



UNIVERSITÀ DELLA CALABRIA


"La presente tesi è cofinanziata con il sostegno della Commissione Europea, Fondo Sociale Europeo e della Regione Calabria. L'autore è il solo responsabile di questa tesi e la Commissione Europea e la Regione Calabria declinano ogni responsabilità sull'uso che potrà essere fatto delle informazioni in essa contenute"

UNIVERSITÀ DELLA CALABRIA

UNIVERSITA' DELLA CALABRIA

Dipartimento di Chimica e Tecnologie Chimiche – CTC

Scuola di Dottorato

Scienza e Tecnica "Bernardino Telesio"

Indirizzo

Metodologie Chimiche Inorganiche

Con il contributo di

Commissione Europea, Fondo Sociale Europeo e della Regione Calabria

CICLO XXVII

TITOLO TESI

Theoretical Investigation Of The Enzyme Promiscuity Within The Carbonic Anhydrase's Family

Settore Scientifico Disciplinare CHIM/03 CHIMICA GENERALE E INORGANICA

Direttore: Ch.mo Prof. Roberto Bartolino

Firma _____

Supervisore: Prof.ssa Tiziana Marino

Firma _____

Dottorando: Dott. Paolo Piazzetta

Firma _____

Theoretical investigation of the enzyme promiscuity within the Carbonic Anhydrase's family

Paolo Piazzetta

Supervisor: Prof.ssa Tiziana Marino

Università degli Studi della Calabria

Dipartimento di Chimica e Tecnologie Chimiche – CTC

Rende, Italy, 2014

Introduzione

Gli enzimi sono importanti catalizzatori che mediano, in organismi biologici, molte trasformazioni chimiche e sono coinvolti direttamente nella crescita delle cellule, nel metabolismo degli alimenti, nella segnalazione, regolazione, e trasduzione di energia. La presenza di ioni metallici come cofattori, gioca un ruolo sia catalitico che strutturale per catalizzare reazioni che spesso risultano difficili da realizzare. I metallo-enzimi presentano caratteristiche comuni, cioè che lo ione metallico legato alla proteina attraverso un sito di coordinazione labile è situato in una 'sacca' la quale forma, di solito, è compatibile con il substrato.
Recentemente è stato ipotizzato che un'estesa specificità, o funzione promiscua, di alcuni enzimi ha permesso loro di processare una moltitudine di funzioni sono state necessarie per preservare organismi ancestrali.
La presente tesi è principalmente incentrata sul meccanismo di reazione dell'anidrasi carbonica (CA, EC 4.2.1.1) atto a processare substrati e reazioni promiscue.
Un ampio range di metodi teorici, da quelli basati sul funzionale della densità (DFT) come gli approcci Quantum Mechanical (QM) e Quantum Mechanics/Molecular Mechanics (QM/MM), a quelli classici, come le simulazioni di dinamica molecolare, sono stati usati per investigare i meccanismi di reazione proposti.
In questa tesi, il Capitolo 1 è incentrato sugli aspetti teorici mentre nel Capitolo 2 sono descritti i metodi computazionali usati.

Gli aspetti promiscui dell'attività catalitica della CA sono elucidati nel Capitolo 3, mentre, un'esaustiva e dettagliata descrizione dei vari casi presi in esame è contenuta nel Capitolo 4.
In dettaglio sono stati investigati:

i) il comportamento catalitico della CA nei confronti di due substrati isoelettronici la CO_2, carbodimmide ed acido cianico, a livello QM e QM/MM in parallelo con il processo catalitico naturale (Paper I e II)

ii) la CA come carbonil solfuro idrolasi nell'efficiente conversione del COS in H_2S e CO_2 attraverso l'uso dell'approccio cluster (QM) (Paper III)

iii) la promiscuità catalitica indotta dalla sostituzione dello ione nativo con il Rh(I) al fine di catalizzare l'idrogenazione diretta del diossido di carbonio.

iv) gli eventi di rilascio dei prodotti di reazione inerenti il processo nativo (bicarbonato) e promiscuo (ureato) attraverso tecniche denominate SMD (steered molecular dynamics) (Paper V)

In fine, con l'intento di esplorare l'attività esterasica della CA è stato intrapreso uno studio basato su gli approcci QM e QM/MM. La scelta del modello per riprodurre il sito catalitico e la descrizione del complesso di Michaelis Menten sono stati riportati (In fase di studio).

Introduction

Enzymes are important biocatalysts that mediate most chemical transformations in organisms and are involved in cell growth, food metabolism, signaling, regulation, and energy transduction. The presence of cofactors as metal ions, plays a structural and catalytic roles in the enzymatic chemistry for catalyzing reactions that are often difficult to be realized. Metalloenzymes present a common feature, namely that the metal ion is bound to the protein with one labile coordination site located in a pocket whose shape fits usually the substrate.

Recently has been hypothesized that a broad specificity, or promiscuous functions, of some enzymes enabled them to perform a multitude of functions that were necessary to maintain ancestral organisms.

The present thesis has its main focus on the working mechanisms followed by carbonic anhydrase (CA, EC 4.2.1.1) into processing promiscuous substrates and promiscuous reactions.

A wide range of theoretical methods from that based on density functional theory (DFT), including Quantum Mechanical only (QM) and combined Quantum Mechanics/Molecular Mechanics (QM/MM) approaches, to that classical, such as molecular dynamics simulations, has been used to capture the main features of the proposed mechanisms.

In this thesis, a brief outline of theoretical background is given in Chapter 1 and the used computational methods are described in Chapter 2.

Some features characterizing the promiscuous aspect of the catalytic activity of CA will be highlighted in Chapter 3, while the comprehensive investigation of the various promiscuous considered cases of CA is described in Chapter 4. In particular, detailed insights have been obtained:

i) on the catalytic behavior of CA toward two promiscuous substrates as carbodiimide and cyanic acid in comparison with that shown toward the native carbon dioxide by means of fully QM and combined QM/MM investigations; (**Paper I, II**)

ii) on the working mechanism of CA acting as a carbonyl sulfide hydrolase that efficiently converts COS in H_2S and CO_2 by using the cluster model approach; (**Paper III**)

iii) on the CA catalytic promiscuity induced by replacing the native zinc cation with a Rh(I) in order to catalyze the direct hydrogenation of carbon dioxide. (**Paper IV**)

Furthermore, in order to gain insight on the restoring of the catalytic cycle of CA and to unravel at atomic level the factors involved during the releasing process of the products for both native and promiscuous (carbodiimide) substrates, steered molecular dynamics simulations have been carried out. (**Paper V**)

At the end with the intention of exploring the esterase catalytic activity of CA toward activated esters a theoretical study based on QM cluster model and QM/MM approaches has been undertaken. The choice of the model for mimic the active site of CA in this particular catalytic process, along with the description of the Michaelis Menten complex obtained is reported. (***Work in progress***)

Acknowledgements

I am using this opportunity to express my gratitude to the Theoretical and Computational Chemistry Group at department of chemistry, university of Calabria, which supported me throughout the course of these three years.

I owe my sincere thanks to my supervisor Prof.ssa Tiziana Marino. I am thankful for her aspiring guidance, invaluably constructive criticism and friendly advice during the project work.

I would like to express my special appreciation and thanks to the head of the group Prof. Nino Russo, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on research have been priceless.

I would also like to thank Prof. Dennis R. Salahub who gave me the opportunity to spend a formative training period in his research group at University of Calgary, Canada.

I also thank Prof. Michele Parrinello at Department of Chemistry and Applied Biosciences, ETH Zürich USI-Campus, who introduced me to the methods for the accurate binding free-energy calculation.

List of Papers included in this thesis

I. **Promiscuous Ability of Human Carbonic Anhydrase: QM and QM/MM Investigation of Carbon Dioxide and Carbodiimide Hydration**

Paolo Piazzetta, Tiziana Marino, and Nino Russo
Inorg. Chem. 2014, *53*, 3488–3493

II. **Insight into the promiscuous activity of human carbonic anhydrase against the cyanic acid substrate from a combined QM and QM/MM investigation**

Paolo Piazzetta, Tiziana Marino, and Nino Russo
Phys. Chem.Chem.Phys. 2014, *16*, 16671-16676

III. **The working mechanism of the β-carbonic anhydrase degrading carbonyl sulfide (COSase): a theoretical study**

Paolo Piazzetta, Tiziana Marino, and Nino Russo
Submitted manuscript

IV. **Direct Hydrogenation of Carbon Dioxide by an Artificial Reductase Obtained Substituting Zinc with Rhodium in the Carbonic Anhydrase Catalytic Center. A mechanistic study**

Paolo Piazzetta[1], Tiziana *Marino*[1], Nino Russo[1], D. R. Salahub[2]

[1]Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Rende, Italy

[2]IQST – Inst. for Quantum Science and Technology, CMS – Centre for Molecular Simulation, ISEEE – Inst. for Sustainable Energy, Environment and Economy BI 556
University of Calgary, 2500 University Drive NW
Calgary, Alberta, Canada T2N 1N4

Manuscript in preparation

V. **Theoretical investigation on the restoring step of the carbonic anhydrase catalytic cycle**

Paolo Piazzetta, Tiziana Marino, and Nino Russo
Manuscript in preparation

Abbreviations and Acronysms

AFM	Atomic Force Microscopy
AMBER	Assisted Model Building and Energy Refinement
ASC	Apparent Surface Charge
B3LYP	Becke 3 parameter Lee – Yang – Parr functional
CA	Carbonic Anhydrase
CC	Coupled Cluster theory
CHARMM	Chemistry at HARvard Molecular Mechanics
CI	Configuration Interaction
COS	Carbonyl Sulfide
DFT	Hybrid Functional Theory
EVB	Empirical Valence Bond
FES	Free Energy Surface
FF	Force Field
FM	Funnel-Metadynamics
GGA	Generalized Gradient Approximation
GROMOS	GROningen MOlecular Simulation
GVB	Generalized Valence Bond
HF	Hartree-Fock
HMDFT	Hybrid Meta Density Functional Theory method
IRC	Intrinsic reaction coordinate
KS	Kohn-Sham
L(S)DA	Local (Spin) Density Approximation
LJ	Lennard-Jones potential
MCSCF	Multiconfigurational Self-Consistent Field
MD	Molecular Dynamics
MM	Molecular Mechanics
MP	Møller–Plesset
MPW	modified Perdew and Wang exchange functional
NBO	Natural Bond Orbital
ONIOM	Our own N-layered Integrated molecular Orbital and molecular Mechanics
OPLS	Optimized Potential for Liquid Simulations
PCM	Polarizable Continuum Model
PDB	Protein Data Bank
PES	Potential Energy Surface
PME	Particle Mesh Ewald

PMF	Potential Of Mean Force
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics
RC	Reaction Coordinate
RMSD	Root Mean Square Deviation
SAS	Solvent Accessible Surface
SCF	Self-Consistent Field
SMD	Steered Molecular Dynamics
TST	Transition State Theory
UFF	Universal Force Field
VWN	Vosko, Wilk and Nusair
ZPE	Zero point energy

Amino acids abbreviations

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

Contents

1 Theoretical Background

Quantum mechanics has been the most powerful theory used to describe the chemical properties. A general view concerning quantum biochemistry is traced in this chapter with the aim to briefly introduce the theory basics which acts as pillars of the modern quantum chemical methods. A more comprehensive description is available in standard textbooks [1-7].

1.1 Hartree-Fock Methods

The main aim of studying theoretical chemistry is to define the properties of molecular systems in term of electronic structure. Electrons are light particles that don't obey classical behaviour so they should be treated quantum mechanically. In a treatment in which time – dependent interaction are not taken in account, the system can be described by the *time – independent Schrödinger equation* [8]:

$$\hat{\mathbf{H}}\boldsymbol{\Psi} = E\boldsymbol{\Psi} \qquad \textbf{(1.1)}$$

Even applying the common limitations of *non – relativistic* Hamilton operator $\hat{\mathbf{H}}$ and *Born – Oppenheimer approximation* [9], the Schrödinger equation can be solved exactly only for one electron system at most, such as hydrogen atom or H_2^+.

The Hartree-Fock (HF) method laid a basis for *ab initio* approach by considering the trial wave function as a product of orthonormal spin-orbitals $\Psi_i(\mathbf{r}_i, s_i)$, where each electrons move in the average field generated by all the others electrons [10-12]. Here, $\mathbf{r}_i = (x, y, z)$ is the vector variable and $s_i = \frac{1}{2}, -\frac{1}{2}$ is the spin variable of the i-th electron. Spin was included as an *ad hoc* addition to nonrelativistic model. Hereafter the method was improved by introducing the *Pauli exclusion principle* [13,14]

with the successive introduction of the *Slater Determinant* to represent the n-electron wave function:

$$\boldsymbol{\Psi}_n^{\mathrm{SD}} = \frac{1}{\sqrt{n!}} \begin{vmatrix} \Psi_1(\mathbf{r}_1, s_1) & \Psi_2(\mathbf{r}_1, s_1) & \cdots & \Psi_n(\mathbf{r}_1, s_1) \\ \Psi_1(\mathbf{r}_2, s_2) & \Psi_2(\mathbf{r}_2, s_2) & \cdots & \Psi_n(\mathbf{r}_2, s_2) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_1(\mathbf{r}_n, s_n) & \Psi_2(\mathbf{r}_n, s_n) & \cdots & \Psi_n(\mathbf{r}_n, s_n) \end{vmatrix} \qquad \textbf{(1.2)}$$

Having the general form of the solution and applying the *variational principle* to it leads to HF equation that can be processed iteratively, resulting in solution to a certain degree of accuracy, called *self-consistent field* (SCF).

However with the HF approach the motion of the electrons with opposite spins is not correlated. Electron correlation can be included explicitly with well-known extensions collectively called post-Hartree-Fock methods [15] like *Møller–Plesset perturbation theor*y (MP), the *generalized valence bond* method (GVB), *multiconfigurational self-consistent field* (MCSCF), *configuration interaction* (CI) *and coupled cluster theory* (CC). These approaches improve the level of accuracy but become computationally much more demanding, and thus are only suitable for relatively small systems. To handle larger systems an alternative approach has been developed.

1.2 **Density Functional Theory**

Density Functional Theory (DFT) operates with the electron density $\rho(\mathbf{r})$ as basic variable, in contrast to the wave function $\boldsymbol{\Psi}_n(\mathbf{r}_1, s_1, \mathbf{r}_2, s_2, \ldots, \mathbf{r}_n, s_n)$. Hohenberg-Kohn (1964) [16] introduced that there exist a unique relationship between ρ and all fundamental properties of a given system, such as energy and wave function. For many-electron systems, dealing with only 3 variables of ρ is much more handy, than $3n$ variables of $\boldsymbol{\Psi}_n$, excluding the spin variable. So the first Hohenberg-Kohn theorem states that every observable of a stationary quantum mechanical system can be calculated, in principle exactly, from the ground-state density $\rho(\mathbf{r})$, i.e., every observable can be written as a functional of the ground-state density. Consider an overall energy E_v for a system with an external potential $v(\mathbf{r})$ acting on the electrons due to the nuclear charges. Omitting the interaction between nuclei, E_v can be

splitted into three terms which are kinetic energy T, electron-electron repulsion E_{ee} and the nuclei-electron attraction V:

$$E_v[\rho] = T[\rho] + E_{ee}[\rho] + V[\rho] = T[\rho] + E_{ee}[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \qquad \textbf{(1.3)}$$

Because the operators T and E_{ee} depend exclusively on the coordinates of the electrons and their forms are the same for all systems, depending only on the number of electrons, they are grouped together into the *universal functional* F[ρ]. Trivial dependence on $v(\mathbf{r})$ thus should be left out:

$$F[\rho] = E_v[\rho] - V[\rho] = T[\rho] + E_{ee}[\rho] \qquad \textbf{(1.4)}$$

An SCF-like procedure allows to optimize a set of *Kohn-Sham* (KS) [17] orbitals, similarly to HF. Notwithstanding one-electron KS orbital ${\Psi_i}^{KS}$cannot carry the significance of HF orbitals and their densities just add up to form the total density ρ^{KS} of an n-electron system:

$$\rho^{\mathrm{KS}} = \sum_i^n \sum_s \left|{\Psi_i}^{KS}(\mathbf{r}, s)\right|^2 \qquad \textbf{(1.5)}$$

Previous equation reflects an essential feature of the KS method that orbitals are restricted to have occupation number 0 and 1, which holds true for the determinant wave function with independent particles. The corresponding component of kinetic energy in eq.(1.4) is:

$$T_S[\rho] = \sum_i^n \left\langle {\Psi_i}^{KS} \middle| -\frac{1}{2}\nabla^2 \middle| {\Psi_i}^{KS} \right\rangle \qquad \textbf{(1.6)}$$

The S subscript of $T_S[\rho]$ implies energy to be calculated from Slater determinant. The trivial component of the electron interaction energy E_{ee} is classical *coloumb* $J[\rho]$ term which represent the electron-electron repulsion and can be defined explicitly:

$$J[\rho] = \frac{1}{2}\int\int \frac{\rho[\mathbf{r}]\rho[\mathbf{r}']}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}\, d\mathbf{r}' \qquad \textbf{(1.7)}$$

Basing on that definition, eq.(1.4) can now be reformulated as:

$$F^{KS}[\rho] = T_S[\rho] + J[\rho] + E_{xc}[\rho] \qquad \textbf{(1.8)}$$

where

$$E_{xc}[\rho] = (T[\rho] - T_S[\rho]) + (E_{ee}[\rho] - J[\rho]) \qquad \textbf{(1.9)}$$

$E_{xc}[\rho]$ is called the *exchange-correlation energy functional*, which can be considered as the difference between $T[\rho]$ and $T_S[\rho]$ plus the non-classical part of $E_{ee}[\rho]$. The KS methods moves all the nontrivial terms of the full

functionals $F[\rho]$ to $E_{xc}[\rho]$ [6]. It is common in DFT separate $E_{xc}[\rho]$ in two parts, a pure exchange $E_x[\rho]$ and a correlation $E_c[\rho]$. The simplest approach to calculate the exchange-correlation energy $E_{xc}[\rho]$ is based on assuming that the density ρ 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) [7]. The exchange energy of a uniform electron gas $E_{xc}[\rho]$ can be computed exactly. *Gradient Corrected* or *Generalized Gradient Approximation* (GGA) makes a further step, utilizing not only the density ρ, but also the gradient ∇_ρ. Finally, the best methods of modern DFT are hybrid type: they combine functionals of LSDA, corrections from GGA, and also piece of HF exchange ${E_x}^{HF}$, calculated via KS orbitals. One popular group of hybrid methods is Becke 3 parameter functional (B3) [18,19], with the three empirical fitted parameters A, B and C:

$$E_{xc}^{\mathrm{B3}} = (1-A)E_x^{\mathrm{Dirac-Slater}} + A{E_x}^{\mathrm{HF}} + B{E_x}^{\mathrm{B88}} + {E_c}^{VWN} + C\Delta{E_c}^{GGA} \qquad \textbf{(1.10)}$$

where Δ signifies that $\nabla\rho$-dependent fractions of corresponding functionals are included. B88 is a widely used gradient correction to exchange by Becke [20]. LSDA exchange ${E_x}^{\mathrm{Dirac-Slater}}$ is given by the Dirac-Slater formula for a uniform electron gas, and LSA correlation energy ${E_c}^{VWN}$ is the functional by Vosko, Wilk and Nusair (VWN) [21].

