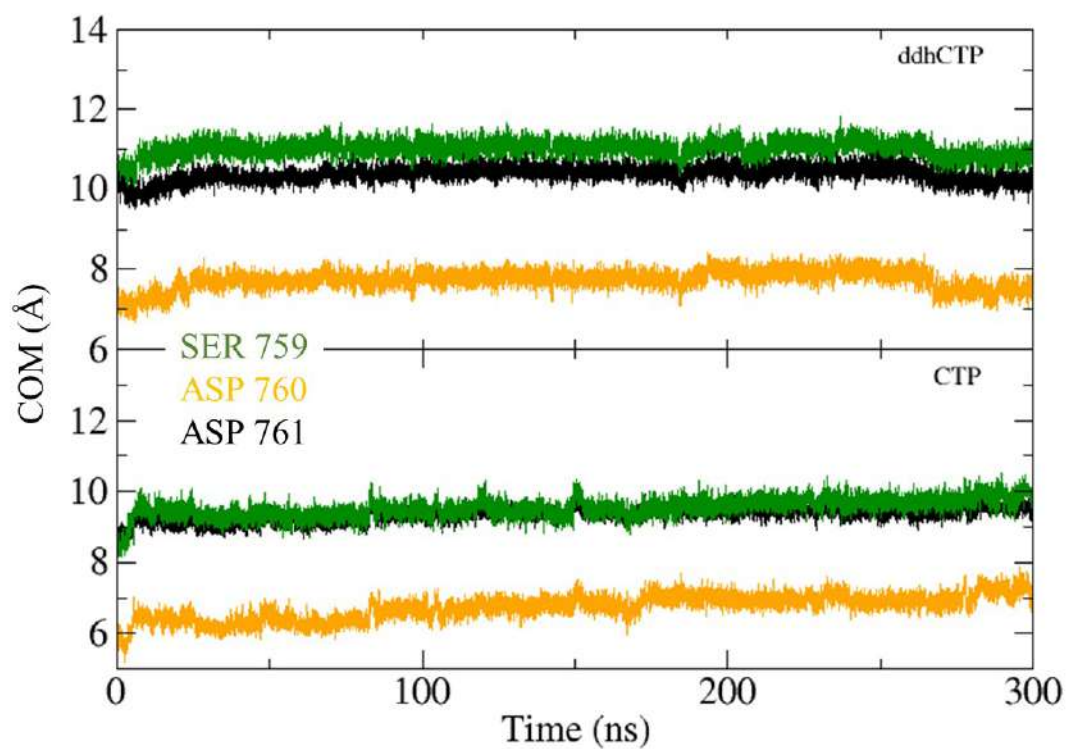


Figure S8. Distances in Å between the center of mass (COM) of amino acid residue Ser759 (in green), Asp757 (in yellow), Asp759 (in black) and the COM of ddhCTP and CTP in a) and b) respectively.