The case when $\Delta{E_c}^{GGA}$ gradient correction to correlation in eq.(1.10) is Lee, Yang and Parr functional (LYP) [22] is known as **B3LYP**. So A determines the extent of replacement of the Slater local exchange ${E_x}^{\mathrm{Dirac-Slater}}$ by the exact HF exchange ${E_x}^{\mathrm{HF}}$; B controls the addition of Becke's gradient-correction to the exchange functional ${E_x}^{\mathrm{B88}}$; C defines the inclusion weight of the LYP correlation ${E_c}^{\mathrm{LYP}}$ and the VWN correlation ${E_c}^{VWN}$ functionals.

Concluding, DFT has a computational cost which is similar to HF and scales as N^{3-4}, with a possibility of more accurate, *a priori* exact results.

In this thesis, the most popular hybrid functional B3LYP is used to investigate the reaction mechanisms. Various assessments have been reported in the literature regarding the accuracy of B3LYP with respect to geometries and energies, which are the two crucial properties for the study of reaction mechanisms [23]. It performs well for the structure determination, for the calculation of atomization energies and other thermochemical quantities, still remaining the method of choice to obtain

geometrical parameters. Furthemore, we performed single point calculation testing the hybrids **B97-D**[24] and the meta-hybrids **MPWB1K**[25-28], **M06**[29] and **M06-2X**[29]. The first one is based on Becke's power-series ansatz from 1997 and is explicitly parameterized by including damped atom-pairwise dispersion corrections. MPWB1K is an *hybrid meta density functional theory method* (HMDFT) modelled for kinetics and it is based on the *modified Perdew and Wang exchange functional* (MPW) and Becke's 1995 correlation functional. M06 and M06-2x are based on meta-GGA approximations where M06 functional is parametrized including both transition metal and non-metals and the M06-2X functional is a highnonlocality functional with double the amount of nonlocal exchange (2X), and it is parametrized only for nonmetals.

1.3 **Transition State Theory**

One of the parameters used to monitoring a chemical reaction is the rate of the process, which often can be measured experimentally. Classical *transition state theory* (TST) [30], developed by Eyring, Polanyi, and Evans, is a statistical mechanical theory for the prediction of reaction rate constants, widely used in the interpretation of reactivity of chemical processes. With this approach the structural features of colliding species is not explicitly considered in TST, which operates in terms of energy pathway related to the *reaction coordinate* (RC). RC is the minimum energy pathway, connecting the minima of reactants **R** and products **P** on the multidimensional *potential energy surface* (PES). A *transition state* **TS** is a maximum on RC, o a saddle point on PES.

Considering a bimolecular reaction with

$$\boldsymbol{A + B} \;\rightleftharpoons\; \boldsymbol{R} \overset{K^{\neq}}{\rightleftharpoons} \boldsymbol{TS} \overset{k'}{\rightarrow} \boldsymbol{P} \;\rightleftharpoons\; \boldsymbol{C + D} \qquad \textbf{(1.11)}$$

where the equilibrium constant $K^{\neq}$ is represented as

$$K^{\neq} = \frac{[\mathbf{TS}]}{[\mathbf{R}]} \qquad \textbf{(1.12)}$$

with the concentrations of the species given in the square brackets. The distribution of energy in the reactant region and the transition state follows the Boltzmann distribution law, which means that the transition

state species is in a rapid equilibrium with reactants. This evidence permits to introduce the free energy equivalent of the equilibrium constant $K^{\neq}$**:**

$$K^{\neq} = \mathrm{e}^{-\frac{\Delta G^{\neq}}{RT}} \qquad \textbf{(1.13)}$$

where $\Delta G^{\neq}$ is the Gibbs free energy difference between **TS** and **R**, T is the absolute temperature, and R is the gas constant. In the TST the rate of reaction is assumed to be directly proportional to the concentration of the TS. Thus, since [**TS**] is small, its decomposition is the rate-determing step process of the reaction. The first-order rate constant k' in eq.(1.12) is related to the decomposition of **TS** and **P** as follow:

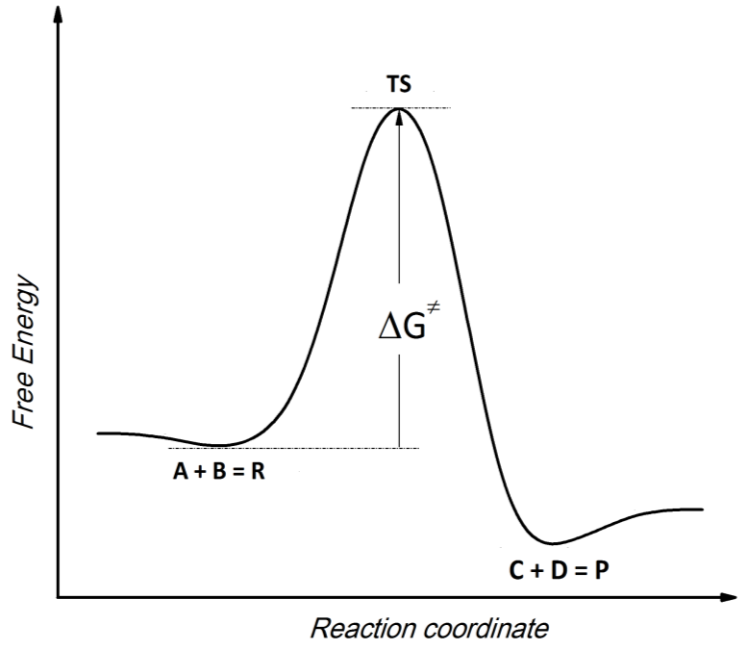


Figure 1.1 Potential Energy Surface of an elementary bimolecular reaction

$$\frac{d[\mathbf{P}]}{dt} = k[\mathbf{R}] = k'[\mathbf{TS}] \qquad \textbf{(1.14)}$$

with k first order rate constant for the reactants **R**. Eyring proposed that the evaluation of k' is related to the activated complex breakdown that occurs during its first vibrational excursion:

$$k' = k\upsilon \tag{1.15}$$

when υ is the vibrational frequency of the bond to be broken and k the transmission coefficient which represent the probability that the immediate breakdown of **TS** will end in product **P** formation. The υ value can be determined form Planck's law:

$$\upsilon = k\frac{k_B T}{h} \tag{1.16}$$

and through the eq.(1.13) it is simple to obtain the final expression for the rate constant k in the form of Arrhenius equation:

$$k = k\frac{k_B T}{h} e^{-\frac{\Delta G^{\neq}}{RT}} \tag{1.17}$$

where k_B and h are the Boltzmann and Planck's constant, respectively.

At room temperature (298.15 K), a rate constant of 1 s^{-1} corresponds to a barrier of 17.4 kcal/mol, and an increase of $\Delta G^{\neq}$ by 1.4 kcal/mol leads to a decrease by one order of magnitude in reaction rate constant.

1.4 **The principles of enzymatic catalysis**

Enzymes are complex molecular systems of a size frequently exceeding thousands of atoms, which show an high specificity for particular reactions. Enzymatic processes are typically faster as compared to uncatalyzed ones and the reason is that they can reduce the related activation barriers, with direct impact on the reaction rates.

For the uncatalyzed reaction, the barrier, $\Delta G^{\neq}$, is the relative free energy of the transition state (**TS**) to the reactant (substrate, **S**). In the corresponding enzymatic reaction, the substrate first binds to the enzyme to form an enzyme-substrate complex (**ES**), and then reacts via one (or several) transition state (**TS**) resulting in an enzyme-product complex (**EP**), from which the product (**P**) is finally released and the enzyme is set free again. According to eq.(1.17), reaction rate increase by a catalyst implies a stabilization of the activated complex, or, equivalently decrease of the free energy barrier from $\Delta G^{\neq}_{\mathbf{uncat}}$ to $\Delta G^{\neq}_{\mathbf{cat}}$.
The enzyme can provide the catalytic power to make $\Delta G^{\neq}_{\mathbf{cat}} < \Delta G^{\neq}_{\mathbf{uncat}}$, and consequently accelerates the rate of the overall reaction. There are many factors that contribute to the catalytic efficiency of enzymes. The initial "*lock and key*" model proposed by Fisher assumed that the binding of a

substrate to the enzyme activates the substrate to its reactive conformation [31]. However with the introduction of the transition state theory, it became obvious that enzymes provide a pre-organized environment that can stabilize transition states. Electrostatic effects due to the polar enzymatic surrounding are the major contributing factor to the stabilization of the transition state in many enzymatic reactions.

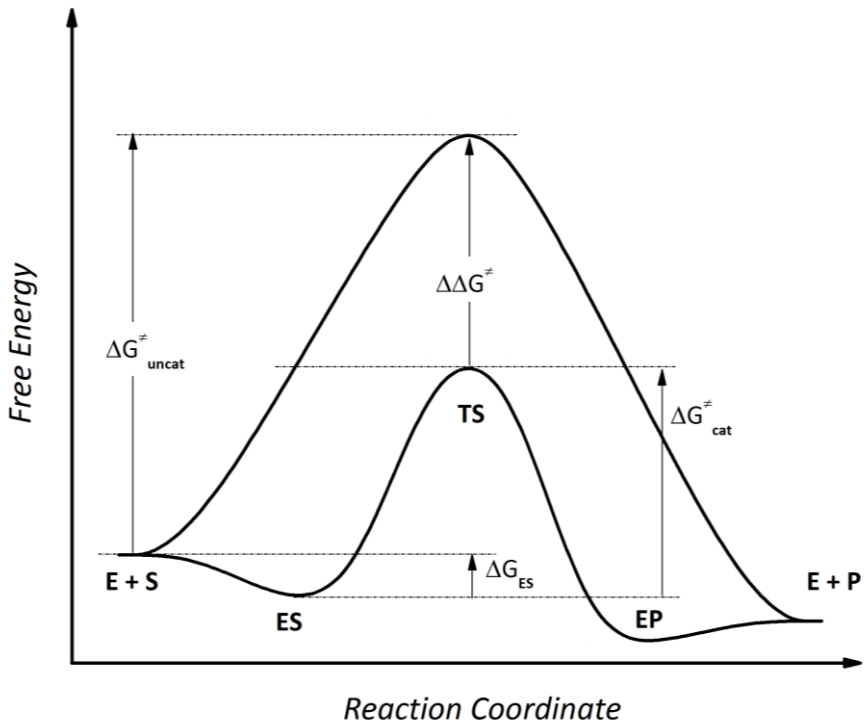


Figure 1.2 Free energy profile of activated complex during enzyme catalysis as compared to uncatalysed reaction.

For a multiple-step enzymatic reaction with several intermediates of various stability, the largest energy difference between an intermediate and a transition state in the forward direction corresponds to the rate-limiting step. The overall rate of an enzyme-catalyzed reaction can be described by a rate constant $k_{\mathbf{cat}}$ that is the equivalent to the rate constant of the rate-limiting step during the reaction. $k_{\mathbf{cat}}$ is also known as *turnover number*, as it is the number of successive reactions catalyzed by an enzyme molecule per unit of time in excess of substrate. In the 1930

L. Michaelis and M. Menten characterized the kinetics of simple reaction. K_M is known as *Michaelis constant*, and it is the concentration of substrate at which the reaction velocity is half-maximal (V_{max}) and are both experimentally measurable. The constant $k_{\mathbf{cat}}/K_M$ is a measure of how efficiently an enzyme converts a substrate into product. Nevertheless there is a physical limit on enzymes efficiency: when formation of product **P** from enzyme-substrate **ES** complex is fast compared to its decomposition back to **E** and **S**, the reaction rate is equal to the rate of **ES** formation. This is called *diffusion-controlled limit*, which ranges between 10^5 and 10^7 mM^{-1}s^{-1}. Enzymes attained this limit are considered to be perfect from the chemical point of view. Carbonic anhydrase, that is the enzyme considered in this thesis, belongs to this 'perfect' enzymes group.

Table 1.1 The K_M, $k_{\mathbf{cat}}$ and $k_{\mathbf{cat}}/K_M$ values related to Carbonic anhydrase [32].

Isozyme	K_M (M)	$k_{\mathbf{cat}}$ (s^{-1})	$k_{\mathbf{cat}}/K_M$ ($M^{-1}s^{-1}$)
Carbonic anhydrase	5.0×10^{-6}	8.10×10^{2}	1.6×10^{8}

1.5 Solvation and dielectric effects

One of the factor to be considered in an accurate calculations in quantum chemistry is the interaction between solute and solvent. The solute properties as geometry, relative energy of isomer, vibrational frequency and electronic spectrum depend on the solvent, particularly the polar one. The common approach is to include the solvent as a continuum with a dielectric permittivity constant ε. The cavity in the continuum with the model system included can be generated using different approaches, such *van der Waals surface*, *solvent accessible surface* (SAS) or *isodensity surface*. To describe the electrostatic interaction between an arbitrary charge density $\rho(\mathbf{r})$ and a continuum dielectric it can be used the *Poison equation*:

$$\nabla^2\phi(\mathbf{r}) = \frac{-4\pi\rho(\mathbf{r})}{\varepsilon} \qquad \textbf{(1.18)}$$

where $\phi(\mathbf{r})$ is the electrostatic potential of charge density. During a quantum chemical calculation, eq.(1.18) is solved numerically, so the solvent charges distribution is then returned to the SCF procedure which

is performed accounting for the solvent charges. This process is called *self-consistent reaction field* (SCRF) method because it is repeated until self-consistency is reached.

The *Polarizable Continuum Model* (PCM) [33] is one of the most frequently used continuum solvation methods and has seen numerous variations over the years. The PCM model calculates the molecular free energy in solution as the sum over three terms:

$$G_{sol} = G_{es} + G_{dr} + G_{cav} \quad \textbf{(1.19)}$$

These components represent the electrostatic (**es**) and the dispersion-repulsion (**dr**) contributions to the free energy, and the cavitation energy (**cav**). All three terms are calculated using a cavity defined through interlocking van der Waals-spheres centered at atomic positions. The reaction field is represented through point charges located on the surface of the molecular cavity (*Apparent Surface Charge* (ASC) model).

The dielectric constant in biochemical models is an empirical parameter, and normally chosen to be equal to 4, as approximated between ε =3 for the protein itself, and ε =80 for the surrounding water. Usually solvent effects are computed upon the gas-phase optimized geometries.

2 Computational Methods

A theoretical modeling of an enzyme-catalyzed reaction mechanism can, in general, be accomplished at different levels of approximation. In this section different approaches are described, following a multi-scale guideline starting from the methods used at the basic level and continuing with some more advanced ones. The goal is to give an overview of the different techniques without going into detail, but simply to show how these can be used as a powerful tool of investigation.

2.1 Cluster Models

In the cluster model approach a model is constructed based on the crystal structure of the enzyme by selecting a discreet number of atoms from the residues that make up the active site [34-38]. This procedure allows to apply an accurate quantum chemical methods to study a reaction mechanism considering a small cluster of enzyme active sites. In addition to the intrinsic chemistry of the catalyzed reaction, a cluster model also tries to take in consideration the effect of surrounding residues on the reacting region. To compensate for the lack of steric effects, selected atoms are kept constrained at their crystallographic positions [39] during the geometry optimizations. A quantum chemical study of a reaction mechanism implies the determination of all intermediates and transition states along the reaction path. In this context cluster models are easiest to setup for systems like metalloenzymes where the chemistry at the active site is dominated by strong electrostatic interactions due to the presence of the central metal ion. The primary coordination sphere of ligating atoms that are directly attached to the metal ions must be

included in the model. The secondary coordination sphere, which includes residues that are near but not directly attached to the metallic center, in some extent, can be considered as part of the scaffold, since the properties of some of them can strongly influence the reactivity of the metal ions. To simulate the electrostatic polarization effects exerted by the protein surrounding on the cluster, dielectric cavity techniques can be used, for example the polarizable-like continuum model (PCM) that has been described in the previous chapter. The missing outer parts of the enzyme are mimicked by a homogeneous medium with a dielectric constant usually set to 4. For large cluster models, the solvation effects usually saturate, i.e. with an increasing size of the model the choice of the dielectric constant becomes less critical.

2.2 Docking

If only a structure of the substrate-free form of the enzyme is available, molecular docking can be used to position the substrate in the active site [40,41]. To predict how two molecular objects interact, the first requirement is a structure of the protein of interest, usually stored in a *x-ray* pose.

Mainly, a docking program depends strictly from two components that are responsible of the success of the simulation: *search algorithms* [42] and the *scoring function* [43].

Search algorithms is the part of the docking program that is responsible for generating different conformations of the ligand. It can be divided into three categories:

- systematic or stochastic torsional searches which give an exhaustive search of the ligand conformations according to rotatable bonds;
- Genetic Algorithms include a random factor in the generation of conformations and consequently a large number of different conformations that can be achieved with a small loss of exhaustiveness;
- Molecular Dynamics Simulation methods produce time dependent conformations that are correlated with each other.

The *scoring function* takes a pose as input and calculates an approximation to the free energy of binding of the ligand in a certain

conformation and position. Most scoring functions are physics-based molecular mechanics force fields that estimate the energy of the pose. An alternative approach is to derive a statistical potential for interactions from a large database of protein-ligand complexes, such as the *Protein Data Bank* (PDB), and evaluate the fit of the pose according to this inferred potential [44,45].

In the enzymatic catalysis, docking methodologies play a key role in finding the best placement for the substrate in the active center, when it is not present in the crystallographic structure. The strategy of docking the ligand is to choose the best calculated conformation by the scoring function, and use this result as starting point for mechanistic studies. Structures are optimized in the course of that process, but these optimizations will most probably fall on a local minimum not related with the best substrate conformation. To bridge this gap, mechanistic constraints referred from other enzymes with the substrate in the active site can be used, otherwise, when less is known about the binding, the conformations can be refined via minimizations with molecular mechanics.

2.3 **Molecular Mechanics Overview**

A simple and fast model to describe molecular systems is *Molecular Mechanics* (MM) [46,47]. The description of a system is done by the combination of a *force field* (FF) and set of parameters. In general in MM the cumulative physical forces can be used to describe molecular geometries and energies. The special conformation ultimately obtained is then a natural adjustment of geometry to minimize the total internal energy. A molecule is considered as a collection of masses centered at nuclei (atoms) connected by springs (bonds); in response to inter and intra-molecular forces, the molecule stretches, bends, and vibrates about those bonds. This simple description of a molecular system as mechanical body is usually associated with “classical” systems. So, this “classical” mechanics description works well for describing molecular structures and processes, with the exception of bond-breaking event.

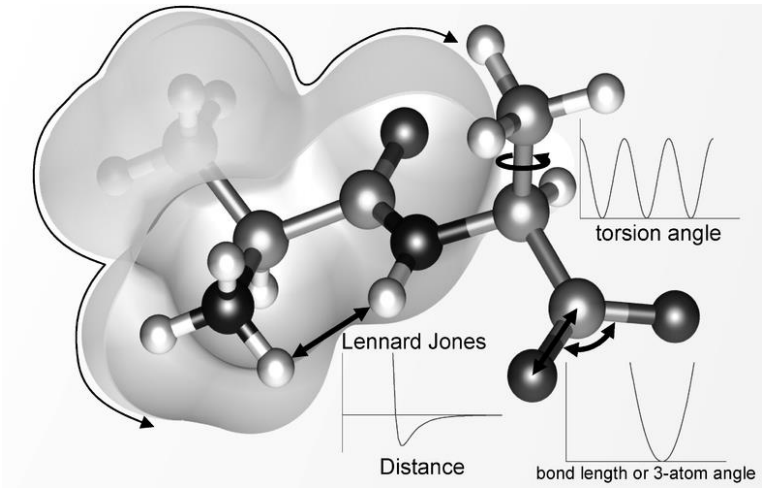


Figure 2.1 A molecule is considered a mechanical system in which particles are connected by springs, and where simple physical forces dictates its structures and dynamics [48].

2.3.1 Force Fields

The basic functional form of a FF encapsulates both bonded terms relating to atoms that are linked by covalent bonds, and nonbonded (also called "*noncovalent*") terms describing the long-range electrostatic and van der Waals forces. The specific decomposition of the terms depends on the force field, but a general form for the total energy in an additive force field can be written as

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

where the components of the covalent and non-covalent contributions are given by the following summations:

$$E_{Bonded} = E_{Bond} + E_{Angle} + E_{Dihedral} \quad \textbf{(2.2)}$$

$$E_{Non-Bonded} = E_{Electrostatic} + E_{van\ der\ Waals} \quad \textbf{(2.3)}$$

Generally the bond and angle terms are modelled as harmonic potentials centered around equilibrium bond-length values derived from experiment or theoretical calculations. For accurate reproduction of vibrational spectra, the Morse potential can be used instead, at computational cost.

The dihedral or torsional terms typically have multiple minima and thus cannot be modeled as harmonic oscillators, though their specific functional form varies with the implementation. This class of terms may include "improper" dihedral terms, which function as correction factors for out-of-plane deviations. The non-bonded terms are much more computationally costly to calculate in full, since a typical atom is bonded to only a few of its neighbours, but interacts with every other atom in the molecule. Fortunately the van der Waals term falls off rapidly – it is typically modelled using a "6 –12 *Lennard-Jones potential*" (LJ) [49], which means that attractive forces decay with distance as r^{-6} and repulsive forces as r^{-12}, where r represents the distance between two atoms. The repulsive term r^{-12} is however unphysical, because repulsion increases exponentially. Generally a cutoff radius is used to speed up the calculation so that atom pairs whose distances are greater than the cutoff have a van der Waals interaction energy of zero. The electrostatic terms are more difficult to calculate accurately because they do not decay rapidly with distance, and long-range electrostatic interactions are often important features of the system (especially for proteins). The basic functional form is the Coulomb potential, which only falls off as r^{-1}. A variety of methods are used to address this problem, the simplest being a cutoff radius similar to that used for the van der Waals terms. However, this introduces a sharp discontinuity between atoms inside and atoms outside the radius. *Switching* or *scaling functions* that modulate the apparent electrostatic energy are somewhat more accurate methods that multiply the calculated energy by a smoothly varying scaling factor from 0 to 1 at the outer and inner cutoff radii.

While improvement of potential energy function – both in terms of functional form and parameters – has been an ongoing enterprise, the current, "*second generation*" molecular mechanics and dynamics force fields are more sophisticated than those originating from pioneering works. In particular parameterization depends quite significantly now on quantum mechanical calculation. "*Third generation*" FF that account more accurately for electronic polarizabilities are already emerging [50].

2.3.2 Many FF Choices

At present, there is no universal force field able to reproduce accurately different model systems, nor are the many force fields in use close to converging to one another in some sense.
Protein modelers often use the CHARMM package with the related FF [51] developed by Martin Karplus and coworkers while nucleic acid modelers might prefer the AMBER [52] program developed by Peter Kollman. In general therefore the choice of the FF is strictly linked to the model and to what the user want to describe. It is often necessary to underline a particular property of a defined system that could be relevant in a conformational study, for example hydrogen bonds. Some macromolecular FFs use explicit hydrogen-bonding potentials, through these potentials were more common in older version that did not consider hydrogens explicitly. The strength of a hydrogen bond is determined by its geometry, which depends on the distances associated with the Donor-Hydrogen – Acceptor angle (ϑ_{dha}) formed about hydrogen atom. For linear hydrogen bonds, ϑ_{dha} has the ideal value of 180°. In the current bio-molecular AMBER and CHARMM force fields, the proper dependence of hydrogen bonding on distance and angular orientations is adequately treated with the electrostatic LJ terms. Classical electrostatic and bonding forces can account for hydrogen-bonding interactions. Thus much work for small-molecule FF.
We could conclude that there are some general limitation of current molecular FFs, in the sense that each one have deficiencies. User must choose the best FF that fits with the model in agreement with the scope of the investigation.

List of classical force fields:

- AMBER (*Assisted Model Building and Energy Refinement*) - used for proteins and DNA
- CHARMM (*Chemistry at HARvard Molecular Mechanics*) - used for both small molecules and macromolecules
- CVFF - used for small molecules and macromolecules [53]
- COSMOS-NMR - hybrid QM/MM forcefield adapted to a variety of inorganic compounds, organic compounds and biological macromolecules, including semi-empirical calculation of atomic

charges and NMR properties. COSMOS-NMR is optimized for NMR based structure elucidation and implemented in COSMOS molecular modelling package. [54]
- GROMOS - a force field that comes as part of the GROMOS (*GROningen MOlecular Simulation* package), a general-purpose molecular dynamics computer simulation package for the study of biomolecular systems. [55]
- OPLS (*Optimized Potential for Liquid Simulations*) (variations include OPLS-AA, OPLS-UA, OPLS-2001, OPLS-2005). [56]
- ENZYMIX – a general polarizable force field for modeling chemical reactions in biological molecules. This force field is implemented with the *empirical valence bond* (EVB) method and is also combined with the semimacroscopic PDLD approach in the program in the MOLARIS package.
- ECEPP/2 - first force field for polypeptide molecules - developed by F.A. Momany, H.A. Scheraga and colleagues.
- QCFF/PI – A general force field for conjugated molecules. [57,58]
- UFF-A general force field with parameters for the full periodic table. [59]

2.3.3 Molecular Dynamics

Molecular Dynamics (MD) simulations represent the computer approach to statistical mechanics. As a counterpart to experiment, MD simulations are used to estimate equilibrium and dynamic properties of complex systems that cannot be calculated analytically.
The molecular dynamics approach is simple in principle, i.e. to simulate motion of a system under the influence of a specified force field by following molecular configurations in time according to Newton‘s equation of motion. A system of N atoms can be described with a first-order differential equation:

$$\begin{aligned} \mathbf{M}\dot{V}(t) = F(X) = -\nabla E\big(X(t)\big) + \ \dots\,, \\ \dot{X}(t) = V(t) \end{aligned} \qquad \textbf{(2.4)}$$

where $X \in \boldsymbol{R}^{3N}$ denotes the collective Cartesian vector of the system; V is the corresponding collective velocity vector; $\mathbf{M}$ is the diagonal mass

matrix and the dot superscript denote differentiation with respect to time.

The force F in the right-hand-side is composed of the systematic force, which is the negative gradient of the potential energy E, additional terms that mimic the environment. Each gradient component $i, i = 1, \ldots 3N$ is given by:

$$\nabla E(X)_i = \partial E(X) / \partial \alpha_i \qquad \textbf{(2.5)}$$

where α_i denoted x, y and z component of an atom. These equation must be integrated numerically since analytic (closed-form) solution are known for the simplest systems. Such numerical integrations and velocity pairs, $\{X^n, V^n\}$, for integers n that represent discrete times $t = n\Delta t$ at intervals (timestep) Δt.

Molecular dynamics simulations generate information at the microscopic level, including atomic positions and velocities. The conversion of this microscopic information to macroscopic observables such as pressure, energy, heat capacities, etc., requires statistical mechanics. Statistical mechanics is fundamental to the study of biological systems by molecular dynamics simulation.

The thermodynamic state of a system is usually defined by a small set of parameters, for example, the temperature, T, the pressure, P, and the number of particles, N. Other thermodynamic properties may be derived from the equations of state and other fundamental thermodynamic equations. An ensemble is a collection of all possible systems which have different microscopic states but have an identical macroscopic or thermodynamic state.

There are a few choices for the type of ensemble conditions used in molecular dynamics. The most used are:

- the microcanonical ensemble (NVE), where the number of particles, the volume of the box and the energy are kept constants;
- the canonical ensemble (NVT), where the number of particles, volume of the box and temperature are kept constant;
- the isothermal isobaric ensembles (NPT), where the number of particles, the pressure and the temperature are kept constant.

2.4 MD in Enzymatic Catalysis

The application of molecular dynamics in enzymatic catalysis is, mainly, to stabilize the structure of the enzyme-substrate complex after some modeling task has taken place. Usually, to proceed with the study of the catalytic mechanism itself, one needs the structure of the enzyme in the reactants state, but the crystallographic structure is rarely found exactly at this state. When the substrate is changed or protein residues are mutated, or the substrate or part of the protein are missing, further steps are necessary. Small modifications, such as transform a mutated substrate, can be done directly by adding or substituting atoms in the structure. In these cases, a structure minimization can be sufficient to wield the correct enzyme-substrate complex. Larger modifications, such as the docking of the ligand in the active center, usually require larger movements of the structure to achieve relaxation. These could not be achieved by simple minimization, and a molecular dynamics simulation is done.

In the present thesis, MD based approaches have been used mainly to the study of the dynamics of binding – unbinding events related to biomolecular systems because calculation of free energy is of great importance for understanding the kinetics and the structural determinants of these processes. However since they require thorough sampling of configuration space, free energy calculations are extremely costly for complex systems like biomolecules and efficient calculation of free energy is one of the most challenging tasks in computer simulations.

2.5 Steered Molecular Dynamics

Steered Molecular Dynamics (SMD) induces unbinding of ligands [60,61] and conformational changes in biomolecules on time scales accessible to molecular dynamics simulations. The simulations reveal the details of molecular interactions in the course of unbinding or unfolding, hereby providing important information about the molecular mechanisms underlying these processes. The advantage of SMD over conventional MD is the possibility of inducing relatively large conformational changes in molecules on n*s* time scales accessible to actual computations.

For this purpose, time-dependent external forces are applied to a system and the responses of the system are analyzed. There are two typical protocols of SMD:

- constant velocity pulling
- constant force pulling

In general, in a SMD simulation an external forces is applied ,i.e., to a protein-ligand complex restraining the ligand to a point in space (restraint point) by an external, harmonic, potential. The restraint point is then shifted in a chosen direction forcing the ligand to move from its initial position in the protein and allowing the ligand to explore new contacts along its unbinding path [62-66]. Assuming a single reaction coordinate x, and an external potential

$$\mathrm{U} = K(x - x_0 - vt)^2/2 \qquad \textbf{(2.6)}$$

where K is the stiffness of the restraint, and x_0 is the initial position of the restraint point moving with a constant velocity, v, the external force exerted on the system can be expressed as

$$F = K(x_0 + vt - x) \qquad \textbf{(2.7)}$$

This force corresponds to a molecule being pulled by a harmonic spring of stiffness K with its end moving with velocity v.
In cases where irreversible work, done during unbinding can be attributed to a nondispersive frictional force, a quantitative description of the thermodynamic potentials governing the binding and unbinding processes can be achieved by discounting the irreversible work from the calculated *potential of mean force* (PMF). Irreversibility of the unbinding process can also be accounted for by averaging over an ensemble of SMD trajectories according to the nonequilibrium equality for free energy differences (*Jarzynski's equality*) [67,68]. Jarzynski's equality is a relation between equilibrium free energy differences and work done through nonequilibrium processes. Considering a process that changes a parameter λ of a system from λ_0 at time zero to λ_t at time t. The second law of thermodynamics states that the average work done on the system cannot be smaller than the difference between the free energies corresponding to the initial and the final values of λ:

$$\Delta F = F(\lambda_t) - F(\lambda_0) \;\leq \langle W \rangle \qquad \textbf{(2.8)}$$

where the equality holds only if the process is quasistatic [69]. According to this inequality, a nonequilibrium process provides only an upper limit for the free energy difference. Jarzynski discovered an equality that holds regardless of the speed of the process:

$$e^{-\beta\Delta F} = \langle e^{-\beta W} \rangle \quad \textbf{(2.8)}$$

where $\beta = 1/k_B T$ uncludes the boltzmann constant (k_B) and temperature (T). .Jarzynski's equality opens the possibility of calculating free energies from nonequilibrium processes. This approach as nonequilibrium thermodynamic integration is in opposition to the conventional thermodynamic integration based on quasi-static processes for which $\Delta F = \langle W \rangle$. The average of exponential work appearing in Jarzynski's equality is dominated by the trajectories corresponding to small work values that arise only rarely. An accurate estimate of free energy, hence, requires suitable sampling of such rare trajectories. Therefore, although Jarzynski's equality holds for processes of any speed, practical applications are currently limited to slow processes for which the fluctuation of work is comparable to the temperature.

2.5.1 Reversible pulling

Some processes, like some folding – unfolding or binding – unbinding events with competitive substrates, can be simulated applying the same pulling speed in both directions. If a pulling simulation is performed very slowly, then the process is reversible. The work done during such a reversible pulling is equal to the free energy difference of the system between initial and final states. Thus, a reversible work curve can be considered as an exact PMF [70]. Therefore, eq.(2.8) becomes an equality:

$$F(\lambda_t) - F(\lambda_0) = \langle W_{0\to t} \rangle \quad \textbf{(2.9)}$$

2.5.2 Irreversible pulling

In studying large systems like biomolecules, the time scale accessible to computer simulation is often much shorter than the natural time scale of the process of interest. Therefore, such a process needs to be accelerated

in simulations; in addition, only a small number of trajectories, can be obtained.
After the discovery of Jarzynski's equality, the second order cumulant expansion formula [71,72], identical to the near-equilibrium one, was recognized as an approximation to the exponential average.

Thus, employing Jarzynski's equality eq.(2.8):

$$e^{-\beta\{F(\lambda_t)-F(\lambda_0)\}} = \langle e^{-\beta W}\rangle \quad \textbf{(2.10)}$$

$$F(\lambda_t) - F(\lambda_0) = -1/\beta \log\langle e^{-\beta W}\rangle \quad \textbf{(2.11)}$$

The right-hand side can be expanded in terms of cumulants

$$F(\lambda_t) - F(\lambda_0) = \langle W\rangle - \frac{1}{2T}[\langle W^2\rangle - \langle W\rangle^2] + \ \ldots \quad \textbf{(2.12)}$$

If the distribution of work W is Gaussian, third and higher cumulants are identically zero[73].

Two kinds of error are involved using these approximate formulas:

- a systematic error due to the truncation of higher order terms
- a statistical error due to insufficient sampling.

If an infinite number of trajectories were available, the statistical error would vanish and hence the exponential average eq.(2.11) would give the best PMF estimation. However, since in practice only a limited number of trajectories are sampled, the statistical error may dominate the systematic error. Thus, the approximate formulas may give better results since lower order cumulants are estimated with smaller statistical errors.

2.6 **Metadynamics**

Metadynamics belongs to a class of methods in which sampling is facilitated by the introduction of an additional bias potential (or force) that acts on a selected number of degrees of freedom, named, collective variables (CVs) [74]. The dynamics is driven by the free energy and is biased by a history-dependent potential constructed as a sum of Gaussians deposited along the system from revisiting configurations that have already been sampled. At time t, the metadynamics bias potential can be written as [75]

$$V_G(S,t) = \int_0^t dt' \omega \exp\left(-\sum_{i=1}^{d} \frac{(S_i(R) - S_i(R(t')))^2}{2\sigma_i^2}\right) \quad \textbf{(2.13)}$$

where S is a set of d functions of the microscopic coordinates R of the system; ω is an energy rate and σ_i is the width of the Gaussian for the ith CV. The energy rate is constant and usually expressed in terms of a Gaussian height W and a deposition stride (t_G) can be defied as

$$\omega = \frac{W}{t_G} \quad \textbf{(2.14)}$$

This potential, in time, fills the minima in the free energy surface, i.e. the sum of the Gaussians and of the free energy becomes approximately a constant as a function of the collective variables.