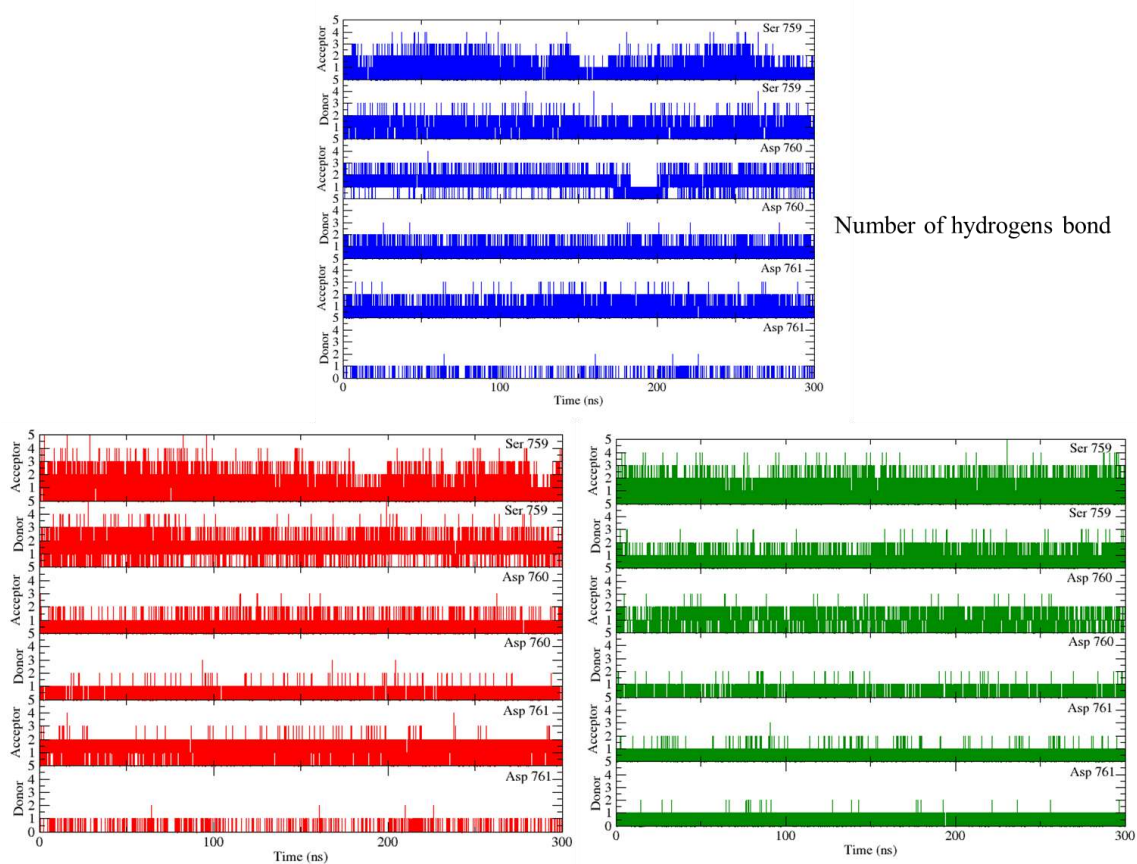


Figure S9. Number of hydrogens bonds involving Ser759, Asp760 and Asp761 (acting as acceptor of H-bond) and protein part (acting as donor) calculate for 300 ns of MD simulation of systems: RdRp-RNA in blue, RdRp-RNA-ddhCTP in green, RdRp-RNA-CTP in red respect to single residue of active site, as reported in each panel.



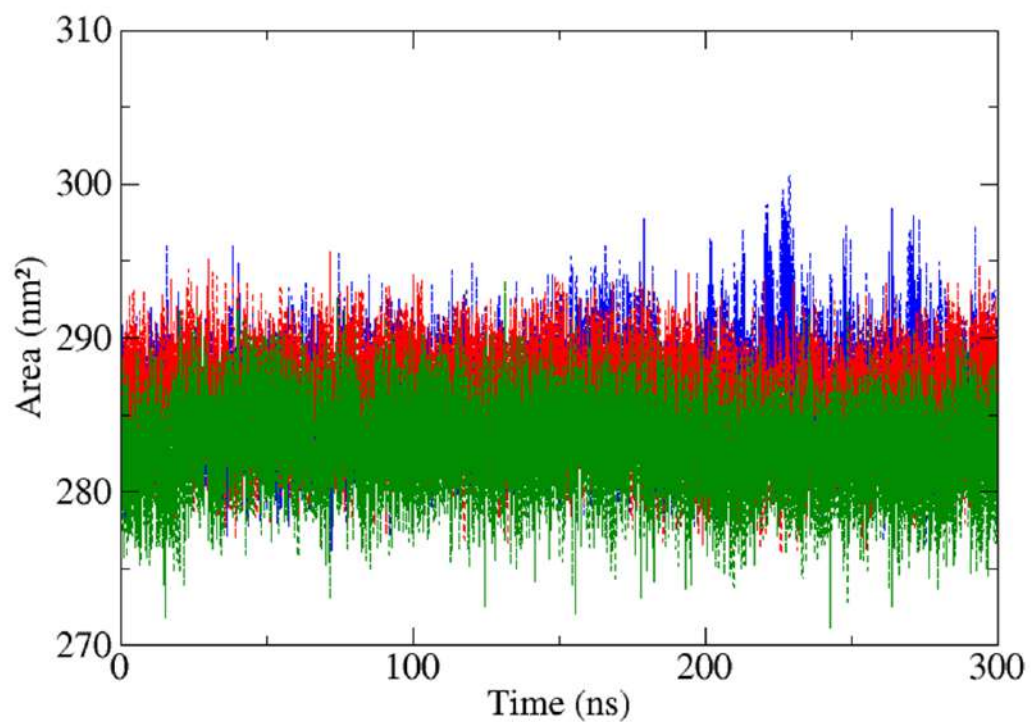


Figure S10. Solvent accessible surface area (SASA) calculated respect to protein catalytic residues Ser 759, Asp 760 and Asp 761, and two  $Mg^{2+}$  ions, of RdRp-RNA in blue, RdRp-RNA-CTP in red and RdRp-RNA-ddhCTP in green.

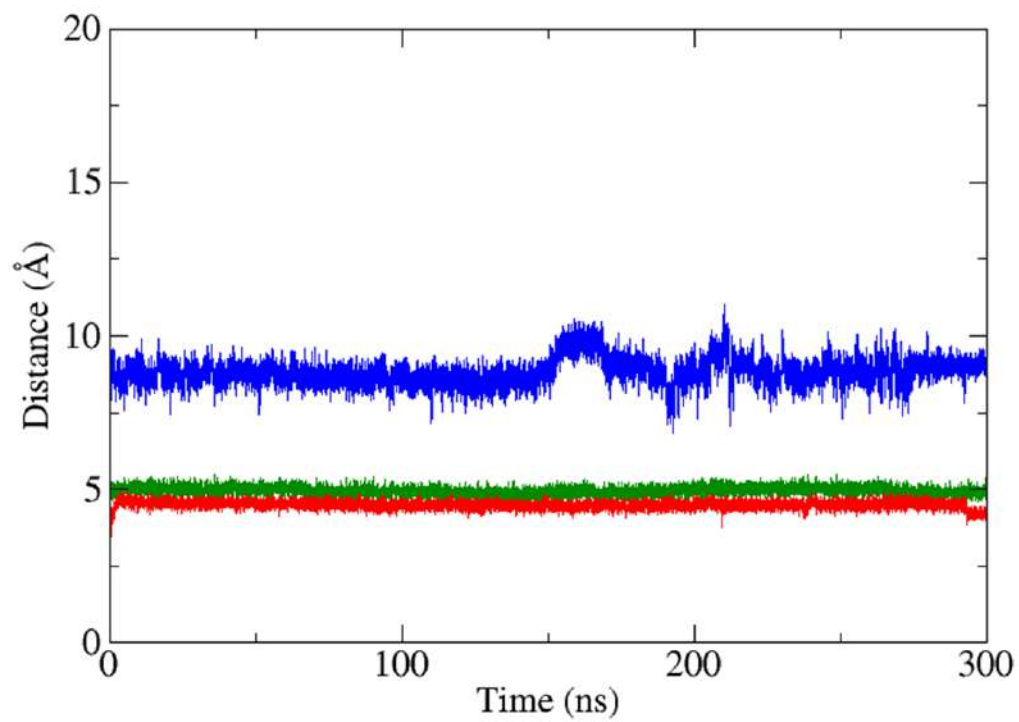


Figure S11. Distances in Å between two magnesium ions along 300 ns of MD simulation of RdRp-RNA in blue, RdRp-RNA-CTP in red and RdRp-RNA-ddhCTP in green.