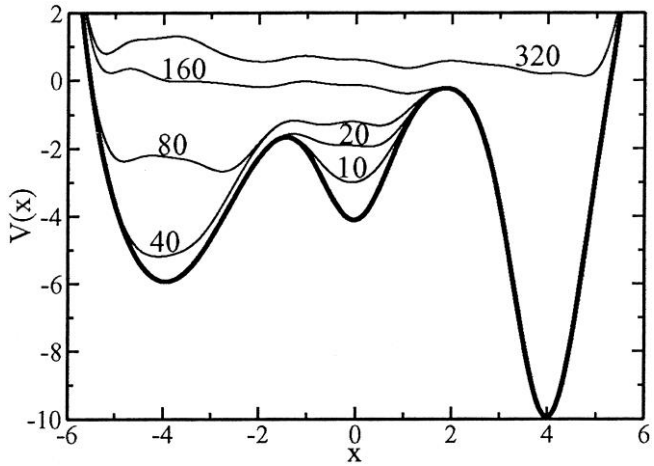


Figure 2.2 Schematic representation of the progressive filling of the underlying potential (thick line) by means of the Gaussians deposited along the trajectory. The sum of the underlying potential and of the metadynamics bias is shown at different times (thin lines) [76].

2.6.1 Funnel Metadynamics

Parrinello and co-workers recently developed a metadynamics-based approach, named *funnel-metadynamics* (FM), which leads to an efficient characterization of the binding free-energy surface and an accurate calculation of the absolute protein–ligand binding free energy [77]. In FM, a funnel-shaped restraint potential is applied to the system, reducing the space to explore in the unbound state. In such a way, the sampling of ligand-bound and-unbound states is highly enhanced, thus leading to an

accurate estimation of the binding *Free Energy Surface* (FES) within a reasonable simulation time. This potential is a combination of a cone restraint, which includes the binding site, and a cylindric part, which is directed toward the solvent. Using the funnel potential during the simulation, the system does not feel any repulsive bias when the ligand explores regions inside the funnel area. As the ligand reaches the edge of the funnel, a repulsive bias is applied to the system, disfavoring it from visiting regions outside the funnel. So the protein-ligand binding free energy is expressed by

$$\Delta G_b^0 = \Delta G - \frac{1}{\beta} ln\left(\pi R_{cyl}^2 C^0\right) \quad \textbf{(2.15)}$$

where ΔG is the free-energy difference between bound and unbound states and πR_{cyl}^2 is the surface of the cylinder used as restraint potential [77].

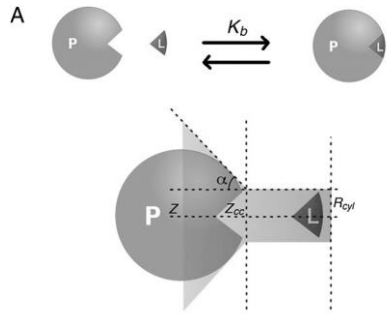


Figure 2.3 Schematic representation of the ligand/protein binding process and the funnel restraint potential. z is the axis defining the exit-binding path of the ligand, z_{cc} is the distance where the restraint potential switches from a cone shape into a cylinder [77].

2.7 QM/MM and ONIOM

The most obvious limitation of cluster approach is to neglect the long distance interactions and the role of the enzymatic conformational changes, during a reaction. In section 2.1 has been explained how to take in account the protein strain with a small model, i.e, keeping fixed some atoms in their crystallographic position. With a full *Quantum Mechanics* (QM) approach it is impossible to treat systems with hundreds of thousands of atoms. On the other side MM is only good for near minima structures because this method can't simulate formation or breaking of

bonds. *Quantum Mechanics/Molecular Mechanics* (QM/MM) [78] hybrid methods have been designed to use the advantages of both levels of theory: the region when the chemistry occurs is described with electronic method (QM) and the rest is treated usually with MM.
One of the popular hybrid QM/MM methods is *Our own N-layered Integrated molecular Orbital and molecular Mechanics* (ONIOM) [79-82]. Because in this thesis the two layered ONIOM was mainly used, few notation should be introduced to explain the technic: ONIOM extrapolates the total energy of the system E into $E_{\mathbf{X},\mathbf{Y}}$ contributions, where **X** is the theory level used, and **Y** is the corresponding part of the system:

$$E_{\mathrm{ONIOM(2-layers)}} = E_{\mathrm{High,Model}} + E_{\mathrm{Low,Complete}} - E_{\mathrm{Low,Model}} \qquad \textbf{(2.16)}$$

To calculate the interaction between the two layers, there are two options available:

- mechanical embedding
- electrostatic embedding

With the mechanical embedding, both *High,Model* and *Low,Model* calculations are done in vacuum. From that, it results that the interaction between the *High* layer and the *Low* layer is considered at the MM level. The interaction is present in the MM calculation of the complete system and is left unaltered. However with using mechanical embedding the QM electronic density is not polarized by the enzymatic environment. This means that the rest of the enzyme is affected by the active center, but the active center is only affected by positional constraints of the environment.
With the electrostatic embedding the calculation of the high layer is not done in vacuum but in the electric field generated by the enzyme. The effect is achieved by including the molecular mechanics point charges of the low layer in the high layer Hamiltonian, as an external potential.
However, the *Low* level method should still be able to describe the atoms in the *Model* part. Usually this is an obstacle in working with metalloproteins, because most of the FFs or semiempirical methods are not parametrized for transition metal atoms.

In our study of carbonic anhydrase, the two-layered ONIOM (B3LYP:UFF) method was used because UFF allows to skip the problem related to parametrization of the metal center.

3 Carbonic anhydrase and enzyme promiscuity

The first zinc enzyme to be discovered was carbonic anhydrase (CA) [83], which catalyzes the hydration of CO_2. The catalytic power of carbonic anhydrase is not limited to CO_2 hydration, as a matter of fact it plays a key role in the relative new scenario of enzyme promiscuity.

3.1 Carbonic anhydrase

Carbonic anhydrase (CA, EC 4.2.1.1) is a zinc-containing enzyme that catalyzes the interconversion of carbon dioxide (CO_2) in carbonic acid (H_2CO_3).

$$CO_2 + H_2O \rightleftharpoons HCO_3^- + H^+ \quad \textbf{(3.1)}$$

Carbonic anhydrase has essential roles in facilitating the transport of carbon dioxide and protons in the intracellular space, across biological membranes and in the layers of the extracellular space.
The CAs are ubiquitous throughout nature and are expressed in eukaryotes eubacteria, and archaea, and are divided into at least three classes based on amino acid homology [84,85]:

- α – CAs from animals (all mammalian CAs), plants, eubacteria, and viruses;
- β – CAs from plants, bacteria, and animals (e.g., C. elegans);
- γ – CAs from bacteria and plants.

Recently, were discovered other two classes of CA:

- δ – CAs TWCA1 from a marine diatom, Thalassiosira weissflogii, which is possibly a homolog of α –CAs[86,87]
- ε – CAs from bacteria, which probably represent a subclass of β - CA that has diverged such that only one of its two domains has retained a viable active site[88,89].

The α – CAs are the most thoroughly studied and the only class present in humans. There are at least 16 members of the α – CA family: CA I, II, III, IV, VA, VB, VI, VII, VIII, IX, X, XI, XII, XIII, XIV, and XV that have wide-ranging cellular localizations [90,91].
In this thesis we mainly have dealt of the hCAII because it is a monomeric, single-chain protein of intermediate molecular weight (~30 kDa) and is extremely well-characterized biophysically. There is, farther, a wealth of structural information available which makes them particularly well suited for biophysical studies. The active site of α – CAs comprises a catalytic Zn^{2+} ion coordinated by three imidazole groups of histidines and by one hydroxide ion (or water molecule), all in a distorted tetrahedral geometry. This grouping is located at the base of a cone – shaped amphiphilic depression, one wall of which is dominated by hydrophobic residues and the other of which is dominated by hydrophilic residues.
Generally CA inhibition is important in the treatment of glaucoma, altitude sickness, and obesity; its overexpression has recently been implicated in tumor growth. Its inhibition in pathogenic organisms might lead to further interesting drugs [92,93]. More than its medical relevance, its tractability and simplicity are what make CA a particularly attractive model enzyme.

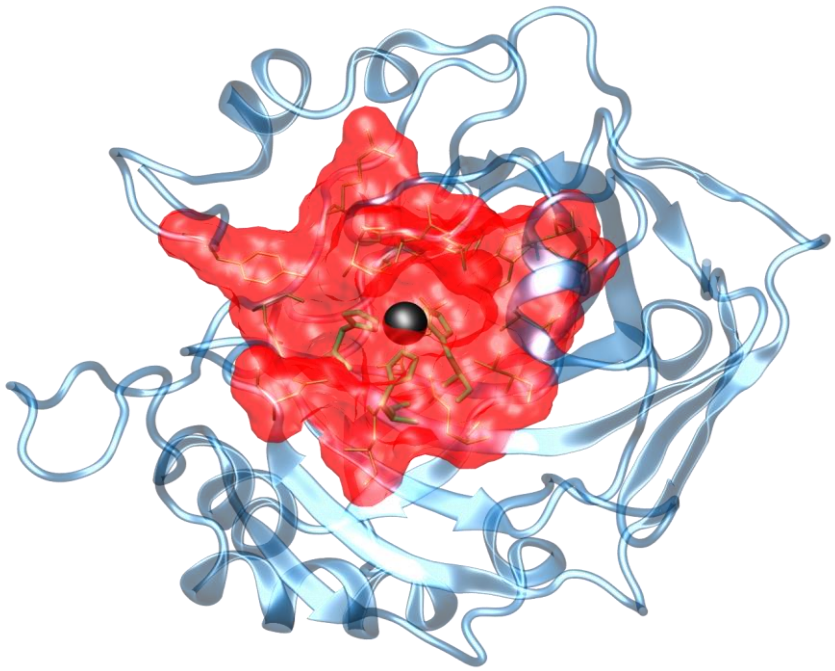


Figure 3.1 Human carbonic anhydrase II with the zinc metal core and three coordinated histidines. The red surface shows the binding cavity.

3.1.1 Catalytic mechanism

All isoforms of CA have a conical cavity at the active site which is roughly 15 Å in diameter at its mouth and 15 Å deep [94,95]. In the *α* – CA family the catalytic Zn^{2+} ion, lies at the apex of the cone near the center of the protein and is coordinated in a distorted tetrahedral arrangement to a water molecule and three histidine residues, His94, His96, and His119.
The binding cavity of CA has *hydrophobic* and *hydrophilic* faces. In particular, the *hydrophilic* face, that was characterized first by Eriksson *et al.* [90], is crucial in the understanding of the catalytic mechanism. Of the eight residues present on this face, Thr199 and Thr200 are nearest to the entrance to the cavity, while His64 is located on the opposite side of this entrance. The other five active site residues, Tyr7, Asn62, Asn67, Gln92, and Glu106, are involved in an intricate network of hydrogen bonds with nine ordered water molecules in the active site [96,97].

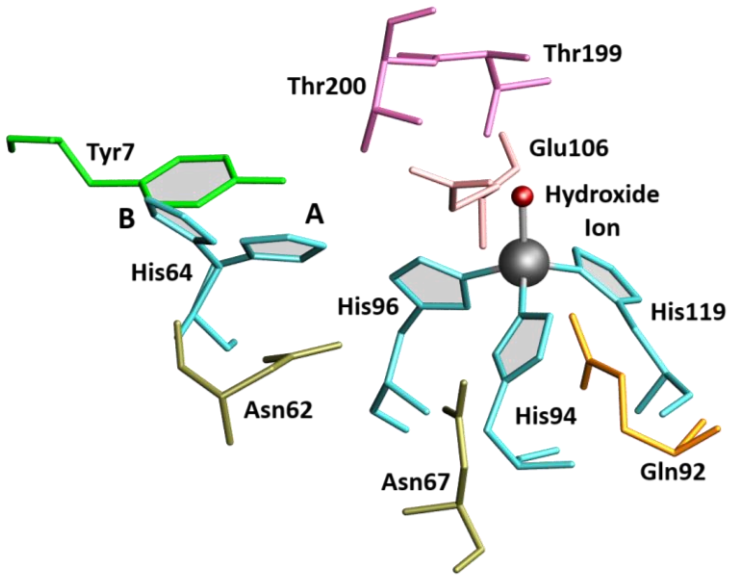


Figure 3.2 The active site region of hCAII generated from PDB file 2CBA [102]. His64, that make up the proton shuttle that regenerates the zinc-bound hydroxyl ion via deprotonation of the zinc-bound water molecule, is represented in both the possible 'in" and "out" positions (**A** – **B**).

Available evidences suggest that all members of the α – CA family share the same ping-pong mechanism for catalysis of the hydration of CO_2.
The currently accepted steps are [98,99]:

- attack on CO_2 by the Zn^{2+} – bound hydroxide to form bicarbonate
- binding of a water molecule to the Zn^{2+} at a position adjacent to the bicarbonate ion
- leaving of bicarbonate
- transfer of a proton from the newly bound water to the buffer

For the Zn-HCO_3^- species stabilization, two different approaches were proposed: the Lipskomb mechanism [100] that hypothesizes a proton transfer between the zinc bound oxygen and another oxygen of the HCO_3^- group and the Lindskog one [101], that suggests an internal rotation of HCO_3^- group that causes the change of the oxygen atom directly linked to the Zn ion. A more detailed description of the mechanism is reported in **Paper I**.

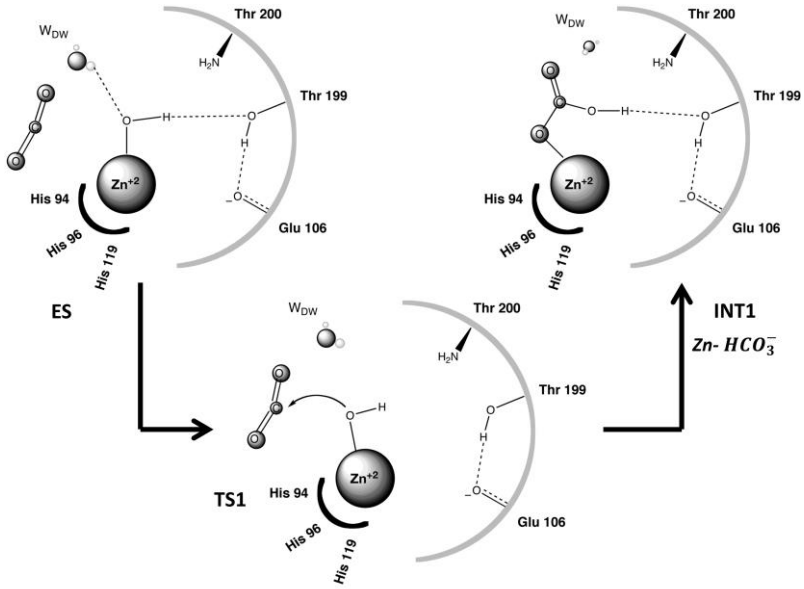


Figure 3.2 Nucleophilic attack on the CO_2 substrate.

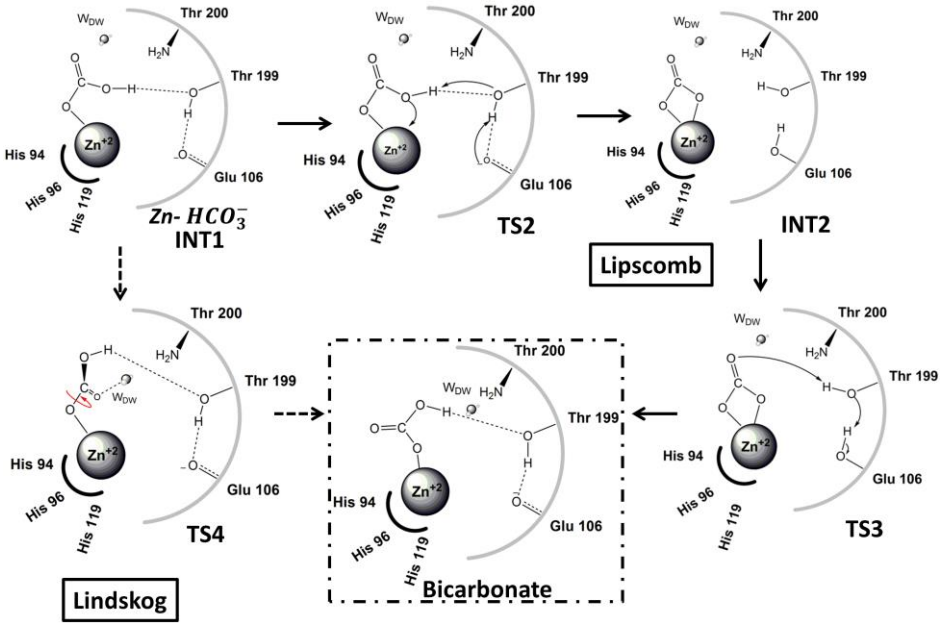


Figure 3.3 Lipscomb and Lindskog mechanisms in the case of CO_2 substrate.

3.2 **Enzyme promiscuity**

Despite their great proficiency of action and robustness, enzymes exhibit a remarkable evolutionary adaptability. In fact, it is known that new enzymatic functions can evolve in a matter of few decades, or even months, as with enzymes that degrade synthetic chemicals that first appeared on this planet during the 20th century, and the alarming evolution of drug resistance, in agreement with oft-forgotten essence of Darwinian processes [102,103]. The hypothesis of a broad specificity, or promiscuous functions, of existing enzymes was first formalized by Jensen in the 1974 [104,105] which proposed that primitive enzymes possessed very broad specificities, in contrast to modern ones. This catalytic versatility enabled fewer enzymes to perform the multitude of functions that was necessary to maintain ancestral organisms. Afterwards O'Brien and Herschlag [106], and later Copley [107], highlighted the importance of promiscuity, and surveyed the mechanistic and structural works in this area.

Promiscuity comes in many different forms. Recently, Hult and Berglund introduced the classification of promiscuity into *condition* promiscuity, *substrate* promiscuity and *catalytic* promiscuity[108]:

- *Enzyme condition promiscuity* shown by enzymes with catalytic activity in various reaction conditions different from their natural ones, such as anhydrous media, extreme temperature or pH.
- *Enzyme substrate promiscuity* shown by enzymes with relaxed or broad substrate specificity.
- *Enzyme catalytic promiscuity* shown by enzymes catalyzing distinctly different chemical transformations with different transition states. It can be either *accidental* (a side reaction catalyzed by the wild-type enzyme) or *induced* (a new reaction established by one or several mutations rerouting the reaction catalyzed by the wild-type enzyme).