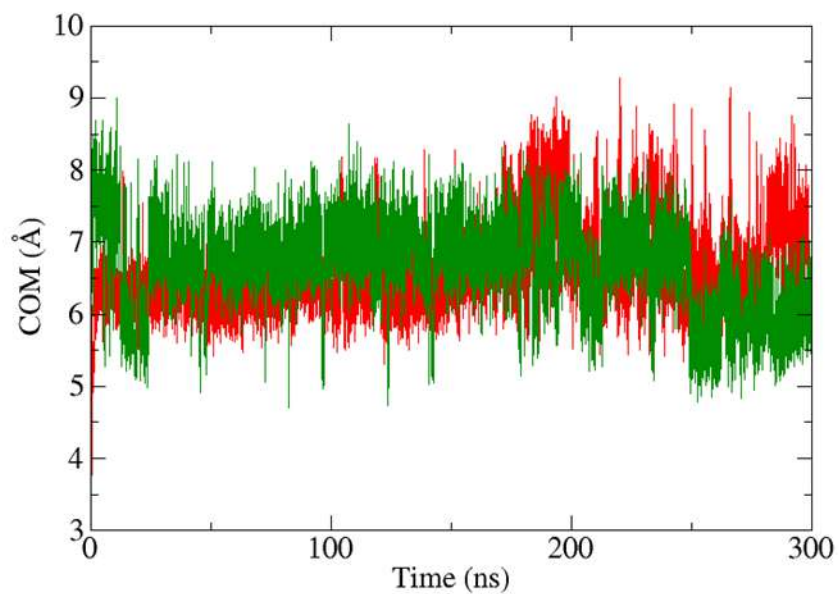


Figure S12. Distances in Å between along 300 ns of MD simulation between primer uracil 3'-terminal 3'-OH group and CTP and ddhCTP, in red and green respectively.

Table S1. Table of 10 structures obtained by RMSD-based clustering of RdRp-RNA 300ns MD simulations. From first column: cluster name, number of frames included in the cluster, population respect to total processed frames in percentage, average distance between frames inside the cluster expressed in Å, standard deviation related to Avg d and the centroid frame.

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	4998	31.6	1.34	0.16	3773	1.85
1	2434	15.4	1.35	0.18	6759	1.79
2	1869	11.8	1.31	0.14	14789	2.22
3	1544	9.7	1.30	0.15	11177	1.84
4	1456	9.2	1.36	0.17	12531	1.77
5	981	6.2	1.37	0.19	9546	1.83
6	895	5.7	1.22	0.16	628	1,87
7	652	4.1	1.35	0.19	10271	1.77
8	627	0.4	1.26	0.14	8890	1.75
9	384	2.4	1.19	0.15	13909	1.89

Table S2. Table of 10 structures obtained by RMSD-based clustering of RdRp-RNA-CTP 300ns MD simulations. From first column: cluster name, number of frames included in the cluster, population respect to total processed frames in percentage, average distance between frames inside the cluster expressed in Å, standard deviation related to Avg d and the centroid frame.

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	6884	45.5	1.60	0.24	10357	6.52
1	2433	16.1	1.5	0.23	5285	6.51
2	1832	12.1	1.53	0.26	1171	6.50
3	1737	11.5	1.49	0.21	4008	6.52
4	988	6.5	1.45	0.22	14854	6.68
5	624	4.1	1.44	0.25	417	6.77
6	474	3.1	1.48	0.24	14070	6.58
7	137	0.9	1.26	0.18	6199	12.65
8	18	0.1	0	0	6366	12.51
9	4	0	0	0	13553	12.54

Table S3. Table of 10 structures obtained by RMSD-based clustering of RdRp-RNA-ddhCTP 300ns MD simulations. From first column: cluster name, number of frames included in the cluster, population respect to total processed frames in percentage, average distance between frames inside the cluster expressed in Å, standard deviation related to Avg d and the centroid frame.

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	3451	23.00	1.42	0.18	11446	1.92
1	2984	19.90	1.38	0.16	7235	1.86
2	2803	18.70	1.38	0.16	3235	1.90
3	2109	14.00	1.35	0.16	14730	1.96
4	1263	8.40	1.40	0.20	5202	1.83
5	1045	7.00	1.43	0.19	619	2.10
6	766	5.10	1.32	0.17	6170	1.95
7	568	3.80	1.35	0.20	1356	1.89
8	14	0.10	0.00	0.00	10492	2.00
9	11	0.10	0.79	0.00	1	2.29

## Supplementary Material

- **Quantum mechanical predictions of the antioxidant capability of moracin C isomers**

•

Angela Parise<sup>1,2</sup>, Bruna Clara De Simone<sup>1</sup>, Tiziana Marino<sup>1</sup>, Marirosa Toscano<sup>1</sup>, Nino Russo<sup>1\*</sup>

1 Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Rende, Italy

2 Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France

\* **Correspondence:**

Nino Russo

nrosso@unical.it

Supplementary Data

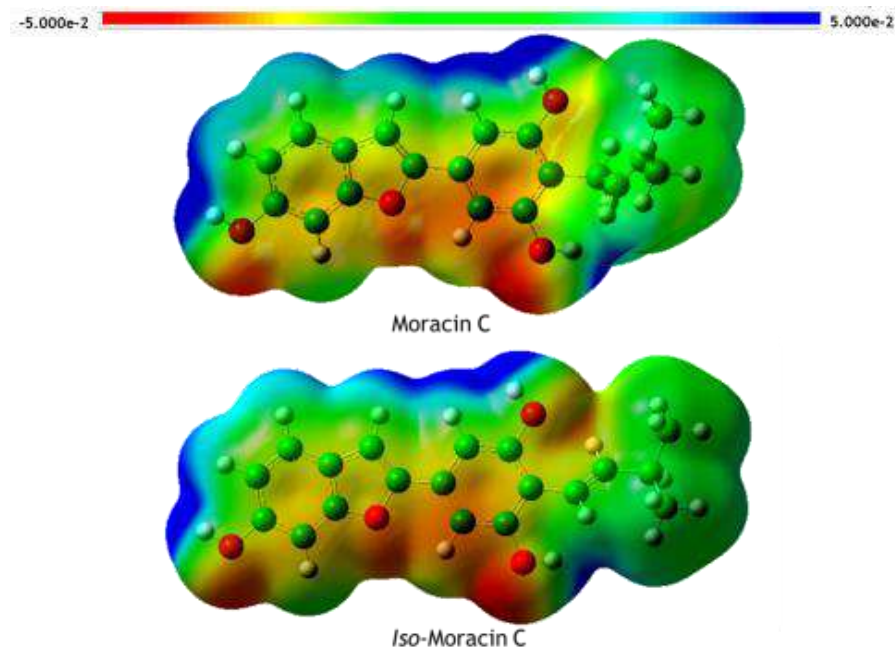


Figure S1. Molecular electrostatic potential (MEP) for moracin C and *iso*-moracin C structures optimized in water at M05-2X/6-311++G(d,p) level of theory.

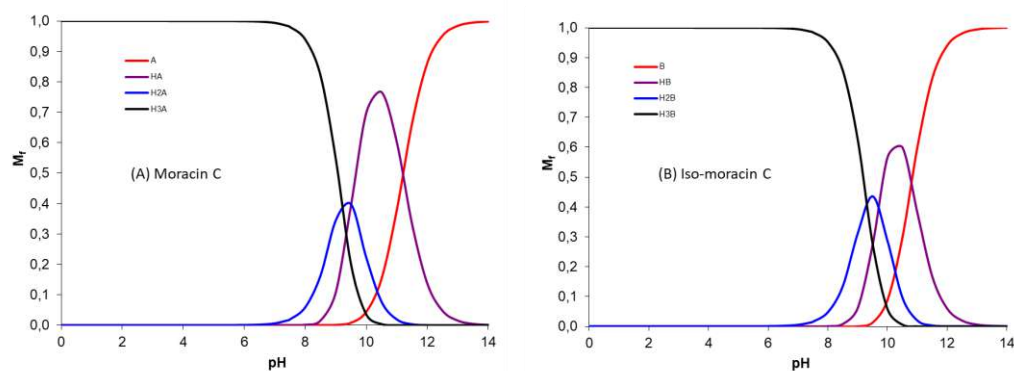


Figure S2. Molar fractions ( $M_f$ ) of the different acid–base species of moracin C and *iso*-moracin C at physiological pH.