This classification is sometimes not straightforward, because it is possible in some special cases, to see different types of promiscuity combined.

3.2.1 Mechanistic aspects

Catalytic machinery and same active site show exquisite specificity with respect to the native substrate but still promiscuously catalyze other, often completely unrelated, reactions. A comprehensive description can be found fixing some aspects related to this specificity-promiscuity dichotomy [109]:

- specificity[1] and promiscuity often coincide within the same active site. Thus promiscuity is linked to diverse conformations whereby the native and the promiscuous functions are mediated by different active site configurations.
- in many cases, promiscuous activities share the same active-site configuration and main active-site features with the native activity. Often comparison between both natural and promiscuous activities revealed that, although the substrates bind in a similar mode and with a similar catalytic mechanism, the natural interaction is much more favourable respect to the promiscuous substrates.
- the same catalytic residue can act in a different protonation state in the native compared to the promiscuous function. In this case the mechanism of catalysis depends on the value of the dissociation constant, that works as 'switch'.
- sometimes although both the original and promiscuous activities reside within the same active site and rely on its major feature other key parts of the catalytic machinery differ, for example, the coordination of natural and promiscuous substrate, in the same catalytic site, can be mediated by different residues.
- the cofactor ambiguity is a particular case in which, changes in enzyme specificity can also be induced by metal substitutions. In carbonic anhydrase, substitution of the native Zn^{2+} by Mn^{2+} enabled the catalysis of styrene epoxidation [110] and rhodium-substituted carbonic anhydrase acts as a hydrogen-utilizing reductase [111].

[1] Multispecific or broad-specificity enzyme perform the same reaction on a whole range of substrates, usually with similar efficiency.

- although the native substrate may interact directly with active-site residues, accidental hydrogen bonds mediated by water molecules may play a role in promiscuous interactions. Water molecules can buffer opposing dipoles or charges between the substrate and active-site residues, or they can act as acid, base, or nucleophile in the catalysis of promiscuous reactions. Indeed, spatially defined active-site water molecules have catalytic power that are comparable to amino acid residues, and localized water molecules may have played a key role in primordial enzymatic active sites.
- differences between the efficiency of promiscuous and native activities can be manifested in differences in either k_{cat} or K_M values; although it is expected that promiscuous substrates that bind weakly will exhibit high K_M values, many promiscuous substrates are characterized by low k_{cat} values. Thus, specificity may result not only from substrate binding interactions, but also from appropriate positioning relative to the catalytic machinery

3.3 **Promiscuity of carbonic anhydrase**

Carbonic anhydrase shows a broad promiscuous activity. It is involved in a series of different catalytic processes that can be enclosed in the conventional classification made by Hult and Berglund [108]. In fact both *substrate* and *catalytic* type of promiscuous activity have been found in the hCA II [112]. A case is the promiscuous esterase activity of hCAII, which probably arise from the mechanistic similarity between hydration of the carbonyl of CO_2 and that of an ester, i.e. nucleophilic attack by a zinc coordinated OH^- ion and the stabilization of the resulting oxyanionic intermediate [113,114].
Other unphysiological reactions related to the promiscuous activity of CA are aldehyde hydration[115] and in general hydrolysisreaction of diverse halogeno derivatives such as 2,4-dinitrofluorobenzene [116], benzyloxycarbonyl-chloride [117] or sulfonyl chlorides [118].
This broad variety of different substrates-CA interactions (*substrate promiscuity*) helped the design of potent inhibitors, some of which possess important clinical applications (enzyme inhibitors as antiglaucoma drugs). Most of the inhibitors, often small anions as cyanate, cyanide and

thiocyanate, act as 'metal poison' [119] but, recently, small inhibitors that do not interact with the zinc ion have been identified [120,121]. The case of isoelectronic substrates of CO_2 as cyanic acid and carbodiimide will described in detail in the next section and in the attached papers. (**Paper I and II**)

A case of *catalytic promiscuity* involving carbonic anhydrase was cited in the previous paragraph: metal substitution of the native Zn^{2+} by Rh^{+} carbonic anhydrase acts as a hydrogen-utilizing reductase. In fact, despite its strong catalytic ability, CA is not able to reduce organic molecules and can't work as hydrogenase. Rhodium – substituted CA can be defined as the first cofactor-independent reductase that reduces organic molecules using hydrogen [111]. This possibility suggested the exploitation of the natural affinity of CA for CO_2 in order to define a new hydrogenation process of CO_2 to formic acid opening an innovative perspective for an efficient chemical reduction of carbon dioxide.

4 Results

The object of this chapter is to guide the reader through the carbonic anhydrase's chemistry focusing on its promiscuous activity. The first two cases involve a comparison between the natural process of CA and two isoelectronic substrates of carbon dioxide (cyanamide - *Paper I* **and cyanic acid -** *Paper II***). Afterwards will be described the activity of carbonic anhydrase as degrading agent of carbonyl sulfide (COS-hydrolase -** *Paper III***) and the rhodium-substituted carbonic anhydrase as a hydrogen-utilizing reductase agent (***Paper IV***).**
SMD procedure were used to determine the the potential energy related to the release processes of the two reaction products: bicarbonate and ureate and will be treated in detail in *Paper V***.**

4.1 **Carbodiimide Hydration** [Paper I]

Carbodiimide and its less stable tautomer cyanamide, isoelectronic with carbon dioxide, are noncompetitive inhibitors toward the hydrolase activity of hCAII but, at the same time, can be catalyzed by hCAII, giving rise to the urea product. In this work, a QM cluster model and QM/MM approaches were used to investigate the hydration of carbodiimide catalysed by hCAII in comparison with the native substrate CO_2.
For both QM and QM/MM calculations the model was built considering the amino acid residues directly and indirectly involved in the catalytic mechanism: it includes the Zn^{2+} cation with its four primary ligands of a water molecule and three histidine residues (His94, His96, and His119) and the hydrophilic binding cavity formed by Glu106, Thr199, and Thr200 residues belonging to the outer coordination sphere. An additional water molecule, called *deep water* (**W**$_{DW}$), was incorporated into the cluster. In the QM cluster, all the amino acid residues were truncated

to reduce the size of the model, and an atom for each residue was fixed during the geometry optimization meanwhile in the QM layer of QM/MM approach, the selected residues were considered without truncation. The remaining portion of the enzyme, along with 219 explicit water molecules, was described at the MM level (UFF). Geometry optimizations were performed using the hybrid density functional B3LYP and in QM calculations the effects of the protein surroundings were mediated by using PCM.

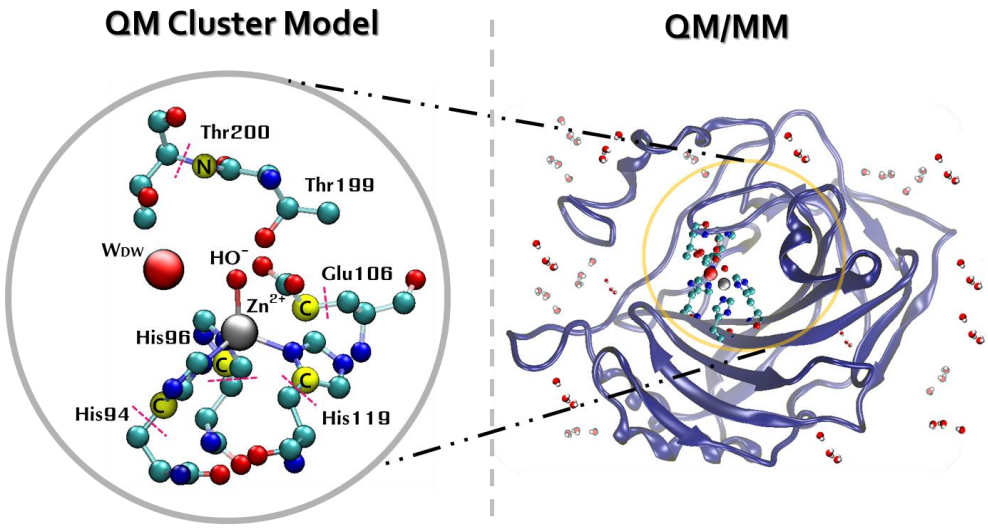


Figure 4.1 QM cluster and QM/MM models; the red dashed lines indicate the truncation of the residues in the QM approach.

4.1.1 Catalytic mechanism

The hydroxyl group (nucleophile species), coordinated to the zinc ion, attacks the carbodiimide substrate, forming an intermediate that subsequently suffers a proton shift at the expense of the **WDW** that gives rise to the ureate anion as final product. The obtained Michaelis–Menten complex (**ES**) in both QM and QM/MM structures is characterized by a coordination of the substrate with zinc ion by one of the nitrogen atoms, contrary to what is observed in the natural process.

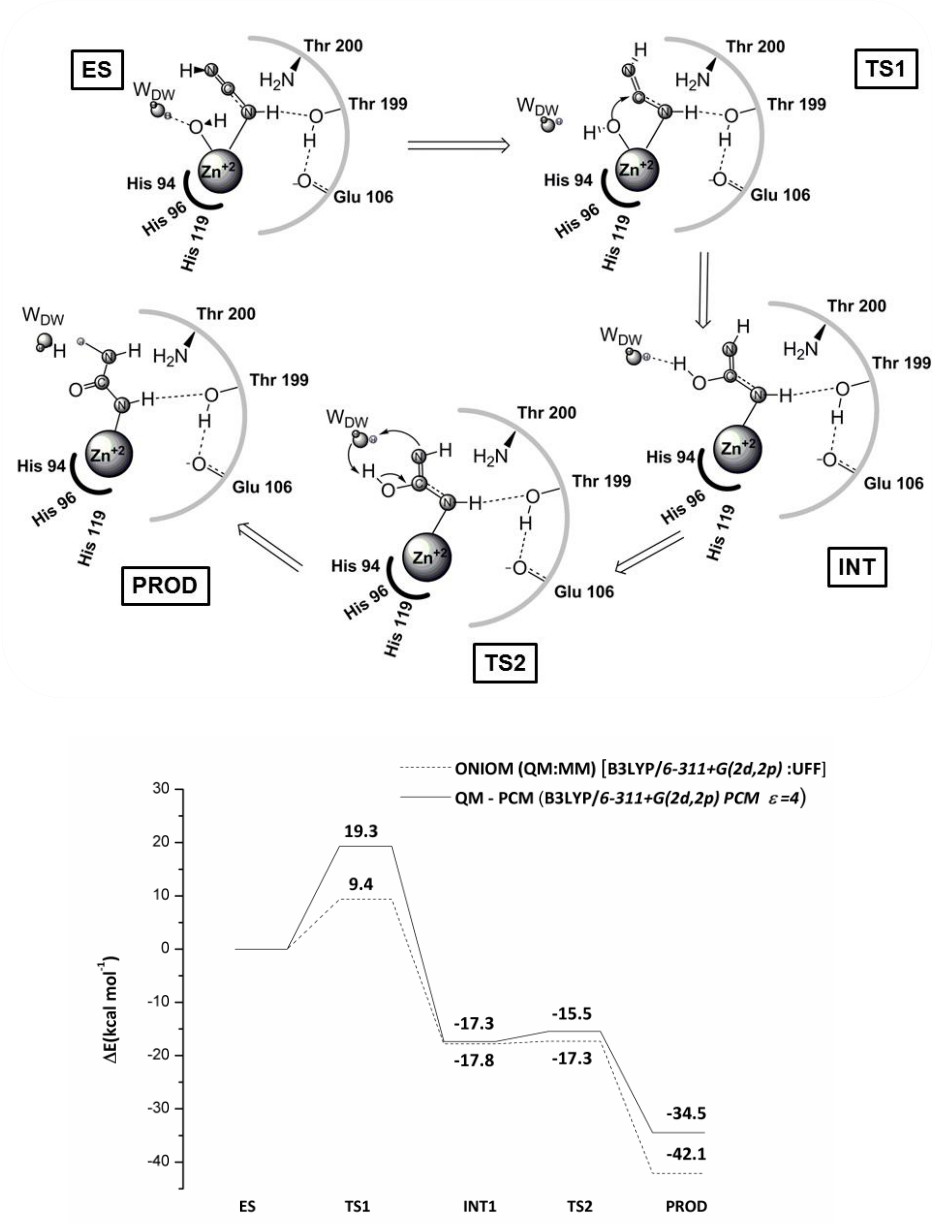


Figure 4.2 Proposed reaction mechanism for addition of water to the carbodiimide and the related PES at QM and QM/MM levels.

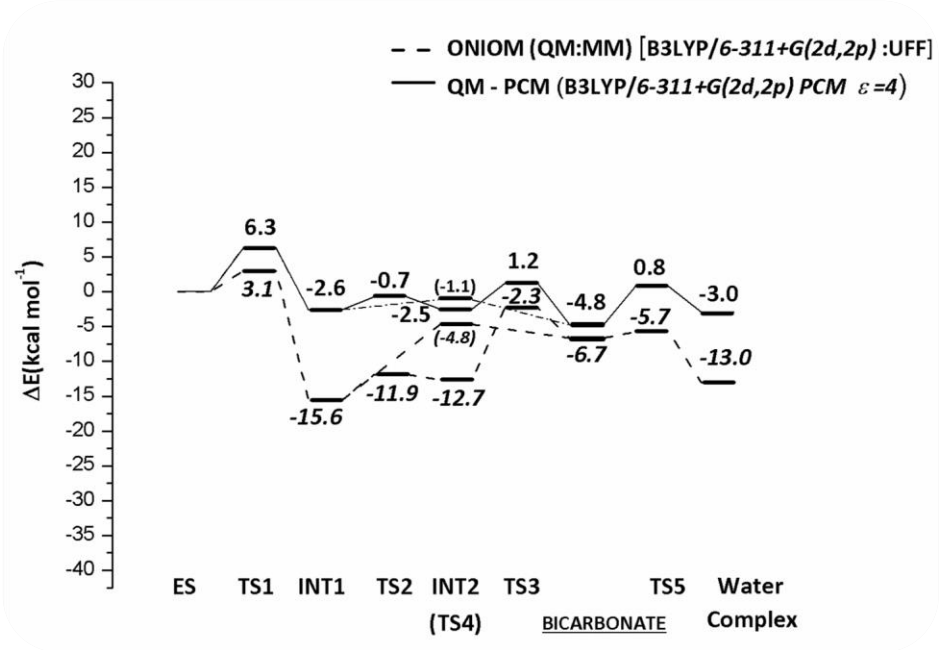


Figure 4.3 PES for the hydrolysis of CO_2 by hCAII at QM and QM/MM levels.

The calculated barrier value for **TS1** related to the carbodiimide hydrolysis (19.3 and 9.4 kcal/mol at QM and QM/MM level respectively) shows that hCAII is able to hydrolyze the substrate with acceptable kinetics for an enzyme-catalyzed reaction. Referring to CO_2 is possible to extract from the measured available kinetic data the corresponding activation barrier which is 4.02 kcal/mol. If we consider the corresponding barrier computed at QM and QM/MM levels (6.3 and 3.1 kcal/mol, respectively), we can conclude that the QM/MM value is closer to the experimental data. In any case the obtained value for QM barrier gives deviations of about 2 kcal/mol, which falls in the range of error previously established for cluster calculations[122].

4.1.2 Benchmark of functionals

In order to verify if the used exchange-correlation functional can quantitatively affect the PES, we have considered other exchange-correlation functionals including that in which the long range effects are

taken into account. Results, reported in **Table 4.1**, show that the barriers for the nucleophilic attack are almost the same with the lowest value obtained at MPWB1K level (17.8 kcal/mol). The reaction energy results to be again sensibly exothermic with the value obtained with MPWB1K closer to the QM/MM value (-38.2 kcal/mol versus -42.1 kcal/mol, respectively).

Table 4.1 The evaluation of the energetic profiles at QM level in protein environment for both substrates, has been done by single points calculations performed on the optimized B3LYP geometries using a larger basis set 6-311+G(2d,2p) and testing other three exchange-correlation functionals.

Carbodiimide	Carbon Dioxide	B3LYP		MPWB1K [25-28]		B97D [24]		M062X [29]	
TS1	[TS1]	**19.3**	[6.3]	**17.8**	[5.5]	**18.9**	[5.3]	**19.1**	[4.2]
INT	[INT1]	**-17.3**	[-2.6]	**-20.9**	[-10.2]	**-10.0**	[0.4]	**-14.3**	[-10.5]
TS2	[TS2]	**-15.5**	[-0.7]	**-17.9**	[-7.6]	**-9.1**	[1.0]	**-12.7**	[-11.1]
PROD	[INT2]	**-34.5**	[-2.5]	**-38.2**	[-10.3]	**-27.0**	[0.8]	**-30.1**	[-10.9]
	[TS3]		[1.2]		[5.5]		[3.6]		[-8.4]
	[Carbonate]		[-4.8]		[-11.2]		[-0.3]		[-10.1]
	[TS5]		[0.8]		[-7.7]		[5.7]		[-7.4]
	[Water – Complex]		[-3.0]		[-10.9]		[2.7]		[-10.9]
	[TS4 (Lindskog)]		[-1.1]		[-9.2]		[-3.2]		[-8.1]

To further verify the reliability of the B3LYP functional, we redid the computation by using the PBE0 exchange-correlation one [123,124]. Results for both of the considered substrates give barriers of 4.5 and 17.7 kcal/mol for CO_2 and carbodiimide, respectively.
An additional comparison was made between PCM and solvation method density (SMD) [125]: 19.3 versus 20.8 kcal/mol at PCM and SMD levels, respectively.

4.1.3 Restoring the enzymatic turnover

In the 4-coordinated bicarbonate complex at both QM and (QM/MM) levels the $\mathbf{W_{DW}}$ attacks the zinc ion allowing the formation of the pentacoordinated **Water Complex** throughout the **TS5** located at 0.8 (−5.7) kcal/mol relative to the reactant; thus, the reaction can proceed releasing the bicarbonate and consequently restoring the catalytic cycle. A different behavior was observed in the carbodiimide. All the attempts to localize a water−ureate final complex failed at both the QM and QM/MM

levels. A potential energy scan on the Zn–O_w bond reveals that as the water approaches the Zn^{2+}, the His119 is detached by the zinc coordination sphere. Even if the Zn–O_w distance is 2.203 Å, similar to that found in the case of carbonate, the Zn–NHis119 one becomes 2.512 Å. In the case of the pentacoordinated complex with bicarbonate, the corresponding Zn-NHis119 distance is 2.043 (2.075) Å at QM (QM/MM) levels.