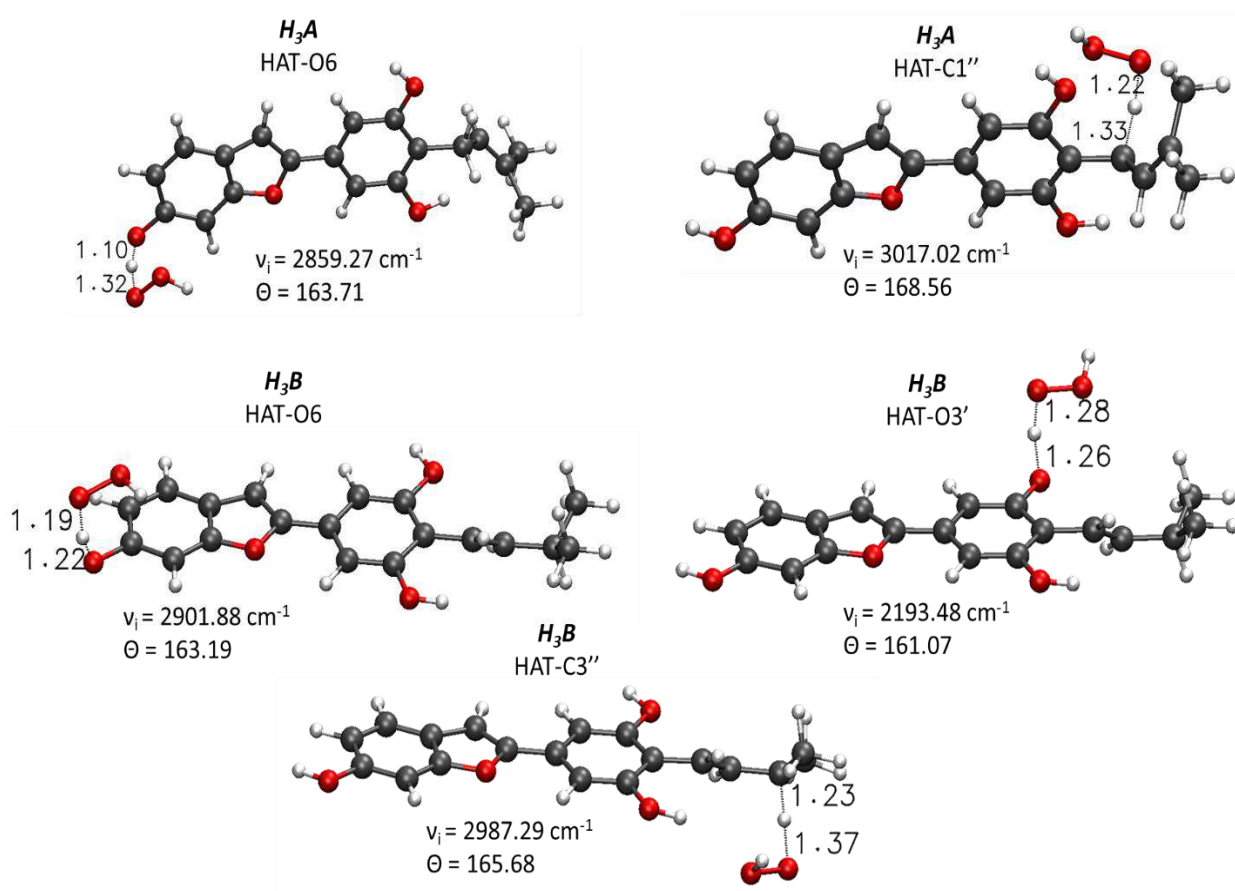
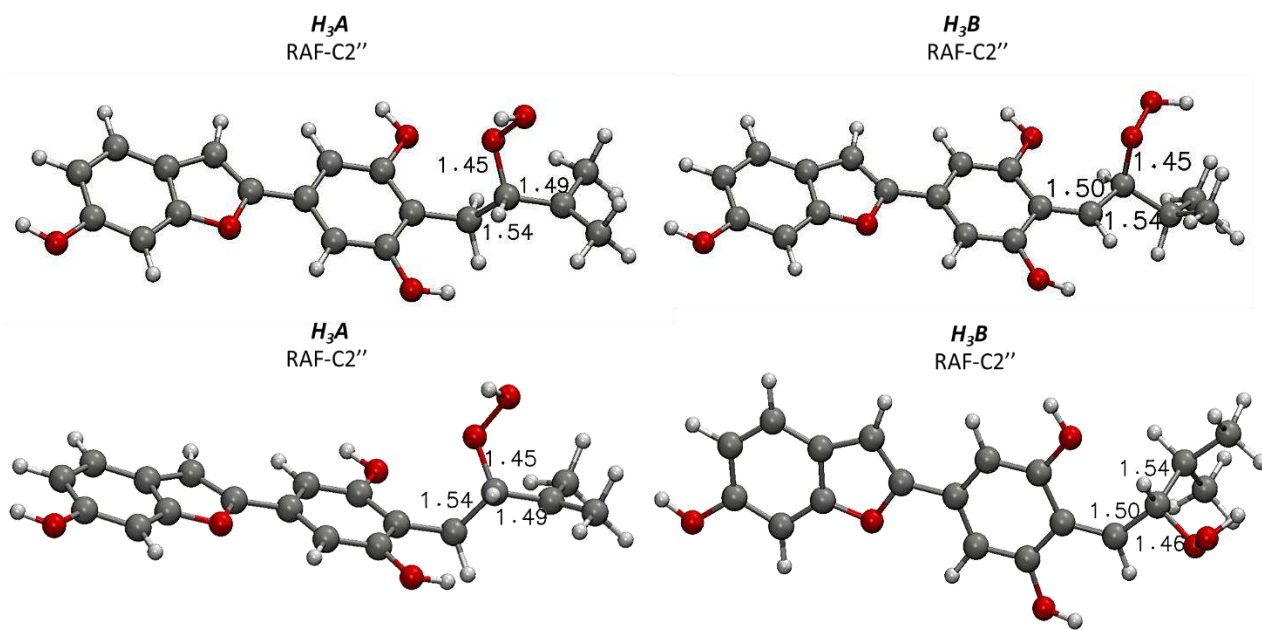


Figure S3. Main geometrical parameters for the optimized TSs structures in lipidic environment for the HAT mechanism of neutral moracin C and *iso*-moracin C. Bond lengths are in Å, angles in degrees and imaginary frequencies in  $\text{cm}^{-1}$ .



**Figure S4.** Main geometrical parameters for the optimized structures in water (on top) and lipidic (down) solvent, concerning the RAF mechanism, for the OOH-moracin C and OOH-iso-moracin C. Bond lengths are in Å.

**Table S1.** NBO charge values in  $|e|$  for the optimized geometries of moracin C and *iso*-moracin C.

<b>NBO</b>	Moracin C	Iso-moracin C
O (6)	-0.761	-0.761
H (6)	0.552	0.552
O (3')	-0.758	-0.754
H (3')	0.556	0.554
O (5')	-0.756	-0.756
H (5')	0.556	0.552
C (4')	-0.152	-0.193
C (1'')	-0.533	-0.301
C (2'')	-0.266	-0.170
C (3'')	-0.029	-0.339
C (4'')	-0.743	-0.704
H1 (C4'')	0.261	0.244
H2 (C4'')	0.255	0.251
H3 (C4'')	0.260	0.252
C (5'')	-0.733	-0.719
H1 (C5'')	0.257	0.246
H2 (C5'')	0.257	0.251
H3 (C5'')	0.255	0.256



**Table S2.** CC bond lengths and atomic spin density values for the OOH addition to the C2'' atom of Moracin C and Iso-Moracin C in water (A) and PE (B) environments