So it is possible to conclude that both levels of theory evidence a great thermodynamic stability of ureate product in the carbodiimide and the substitution with a new water molecule is indispensable for the restoring of the catalytic cycle. For the case of ureate, determination of the pentacoordinated complex, that is the crucial intermediate for the coordination of water and the release of the reaction product, falls. This evidence is an important step in the energetic behavior of the catalysed reaction involving carbodiimide and more specifically in the restoring of the catalytic cycle.

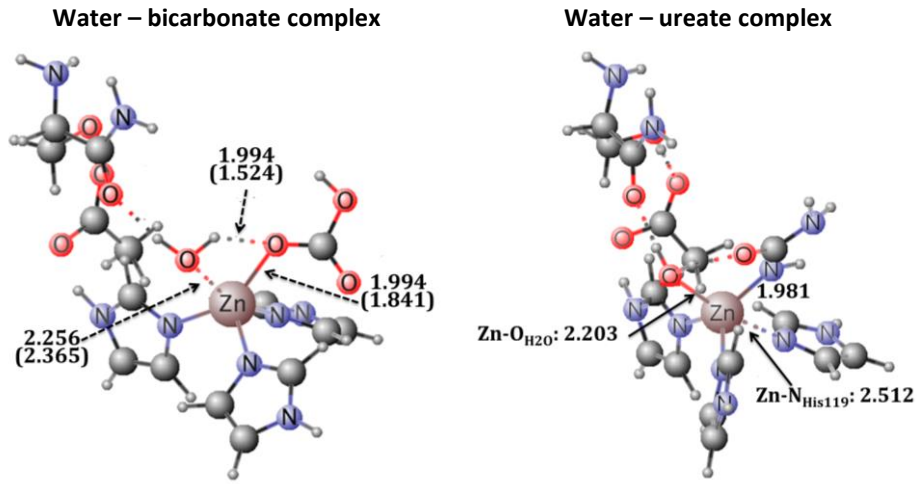


Figure 4.4 Optimized structures of water – bicarbonate complex and water – ureate complex at QM level. The distances in parentheses are referred to QM/MM obtained geometry (distances in Å). The species Water – ureate complex has been not optimized with QM/MM approach.

4.2 **Cyanic acid hydration** [Paper II]

A combined QM and QM/MM level of theory were used to investigate the promiscuous activity of human carbonic anhydrase (hCAII) against a non-physiological cyanic acid substrate. Results show that hCAII is able to hydrate the cyanic acid by a reaction mechanism similar to that of the native substrate (CO_2). Because no crystallographic data related to the cyanic acid – hCAII interaction were available, the QM/MM and QM model has been obtained performing docking simulation. Two different substrate approaches to the enzyme active site, namely carbodiimide (approach1) and CO_2-like (approach 2), have been considered, which are strictly inspired to the geometrical features regards substrates – enzyme interaction computed in the carbodiimide investigation (**Figure 4.5**).

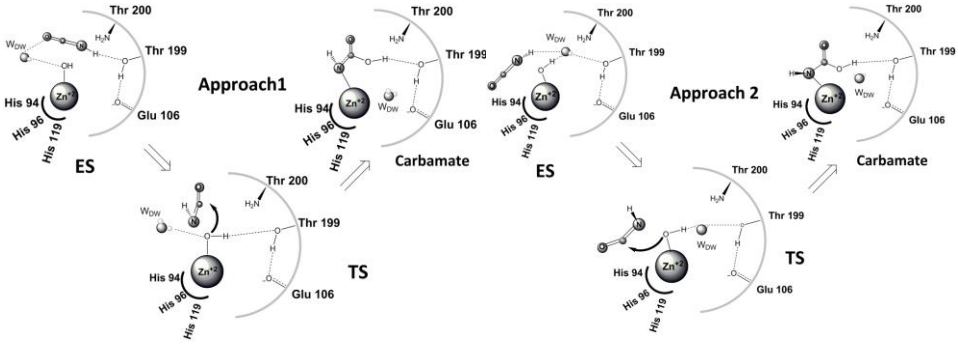


Figure 4.5 Proposed mechanisms for carbodiimide-like (1) and CO_2-like (2) approaches.

The energy barrier for the nucleophilic attack is found to be 15.6 and 4.3 kcal/mol at QM and QM/MM levels, respectively concluding that the cyanic acid prefers approaching the catalytic center in a fashion similar to that of CO_2 molecule - CO_2-like (2).
The release of the product at the expense of water for restoring the catalytic cycle seems to be the critical step with respect to the catalysed natural substrate reaction. A detailed description of the entire investigation can be found in **Paper II**.

4.2.1 Cluster dimension

As previously mentioned we used both QM and QM/MM approaches to simulate the hydration of the cyanic acid considering the same residues of the carbodiimide and CO_2 investigation. So in the QM/MM computation the QM region has been completed retaining the whole and chain for every considered amino acid and the remaining protein with 199 explicit water molecules have been described at MM level. Something different has been done for the QM cluster, taking into account three QM clusters characterized by different sizes:

- 59 atoms (the same cluster used in carbodiimide investigation)
- 83 atoms (considering Thr199 and Thr200 entirely)
- 102 atoms (including Leu 198 in the previous cluster)

Choosing the right cluster size is one of the key point in the metalloenzymes simulation because there is no direct correlation between model size and accuracy of the computations. However metalloenzymes have an advantage in that the metal ions function as anchors in the model, keeping at least the first shell residues in reasonable positions and orientations even without locking. That is related to the fact that increasing the model size has a direct impact in the global flexibility. On the other hand, as previously pointed, the electrostatic contribution of the missing enzyme surroundings can be modeled using an implicit description. In recent years, careful systematic studies using increasing model sizes of the active sites have shown that the solvation effects of the surrounding enzyme environment decrease surprisingly rapidly and, for most cases tested, almost vanish at a size of around 200 atoms [126-130].

Considering the CO_2-like (2) approach, a comparison between the three clusters and the full QM/MM model shows how the cluster size doesn't seem to play a significant role, moving the focus on the protein environment, which is explicitly considered in the ONIOM description.

Table 4.2 Calculated barriers and Reaction Energies (kcal/mol) referred to the B3LYP/6-311+G(2d,2p) zero point corrected energies including solvent effects (PCM, ε=4).

QM – 59 atoms		QM – 83 atoms		QM – 102 atoms		QM/MM	
E_a	ΔE	E_a	ΔE	E_a	ΔE	E_a	ΔE
14.9	-11.4	15.6	-7.9	13.7	-10.3	4.3	-11.2

The cluster composed by 59 atoms gave a barrier of 14.9 kcal/mol, while the larger one with 102 atoms generated 13.7 kcal/mol. A comparison with QM and QM/MM PESs previously obtained for native CO_2 and promiscuous carbodiimide substrates, shows as also in these cases, the barrier obtained at QM level is about twice greater than the corresponding QM/MM value.

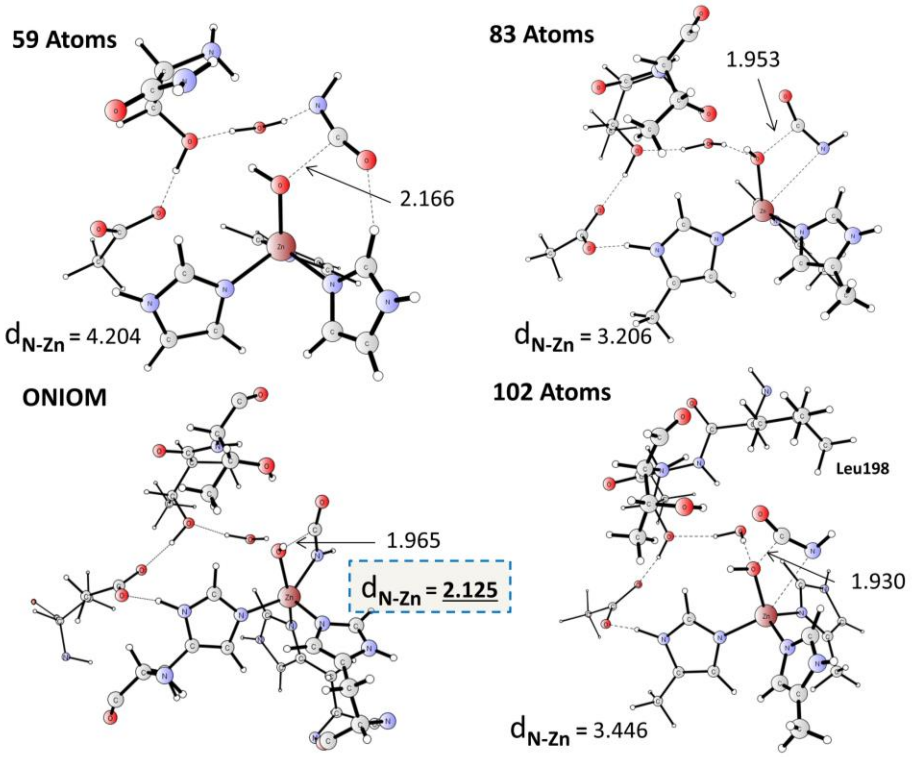


Figure 4.6 QM and QM/MM optimized structures of the TS of cyanic acid for approach 2. In the QM/MM representation the MM portion was omitted. All distances are given in Å.

The increment of the cluster size corresponds to a decrement of the distance between the nucleophilic water molecule and the carbon center of the cyanic acid substrate. However, only in the QM/MM geometry is possible to observe a coordination of the activated substrate to the metal core with a distance of 2.125 Å which can explain, presumably, the low activation energy related to the QM/MM description.

4.3 CA as degrading agent of carbonyl sulfide (COSase) [Paper III]

In order to give insights on the working mechanism of the novel characterized enzyme carbonyl sulfide hydrolase (COSase), that efficiently converts COS in H_2S and CO_2, a detailed theoretical investigation was performed in the framework of density functional theory (B3LYP and M06 exchange-correlation functionals) by using the cluster model approach. The proposed mechanism only partially follows that accepted for human carbonic anhydrase. In the final part of the reaction the metal ion is unable to form a penta-coordinated species. The B3LYP-D3 [131] and M06 potential energy surfaces show a very similar shape. The elucidation of the catalytic reduction of COS assumes also importance in view of the fact that carbonyl sulfide contributes to the global climate change.

4.3.1 Active site model

The $Zn(Cys)_2HisOH$ primary coordination shell of the carbonic anhydrase along with amino acid residues belonging to the outer one has been taken into account for building the cluster (PDB code 1YLK).

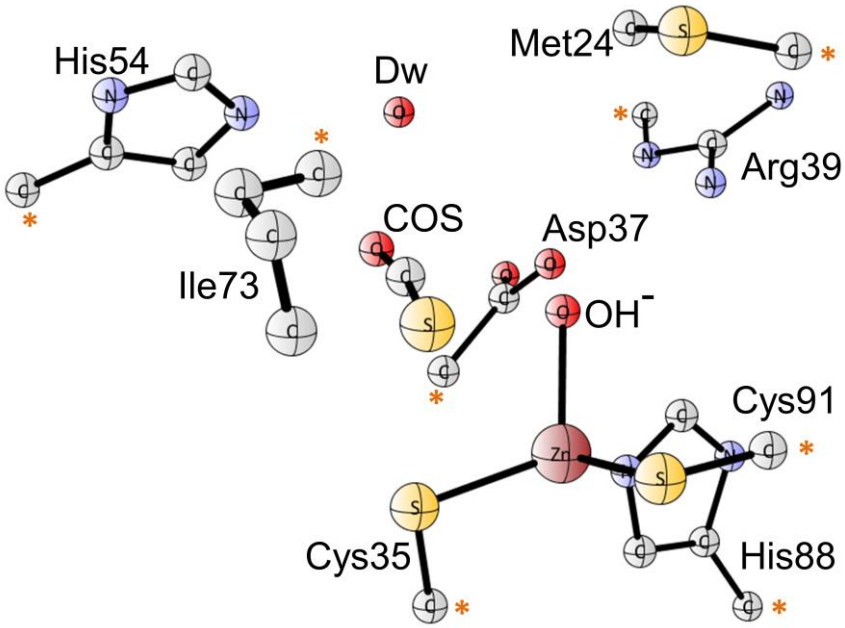


Figure 4.7 Catalytic site model used in the calculations. The stars indicate the atoms kept frozen during the optimization procedure.

In particular, amino acids involved in the network of H-bond present in the active site and essential to warranty the structural flexibility useful during the working mechanism of the enzyme have been retained: His54, Ile73, Met24. Amino acids surrounding the metal center are Asp37 and Arg39. Ile73 residue comes into play in the approaching phase of the substrate because it marks the area surrounding the active center where the COS can be activated by the nucleophile group. The dyad Asp37-Arg39 forms the salt bridge observed in the active site of *β*-carbonic anhydrases arising from other microorganisms. As usual the retained amino acids have been truncated at the Cα atom and, in order to avoid an unrealistic cluster expansion, the terminal carbon atom for all the amino acids has been fixed in their X-ray crystal position.

We have considered two possible mechanisms that essentially differ from the position of the deep water molecule interacting with the salt bridge Asp-Arg. Only in one of them the $\mathbf{W_{DW}}$ in directly involved in the reaction with a protons shift in which it acts simultaneously as hydrogen acceptor and donor (**TS4**).

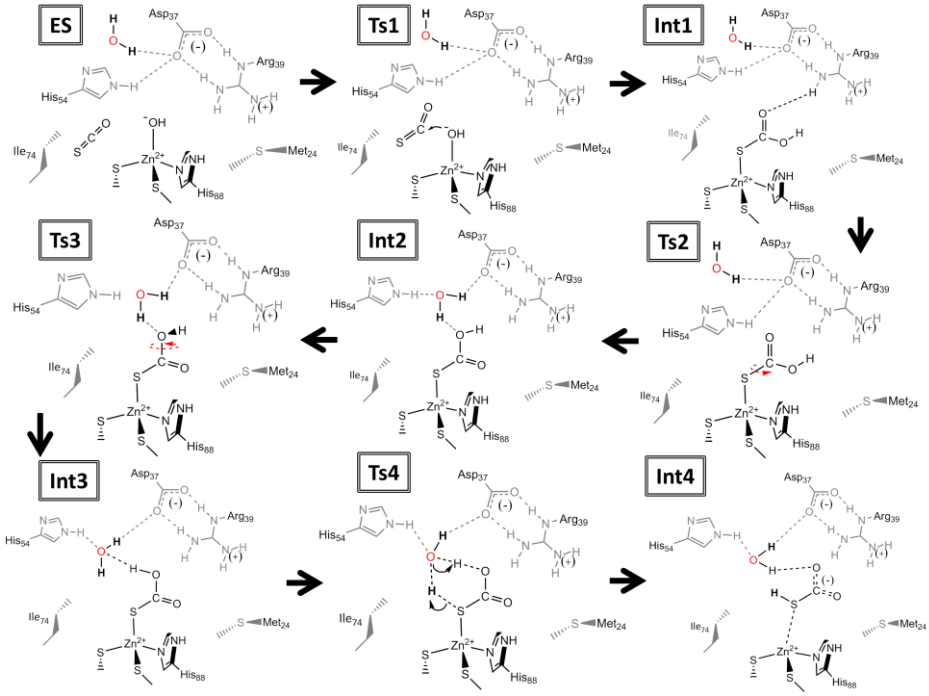


Figure 4.8 Proposed reaction mechanism for the conversion of COS in H_2S and CO_2.

The production of H_2S and CO_2 requires different steps characterized by energetics lying under the formation of the Michaelis-Menten complex (**ES**) except for the first transition state (**TS1**).

The entire investigation with a detailed description of the proposed mechanism and the computed geometries can be found in the attached **PaperIII**.

4.4 **Rhodium-substituted CA as an hydrogen-utilizing reductase** [Paper IV]

The improvement of the enzymatic specificity, making enzymes able to catalyze new, sometimes unnatural, reactions is a promising frontier in biocatalysis. In this context, the natural activity of some enzymes can be enriched by using several strategies to expand reaction specificity. Carbonic Anhydrase (CA) plays a key role in enzymatic catalytic processes owing to its natural promiscuity toward several substrates. Despite its strong catalytic ability, CA is not able to reduce organic molecules and it can't work as hydrogenase. To bridge this gap one of the new approaches is to replace the native metal ion with rhodium making CA the first cofactor-independent reductase that reduces organic molecules using hydrogen [111]. This possibility suggested the exploitation of the natural affinity of CA for CO_2 in order to define a new hydrogenation process of CO_2 to formic acid opening an innovative perspective for an efficient chemical reduction of carbon dioxide. In order to obtain the potential energy surface of the hydrogenation mechanism by rhodium substituted CA (Rh-CA), a computational study at density functional theory level has been undertaken. Our investigation showed how Rh-CA acts as efficient reductase confirming that biological systems are often rising alternatives for chemically difficult conversions.
Different reaction mechanisms were considered such as reductive elimination and σ-bond metathesis. The investigation also involved the release of the formic acid and the restoring of the catalytic cycle. Results show that the σ -bond metathesis potential energy surfaces lies below the reactant species, while the rate determining step is the release for the product with an energy barrier of 12.8 kcal/mol. On the basis of our results we can conclude that this artificial enzyme can efficiently catalize the direct hydrogenation of CO_2.

4.5 Ligand's affinity to hCAII [Paper V]

Qualitative determination of the binding properties between proteins and ligands is fundamental in improving ligands or to determine favourable interaction sites of ligands within proteins. This processes are a transition from one equilibrium state to another which often is a rare event. To characterize rare events is necessary to add external forces to reduce energy barriers and to speed up the kinetics.

Starting from our previous mechanistic investigations on the hCAII activity towards carbon dioxide and carbodiimide (**Paper I**), we validated a strategy based on nonequilibrium SMD simulations method to determine the potential energy related to the release's process of the two reaction products bicarbonate and ureate for CO_2 and carbodiimide respectively. In order to restore the catalytic cycle a water molecule must substitute the reaction product. In the case of native substrate, our QM and QM/MM computations showed that this substitution can easily occur, but on the contrary, in the case of ureate the water coordination on Zn^{2+} and the release of the product seem to be a difficult process and strictly related to the interaction between the reaction product and the metal core.