Atom	Water				PE			
	H <sub>3</sub> A RAF-C2''		H <sub>3</sub> B RAF-C2''		H <sub>3</sub> A RAF-C2''		H <sub>3</sub> B RAF-C2''	
	d (Å)	Spin density	d (Å)	Spin density	d (Å)	Spin density	d (Å)	Spin density
1	-	3.60 10 <sup>-05</sup>	-	1.81 10 <sup>-03</sup>	-	3.60 10 <sup>-05</sup>	-	4.31 10 <sup>-03</sup>
8	1.361	2.34 10 <sup>-04</sup>	1.369	1.24 10 <sup>-02</sup>	1.371	2.34 10 <sup>-04</sup>	1.369	2.83 10 <sup>-02</sup>
7	1.395	-1.64 10 <sup>-04</sup>	1.384	-8.52 10 <sup>-03</sup>	1.394	-1.64 10 <sup>-04</sup>	1.381	-2.00 10 <sup>-02</sup>
6	1.394	1.99 10 <sup>-04</sup>	1.392	1.40 10 <sup>-02</sup>	1.394	1.99 10 <sup>-04</sup>	1.390	3.30 10 <sup>-02</sup>
5	1.410	-6.50 10 <sup>-05</sup>	1.411	-5.02 10 <sup>-03</sup>	1.409	-6.50 10 <sup>-05</sup>	1.411	-1.22 10 <sup>-02</sup>
4	1.391	2.95 10 <sup>-04</sup>	1.392	1.49 10 <sup>-02</sup>	1.382	2.95 10 <sup>-04</sup>	1.393	3.64 10 <sup>-02</sup>
9	1.460	-2.01 10 <sup>-04</sup>	1.460	-1.63 10 <sup>-02</sup>	1.403	-2.01 10 <sup>-04</sup>	1.403	-3.70 10 <sup>-02</sup>
3	1.441	1.22 10 <sup>-03</sup>	1.443	5.25 10 <sup>-02</sup>	1.441	1.22 10 <sup>-03</sup>	1.440	1.25 10 <sup>-01</sup>
2	1.351	-5.32 10 <sup>-04</sup>	1.357	-4.26 10 <sup>-02</sup>	1.357	-5.32 10 <sup>-04</sup>	1.361	-9.33 10 <sup>-02</sup>
1'	1.459	3.23 10 <sup>-03</sup>	1.453	1.17 10 <sup>-01</sup>	1.462	3.23 10 <sup>-03</sup>	1.451	2.46 10 <sup>-01</sup>
2'	1.403	2.45 10 <sup>-03</sup>	1.401	-7.92 10 <sup>-02</sup>	1.403	2.45 10 <sup>-03</sup>	1.411	-1.45 10 <sup>-01</sup>
3'	1.392	-1.35 10 <sup>-03</sup>	1.382	1.26 10 <sup>-01</sup>	1.381	-1.35 10 <sup>-03</sup>	1.384	2.60 10 <sup>-01</sup>
4'	1.404	-2.13 10 <sup>-02</sup>	1.441	-1.03 10 <sup>-01</sup>	1.402	-2.13 10 <sup>-02</sup>	1.439	-2.75 10 <sup>-01</sup>
1''	1.508	1.33 10 <sup>-02</sup>	1.400	8.85 10 <sup>-01</sup>	1.503	1.33 10 <sup>-02</sup>	1.403	7.44 10 <sup>-01</sup>
2''	1.543	-1.59 10 <sup>-02</sup>	1.504	-3.79 10 <sup>-02</sup>	1.544	-1.59 10 <sup>-02</sup>	1.499	-8.85 10 <sup>-02</sup>
3''	1.513	9.26 10 <sup>-01</sup>	1.541	3.04 10 <sup>-03</sup>	1.491	9.26 10 <sup>-01</sup>	1.542	2.64 10 <sup>-02</sup>
4''	1.491	3.23 10 <sup>-02</sup>	1.530	1.51 10 <sup>-03</sup>	1.490	3.23 10 <sup>-02</sup>	1.531	-2.67 10 <sup>-03</sup>