Bibliography

[1] Szabo, A.; Ostlund, N.S. Modern Quantum Chemistry; *Dover Publication*: New York, **1989**.

[2] Cramer, C. J. Essentials of Computational Chemistry; *John Wiley & Sons Ltd.*: New York, **2002**.

[3] Lewars, E. Computational Chemistry*; Kluwer Academic Publishers*: New York, **2004**.

[4] Piela, L. Ideas of Quantum Chemistry; *Elsevier B. V.*: Oxford, **2007**.

[5] Jensen, F. Introduction to Computational Chemistry; *John Wiley & Sons Ltd.*: New York, **2007**.

[6] Parr, R. G.; Yang, W. Density-Functional Theory of Atoms and Molecules; *Oxford University Press*: New York, **1989**.

[7] Koch, W.; Holthausen, M.C. A Chemist's Guide to Density Functional Theory; *John Wiley & Sons Ltd.*: Weiheim, **2000**.

[8] Schrödinger, E. *Ann. Physik.* **1926**, 79, 361.

[9] Born, M.; Oppenheimer, J.: *Ann. Physik.* **1927**, *84*, 457.

[10] Hartree, D. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 89.

[11] Hartree, D. *Proc. Cambridge Phil. Soc.*:**1928**, *24*, 111.

[12] Hartree, D. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 426.

[13] Fock, V. *Z. Physik.* **1930**, *61*, 126.

[14] Pauli, W. *Z. Physik.* **1925**, *31*, 765.

[15] Bartlett, R. J.; Stanton, J. *In Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D.B.: Eds.; *VCH Publishers*: New York, **1994**; Vol. V.

[16] Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864.

[17] Kohn, W.; Sham, L. *J. Phys. Rev. A* **1965**, *136*, 1133.

[18] Becke, A. D. *J. Phys. Phys.* **1993**, *98*, 1372.

[19] Becke, A. D. *J. Phys. Phys.* **1993**, *98*, 5648.

[20] Becke, A. D. *J. Phys. Phys.* **1988**, *38*, 3098.

[21] Vosko, S.H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

[22] Lee, C.; Yang, W.; Parr, R.G. *Phys. Rev. B* **1988**, *37*, 785.

[23] Sousa S. F., Fernandes, P.A.; Ramos M. J. *J. Phys. Chem. A* **2007**, *111*, 10439 and references therein

[24] Grimme, S. *J. Comp. Chem.* **2006**, *27*, 1787.

[25] Perdew, J. P.; Ziesche, P.; Eschig, H.; *Eds. Akademie Verlag*: Berlin, Germany, **1991**, 11.

[26] Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.

[27] Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664.

[28] Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908.

[29] Y. Zhao; D. G. Truhlar: *Theor. Chem. Acc.* **2008**, *120*, 215.

[30] Eyring, H. *J. Chem. Phys.* **1935**, *3*, 107.

[31] Fisher, E. *Ber. Deutsch. Chem. Ges.* **1984**, *27*, 2985.
[32] Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chem. Rev.* **2008**, *108* (3), 946.
[33] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.
[34] Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232.
[35] Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, *14*, 643.
[36] Siegbahn, P. E. M.; Himo, F. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2011**, *1*, 323.
[37] Noodleman, L.; Lovell, T.; Han, W.; Li, J.; Himo, F. *Chem. Rev.* **2004**, *104*, 459.
[38] Chen, S. L.; Fang, W. H.; Himo, F. **2008**, *120*, 515.
[39] Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**: *100*, 421.
[40] Lengauer, T; Rarey, M. *Curr. Opin. Struct. Biol.* **1996**, *6* (3), 402.
[41] Feig, M.; Onufriev, A.; Lee MS; Im. W.; Case, D; Brooks, C. *J. of Comp. Chem.* **2004**, *25* (2), 265.
[42] Halperin, I; Ma, B,; Wolfson, H; Nussinov, R.; *Proteins* **2002**, *47* (4), 409.
[43] Jain, A, *Curr. Protein Pept. Sci.* **2006**, *7* (5), 407.
[44] Böhm, H. J. *J. Comput. Aided Mol. Des.* **1998**, *12* (4), 309.
[45] Muegge, I. *J. Med. Chem.* **2006**, *49* (20), 5895.
[46] Jensen, F. Introduction to Computational Chemistry; *Wiley*, **2006.**
[47] Schlick, T. Molecular Modeling and Simulation: An Interdisciplinary Guide; *Springer*, **2010**.
[48] Boas, F. E.; Harbury, P. B. **2007**, *17*, 199.
[49] Lennard-Jones, J. E. *Proc. R. Soc. Lond. A* **1924**, *106* (738): 463.
[50] Halgren, T. A.; Damm, W.; *Curr Opin Struct Biol.* **2001**, *11*(2), 236.
[51] Vanommeslaeghe, K. et al *J. Comp. Chem.* **2010**, *31*, 671.
[52] Cornell, W.D.; Cieplak, P.; Bayly. C. I.; Gould, I. R.; Merz, K. M. Jr; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
[53] Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. **1998**, *4*, 31.
[54] Möllhoff, M.; Sternberg, U. *J. Mol. Model.* **2001**, *7,* 90.
[55] Oostenbrink, C.; Villa, A; Mark, AE; van Gunsteren, WF. *J. Comput. Chem.* **2004**: *25* (13): 1656.
[56] Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110* (6), 1657.
[57] Warshel, A. *Israel J. Chem.* **1973**, *11,* 709.
[58] Warshel, A.; Levitt, M. QCPE 247, Quantum Chemistry Program Exchange, Indiana University, **1974**.
[59] Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024.
[60] Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Schulten, K.: *Biophys. J.* **1997**, *72*, 1568.

[61] Bayas, M. V.; Schulten, K.; Leckband, D. *Biophys. J.* **2003**, *84*, 2223.
[62] Kosztin, D.; Izrailev, S.; Schulten, K. *Biophys. J.* **1999**, *76*, 188.
[63] Stepaniants, S.; Izrailev, S.; Schulten, K.: *J. Mol. Model.* **1997**, *3*, 473.
[64] Molnar, F.; BenNun, M.; Martinez, T.J.; Schulten, K. *J. Mol. Struct.* (THEOCHEM), WATOC special issue **2000**, *506*, 169.
[65] Grubmüller, H.; Heymann, B.; Tavan, P. *Science* **1996**, *271*, 997.
[66] Marrink, S.J.; Berger, O.; Tieleman, P.; Jähnig, F. *Biophys. J.* **1998**, *74*, 931.
[67] Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690.
[68] Jarzynski, C. *Phys. Rev. E* **1997**, *56*, 5018.
[69] Callen, H. B. *John Wiley & Sons, New York*, 2nd ed. **1995**.
[70] Park, S.; Schulten, K. *J. Chem. Phys.* **2004**, 120, 5946.
[71] Wood, R. H.; Mühlbauer, W. C. F.; Thompson, P. T. *J. Phys. Chem.* **1991**, *95*, 6670.
[72] Hermans, J. *J. Phys. Chem.* **1991**, *95*, 9029.
[73] Marcinkiewicz, J. *Math. Z.* **1939**, *44*, 612.
[74] Laio, A; Parrinello, M. *Proc. Natl. Acad. Sci. USA* **2002**, *99*(20), 12562.
[75] Barducci, A; Bonomi, M.; Parrinello, M. *WIREs Comput Mol Sci* **2011**, *1*, 826.
[76] Laio, A. and Parrinello, M. *PNAS* **2002** *99*, 12562.
[77] Limongelli,V.; Bonomi, M.; Parrinello, M. *PNAS* **2013**, *110*, 6358.
[78] Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
[79] Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *THEOCHEM* **1999**, *461*, 1.
[80] Maseras, F.; Morokuma, *K. J. Comput. Chem.* **1995**, *9*, 1170.
[81] Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K.: *J. Phys. Chem.* **1996**, *100*, 19357.
[82] Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. *J. Chem. Theo. Comp.* **2006**, *2*, 815826.
[83] Keilin, D.; Mann, T. *Biochem. J.* **1940**, *34*, 1163.
[84] Hewett-Emmett, D.; Tashian, R. E. *Mol. Phylogenet. Evol.* **1996**, *5*, 50.
[85] Hewett-Emmett, D. In The Carbonic Anhydrases*: New Horizons*; Chegwidden, W. R.; Carter, N. D.; Edwards, Y. H., Eds.; Birkahüser Verlag: Basel, Switzerland, **2000**; Vol. *90*.
[86] Lane, T. W.; Morel, F. M. M. *Plant Physiol.* **2000**, *123*, 345.
[87] Tripp, B. C.; Smith, K.; Ferry, J. G. *J. Biol. Chem.* **2001**, *276*, 48615.
[88] So, A. K.-C.; Espie, G. S.; Williams, E. B.; Shively, J. M.; Heinhorst, S.; Cannon, G. C. *J. Bacteriol.* **2004**, *186*, 623.
[89] Sawaya, M. R.; Cannon, G. C.; Heinhorst, S.; Tanaka, S.; Williams, E. B.; Yeates, T. O.; Kerfeld, C. A. *J. Biol. Chem.* **2006**, *281*, 7546.
[90] Lehtonen, J.; Shen, B.; Vihinen, M.; Casini, A.; Scozzafava, A.; Supuran, C. T.; Parkkila, A.-K.; Saarnio, J.; Kivela¨, A. J.; Waheed, A.; Sly, W. S.; Parkkila, S. *J. Biol. Chem.* **2004**, *279*, 2719.

[91] Hilvo, M.; Tolvanen, M.; Clark, A.; Shen, B.; Shah, G. N.; Waheed, A.; Ralmi, P.; Hanninen, M.; Hamalainen, J. M.; Vihinen, M.; Sly, W. S.; Parkkila, S. *Biochem. J.* **2005**, *392*, 83.
[92] Supuran, C. T.; Scozzafava, A.; Casini, A. *Med. Res. Rev.* **2003**, *23*, 146.
[93] Nishimori, I.; Minakuchi, T.; Kohsaki, T.; Onishi, S.; Takeuchi, H.; Vullo, D.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 3585.
[94] Pocker, Y.; Sarkanen, S. *AdV. Enzymol. Mol. Biol.* **1978**, *47*, 149.
[95] Eriksson, A. E.; Jones, T. A.; Liljas, A. *Proteins: Struct., Funct.,Genet.* **1988**, *4*, 274.
[96] Liljas, A.; Kannan, K. K.; Bergsten, P. C.; Waara, I.; Fridborg, K.; Strandberg, B.; Carlbom, U.; Jarup, L.; Lovgren, S.; Petef, M. *Nature New Biol.* **1972**, *235*, 131.
[97] Davis, R. P. J. *Am. Chem. Soc.* **1959**, *81*, 5674.
[98] Coleman, J. E. *J. Biol. Chem.* **1967**, *242*, 5212.
[99] Merz, K. M.; Hoffmann, R.; Dewar, M. J. S. *J. Am. Chem. Soc.* **1989**, *111*, 5636.
[100] Lipscomb, W. N. *Annu. Rev. Biochem.* **1983**, *52*, 17.
[101] Lindskog, S.; Spiro, T. G. *Wiley*: New York, **1983**, 77.
[102] Håkansson, K.; Carlsson, M.; Svensson, L.A.: Liljas, A. *J. Mol Biol.* **1992**, *227*(4), 1192.
[103] Raushel, F.M.; Holden, H.M. *Adv Enzymol Relat Areas Mol Biol* **2000**, *74*, 51.
[104] Jensen, R.A. *Annu Rev Microbiol* **1974**, *30*, 409.
[105] Janssen, D.B.; Dinkla, I.J.; Poelarends, G.J.; Terpstra, P. *Environ Microbiol* **2005**, *7*, 1868.
[106] O'Brien, P.J.; Herschlag, D. *Chem Biol* **1999**, *6*, 91.
[107] Copley, S. D *Curr Opin Chem Biol* **2003**,*7*, 265.
[108] Hult, K.; Berglund, P. *TRENDS in Biotechnology* **2007**, *25,* 5.
[109] Khersonsky, O.; Tawfik, D. S. *Annu. Rev. Biochem.* **2010**. *79*, 471.
[110] Fernandez-Gacio, A.; Codina, A.; Fastrez, J.; Riant, O.; Soumillion, P.;. *ChemBioChem* **2007**, *7*, 1013.
[111] Jing, Q.; Okrasa, K., Kazlauskas, R.J *Chemistry*. **2009**. *15*, 1370.
[112] Alterio, V; Di Fiore, A; D'Ambrosio, K; Supuran, CT; De Simone, G. *Chem Rev.* **2012**, *112*(8), 4221.
[113] Gould, S. M.; Tawfik; D. S. *Biochemistry* **2005**, *44*, 5444.
[114] Lopez, M.; Vu, H; Wang, C. K.; Wolf, M. G.; Groenhof, G.; Innocenti, A.; Supuran, C. T.; Poulsen, S-A. *J. Am. Chem. Soc.* **2011**, *133* (45), 18452.
[115] Pocker, Y.; Meany, J.E. *J. Am. Chem. Soc.* **1965***, 87*, 1809.
[116] Henkart, P.; Guidotti, G.; Edsall, J.T.: *J. Biol. Chem.* **1968**, *243*, 2447.
[117] Malmstrom, B. G.; Nyman, P.O.; Strandberg, B., Tilander, B.; Goodwin, T.W.; Harris, J.T.; Hartley, B.S. *New York: Academic Press*, **1964**, 121.
[118] Whitney, P.L.; Folsch, G.; Nyman, P.O.; Malmstrom, B.G. *J. Biol.*

Chem. **1967**, *242*, 4206.
[119] Supuran, C. T.; Scozzafava, A.; Casini, A. *Med. Res. Rev.* **2003**, *23*(2), 146**.**
[120] Maresca, A.; Temperini, C.; Vu, H.; Pham, N. B.; Poulsen, S.-A.; Scozzafava, A.; Quinn, R. J.; Supuran, C. T. *J. Am. Chem. Soc.* **2009,** *131* (8), 3057.
[121] Maresca, A.; Temperini, C.; Pochet, L.; Masereel, B.; Scozzafava, A.; Supuran, C. T. *J. Med. Chem.* **2009**, *53* (1), 335.
[122] Siegbahn, P. E. M.: *J. Biol. Inorg. Chem*. **2006**, *11*, 695.
[123] Perdew, J. P.; Ernzerhof, M.; Burke, K..*J. Chem. Phys.* **1996**, *105*, 9982.
[124] Adamo, C; Barone, V.: *J. Chem.* Phys. **1999** *110* (13), 6158.
[125] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378.
[126] Sevastik, R.; Himo, F. *Bioorg. Chem*. **2007**, 35, 444.
[127] Hopmann, K. H.; Himo, F. *J. Chem. Theory Comput*. **2008**, 4, 1129.
[128] Georgieva, P.; Himo, F. *J. Comput. Chem*. **2010**, 31, 1707.
[129] Liao, R.-Z.; Yu, J.-G.; Himo, F. *J. Chem. Theory Comput*. **2011**, 7, 1494.
[130] Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R-Z. and Siegbahn, P.E.M. *Chem. Rev*. **2014**, *114*, 3601
[131] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys*. **2010**, *132*, 154104.

Work in progress

Promiscuous activity of Human Carbonic Anhydrase II: hydrolysis of derivatized carbohydrates as sulfamates

Paolo Piazzetta, Tiziana Marino, and Nino Russo

b

Esterase activity of hCA

Carbonic Anhydrases has shown to exhibit weak and promiscuous esterase activity toward activated esters such as p-nitrophenyl acetate (k_{cat}/K_M ~10^3 M^{-1} s^{-1} *vs* ~108 M^{-1} s^{-1} for native substrate) [1,2]. However the physiological relevance of CAs catalytic esterase activity is unknown and it is halted in presence of CA inhibitors (*See chapter 2*)[3].
This aspect of promiscuous activity reported for hCA falls under the category of enzyme catalytic promiscuity shown by enzymes catalysing "distinctly different chemical reactions with different transition states" [4,5].
Most recently crystal structures of human CA II in complex with four of the sulfamate inhibitors have been obtained (**Figure 1**)[3]. These inhibitors represent a novel compound class of carbohydrate-based sulfamates where the carbohydrate hydroxyl groups can be free (-OH) or acylated (-OAcyl). From X-ray crystallography is emerged that in addition to the canonical amino acids involved in the interaction with the inhibitor several amino acid residues of hCAII can interact with the bound ligands as reported in **Figure 1**.

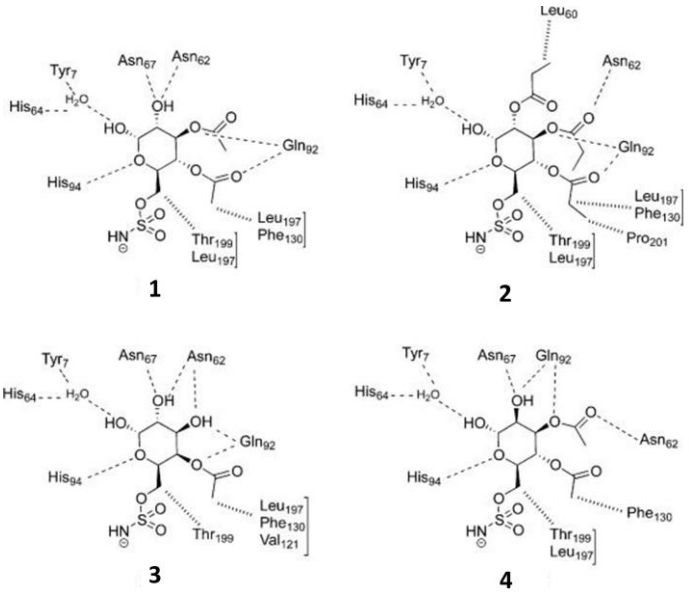


Figure 1. The canonical observed ligand interaction from inhibitor – bound hCAII [3].

Has been proven that derivatized carbohydrates as sulfamates can act as inhibitors of CO_2 hydration and substrates for the esterase. All the proposed inhibitor/substrates species present the C–6 primary hydroxyl group derivatized as sulfamate for the CA recognition. The remaining four sugar hydroxyl groups have been acylated[3]. In the same work has been proposed that, while initially the substrate binding mode is prevalent, after hydrolysis the ligand can also act a pure inhibitor thereby competing with the substrate binding mode[3].
In order to describe this other aspect of the promiscuous activity of hCA by adding further mechanistic details on its working mechanism, we have undertaken a study performed at both QM and QM/MM (ONIOM) [6-10] levels of theory.

Active site Model and computational approaches

The initial structure, generated by a docking procedures with a consequent MD equilibration, has been kindly provided by the authors of the ref. 3. In the QM calculations, the cluster model procedure has been followed. In detail the considered cluster model is characterized by the substrate, the zinc metal ion, a coordinated hydroxyl group, the three histidine residues (94,96,119) and a free water molecule (Deep Water). In this case a more extended number of residues belonging to the second coordination shell and considered fundamental for hold the very bulky substrate has been retained: Tyr7, Asn62, Asn67, Gln92, Glu106, Phe130, Leu198, Thr199 and Thr200 (191 atoms, charge=0). Except the entirely considered Leu197, Thr199 and Thr200 residues, the remaining ones are truncated to the C-alpha ones that are kept frozen to its initial positions during optimization procedure for preserving their spatial arrangement. The geometries optimization has been carried out using B3LYP functional [11,12] associated to 6-31+G(d,p) basis set on all elements except for Zn, which is described by the relativistic compact Stuttgart/Dresden effective core potential (SDD) pseudopotential.[13]
A three layered ONIOM (B3LYP/6-31+G(d,p):AM1:UFF) [6-10] model has been chosen considering the same truncated residues of the QM cluster. The substrate, the zinc with the three imidazole rings, the OH group and the free water molecule are described at the high level (79 atoms), while all the remaining residues are associated to a semiempirical description (AM1 – 100 atoms). The protein environment is modeled using UFF

(3934 atoms). For both QM and QM/MM (ONIOM) calculations, the Gaussian03 package has been used. [14]

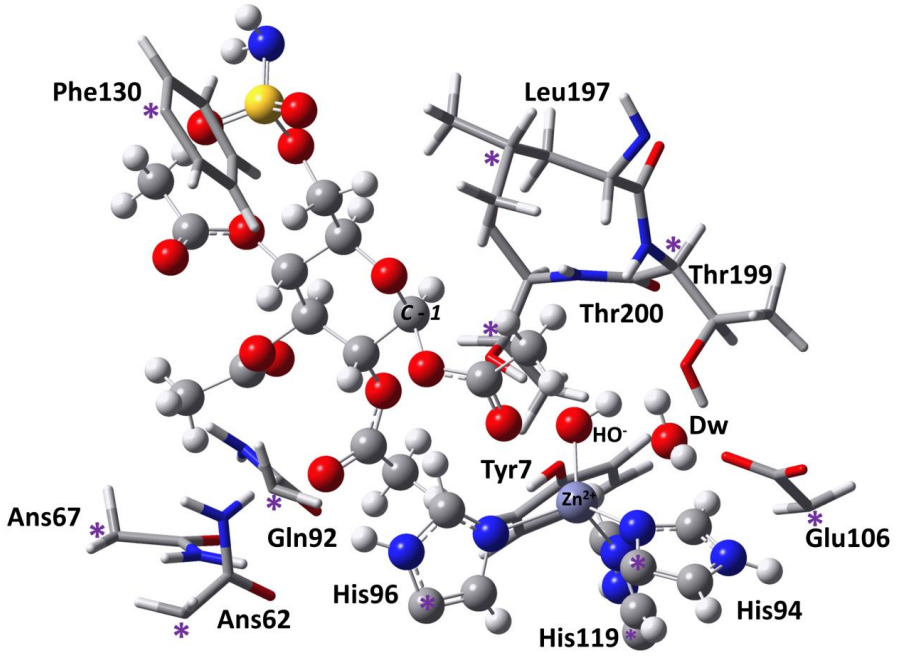


Figure 3. Selected model of the hCAII active site. Stars indicate the atoms that are fixed to their initial position in the QM cluster. The QM representation involves the depicted structure with hydrogen saturations where the catting has been made. As far as the ONIOM representation is concerned the high layer (DFT) is shown in ball and stick; the medium layer in tubes contains the amino acid residues. The remaining protein falling in the MM (UFF) portion is omitted for clarity.

Results and discussion

In **Figure 2** is depicted the proposed mechanism in which hydrolysis occurs via alkyl–oxygen cleavage with formation of oxacarbenium resonance stabilized cation and the consequent departure of the zinc coordinated acyl group[3]. The zinc coordinated hydroxyl then reacts with the oxacarbenium cation leading to the axial anomer and zinc bound acetate. The acetate is displaced by water and the active CA with zinc bound hydroxyl is regenerated [3] (**Figure 2**).

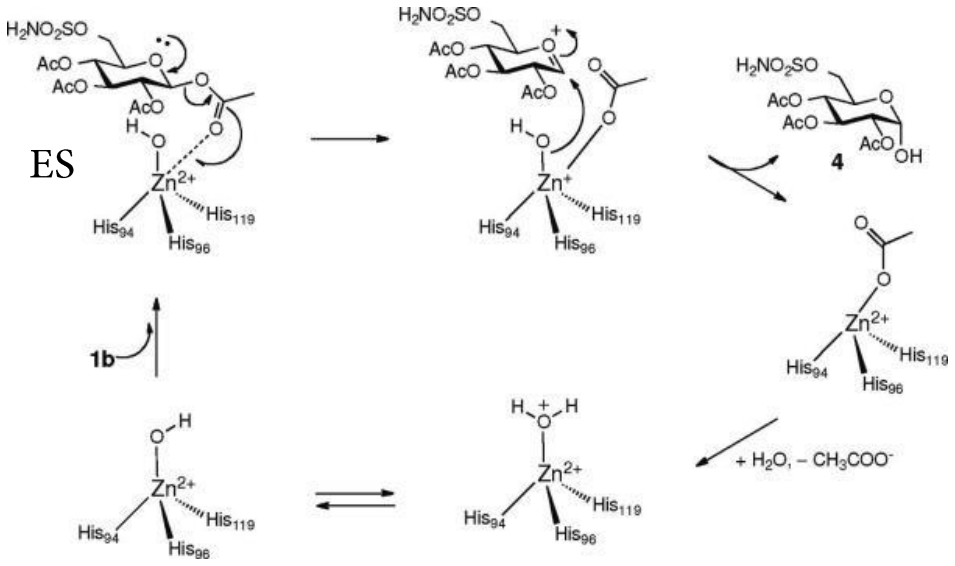


Figure 2. Proposed reaction mechanism by Lopez et *al* [3].

By using this guideline we have started studying the catalytic mechanism to be followed by hCAII in its esterase promiscuous activity.

ES optimization The Michaelis-Menten (ES) complex has been obtained for both QM and QM/MM (ONIOM) approaches. The optimized structures are given in Appendix, while in **Figure 3** are depicted a more simplified version of the QM and ONIOM geometries. Both optimized ES complexes showed comparable behavior, in particular has been observed that the acyl moiety on the C–1 moves away from the Zn^{2+}. The -C=O – Zn^{2+} distance changes from 3.055 Å to 5.853 Å and 5.377 Å in the QM and QM/MM respectively. These findings can be due to the established extended H-bond network of the substrate that pulls out the coordination shell of the metal ion. This can influence the next step represented by the alkyl-oxygen cleavage, as depicted in **Figure 2**, that is triggered by the C–1 acyl coordination to the Zn^{2+} for favoring the formation of the oxacarbenium cation. The search of the related TS is in progress.

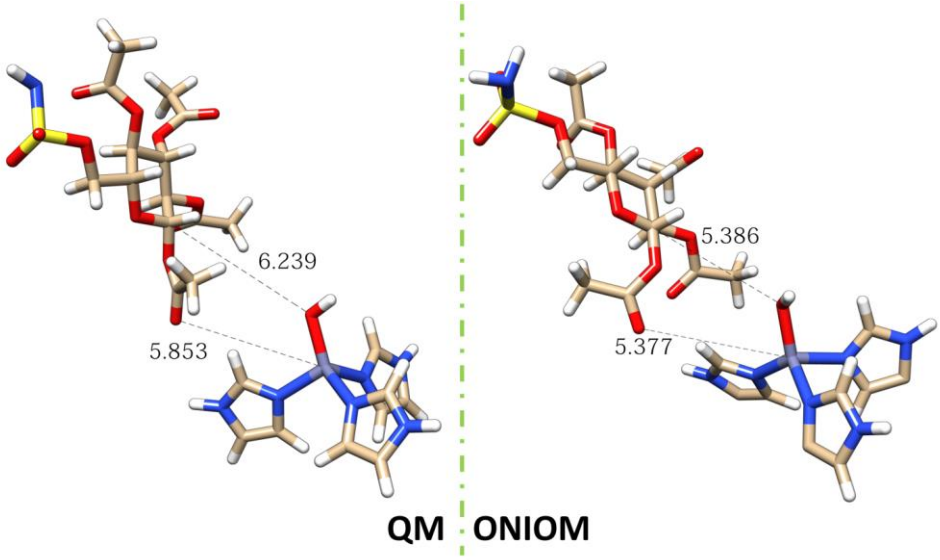


Figure 3. Optimized ES complex representation with both QM and ONIOM. Only the substrate and the metal core with the coordinated histidines were depicted (distances in Å).

References

[1] Krebs, J. F.; Ippolito, J. A.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem*. **1993**, *268*, 27458.

[2] Aharoni, A.; Gaidukov, L.; Khersonsky, O.; McQ Gould, S.; Roodveldt, C.; Tawfik, D. S. *Nat. Genet.* **2005**, *37*, 73.

[3] Lopez, M.; Vu, H; Wang, C. K.; Wolf, M. G.; Groenhof, G.; Innocenti, A.; Supuran, C. T.; Poulsen, S-A. *J. Am. Chem. Soc*. **2011**, *133* (45), 18452.

[4] Hult K.; Berglund, P. *Trends Biotechnol*. **2007**, *25*, 231.

[5] Khersonsky, O.; Roodveldt C.; Tawfik, D. S. *Curr. Opin. Chem. Biol.* **2006**, *10*, 498.

[6] Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.

[7] Ke, Z.; Abe, S.; Ueno, T.; Morokuma, K. J. *Am. Chem. Soc.* **2012**, *134*, 15418.

[8] Martí, S.; Andrés, J.; Silla, E.; Moliner, V.; Tuñón, I.; Bertrán, J. *Angew. Chem. Int. Ed*. **2007**, *46*, 286.

[9] Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.

[10] van der Kamp, M. W.; Mulholland, A. J. *Biochemistry*, **2013**, *52*, 2708.

[11] Becke, A. D. J. *J. Chem. Phys*. **1993**, *98*, 5648.

[12] Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

[13] Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, **2004**.

[14] Andrae, D.; Haussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123.

Coordinates of the optimized ES (QM)

C	-9.12440500	6.96797100	-1.73557100
C	-9.34244300	6.49122200	-0.43609200
H	-9.73272000	7.15767500	0.32855300
C	-9.06748800	5.16327200	-0.10165600
H	-9.23757200	4.79245100	0.90444800
C	-8.56625100	4.28252000	-1.07145500
O	-8.31789900	2.99382800	-0.67522200
H	-7.97589400	2.43415000	-1.39212200
C	-8.34412500	4.74738400	-2.37471700
H	-7.95897400	4.07176900	-3.13503800
C	-8.62387100	6.08192500	-2.69655300
H	-8.44732600	6.42422200	-3.71273500
C	0.09205800	6.93486000	6.45661000
H	0.08729600	7.94429600	6.03818700
H	-0.91908400	6.68466300	6.79254300
C	0.55192200	5.86942200	5.47782400
O	0.66338600	4.68846200	5.82139400
N	0.81311800	6.27646900	4.21397100
H	0.71538400	7.24106600	3.94035000
H	1.08517100	5.61314300	3.49561800
C	0.64675900	1.27114400	8.72473000
H	1.53979100	0.74601200	9.07738100
H	0.77039900	2.34273300	8.90218400
C	0.44300000	0.93446900	7.25675800
O	0.14872700	-0.23344800	6.91417000
N	0.62123800	1.90353700	6.34942300
H	0.78951900	2.88298400	6.55990600
H	0.53911900	1.67325400	5.36418200
C	3.01654200	-2.89474800	6.65419900
H	2.97151600	-3.98458200	6.58166300
H	3.95435100	-2.64111700	7.16016600
C	3.04830200	-2.30844900	5.25214400
O	3.80527000	-2.73104500	4.37925200
N	2.18691000	-1.27764600	5.04954900
H	1.55481100	-0.95306600	5.77235000
H	2.16868300	-0.80584700	4.15735800
C	-2.73431200	-1.60735800	5.50693500
N	-1.58173300	-1.27436700	4.83766200
H	-0.77379500	-0.83856900	5.29230400
C	-1.73012500	-1.63744600	3.54702800
H	-0.99122300	-1.46872900	2.78154700
N	-2.92369700	-2.19575300	3.35836900
C	-3.56351900	-2.17997200	4.58236400
H	-4.56550800	-2.55892900	4.70923600
C	-7.80061700	-1.38780600	2.17992600
N	-7.33409800	-0.52578000	1.20390700
H	-7.85584300	0.18322400	0.70303700
C	-6.04730500	-0.83907000	0.94275800
H	-5.45732800	-0.34399400	0.18967000
N	-5.65306800	-1.84926700	1.70519300
C	-6.74339500	-2.20494200	2.47984600
H	-6.70279300	-3.02637000	3.17908100
C	-8.41755700	-3.26087200	-2.15319100
H	-8.55322100	-2.71803200	-3.09527000

H	-8.92504000	-4.22537900	-2.21492700
C	-6.93897300	-3.44029300	-1.86817900
O	-6.47614000	-4.59877300	-1.76324800
O	-6.24634100	-2.36100300	-1.75379400
C	-4.48226600	-5.57097000	2.75122800
N	-4.36540900	-4.70221700	1.67365800
C	-5.02324400	-5.26330200	0.66252800
H	-5.19926000	-4.85584000	-0.33300200
N	-5.55294500	-6.44405500	1.05572300
H	-6.11676000	-7.03132400	0.45683600
C	-5.22785600	-6.66399500	2.38167600
H	-5.52665100	-7.54948100	2.91951300
C	11.65383300	-4.07269600	2.52949800
C	11.43672300	-5.00805400	1.51231700
H	11.87694900	-5.99935700	1.58212100
C	10.64998000	-4.66755700	0.40684400
H	10.47850300	-5.39529100	-0.38205700
C	10.08071900	-3.39285100	0.31651000
H	9.46793200	-3.12213800	-0.53813700
C	10.29864300	-2.45766800	1.33442300
H	9.85275500	-1.46972400	1.26266800
C	11.08471500	-2.79817900	2.44005900
H	11.25096700	-2.07259000	3.23196700
N	1.38083000	-5.25515400	-4.39063400
H	1.45326700	-5.12322000	-5.39876500
C	0.95955200	-3.99002400	-3.78303500
H	0.69733500	-4.19328000	-2.73711500
C	2.11422400	-2.96440400	-3.83958300
H	1.75508000	-1.99772000	-3.45707900
H	2.34690000	-2.80314000	-4.90026900
C	3.39797800	-3.36624000	-3.08859700
H	3.66516800	-4.37828100	-3.41696100
C	4.54564100	-2.41750300	-3.46799000
H	4.30502100	-1.37892200	-3.20426000
H	5.47472600	-2.68328400	-2.95213700
H	4.74211300	-2.44292100	-4.54615300
C	3.19989700	-3.40053200	-1.56537300
H	4.12632000	-3.68654900	-1.05516400
H	2.90175100	-2.41292100	-1.18705000
H	2.42952600	-4.11826300	-1.26339400
C	-0.25214500	-3.37491100	-4.50701800
O	-0.27016300	-3.28120900	-5.73357900
N	-1.26265800	-2.90287700	-3.71078500
H	-1.28662000	-3.17703900	-2.73583200
C	-2.48451200	-2.32274700	-4.25781200
H	-2.43361300	-2.45348600	-5.34346400
C	-3.73318500	-3.03703300	-3.71338400
H	-4.59818200	-2.50038000	-4.11946500
C	-3.80731600	-4.50838600	-4.12299300
H	-3.84284600	-4.60990700	-5.21413500
H	-2.93910200	-5.06432800	-3.75258800
H	-4.70775500	-4.95460400	-3.69269200
O	-3.72017200	-2.89518300	-2.29190400
H	-4.66368700	-2.73016000	-2.00180400
C	-2.60715900	-0.80855200	-4.01091800
O	-3.59583600	-0.19035900	-4.40953500

j

N	-1.57602300	-0.22184400	-3.34772300
H	-0.82284000	-0.80206400	-3.00621700
C	-1.60670300	1.15856300	-2.93492200
H	-2.34196100	1.67501600	-3.55537100
C	-2.03063000	1.31050200	-1.43040200
H	-1.71561800	2.31476600	-1.11492700
C	-3.54461700	1.20531700	-1.26571500
H	-3.90883500	0.23023900	-1.59782000
H	-3.80594500	1.34008500	-0.21120500
H	-4.05225400	1.97986900	-1.84927800
O	-1.33272200	0.38621700	-0.60989600
H	-1.90110400	-0.38807800	-0.34360500
C	-0.25754400	1.83025500	-3.13163700
O	0.82900200	1.29002100	-2.99195100
Zn	-3.74429700	-2.69608900	1.54637100
O	-2.67785200	-1.90076300	0.20018400
H	-2.89192000	-2.28380500	-0.66870600
C	-0.82089300	5.09034600	0.26933700
H	-1.07355100	6.02822900	0.76442200
H	-1.03660200	5.14467100	-0.80007500
H	-1.43656000	4.28879900	0.69281000
C	0.62718600	4.76128100	0.51375000
O	1.31902000	5.22381000	1.40040400
O	1.07772800	3.85656500	-0.39328300
C	1.31949500	-0.62631700	1.08208500
H	1.08051300	-1.49693200	1.69191100
H	0.57603200	-0.52837200	0.28177700
H	2.30887100	-0.76262600	0.64162900
C	1.26772600	0.60385500	1.94565000
O	0.82080300	0.63119600	3.07651100
O	1.73937800	1.78360500	1.43824000
C	6.36959700	2.84009400	-4.76550200
H	5.44581300	2.94211400	-5.33792400
H	7.12806000	3.53941500	-5.12040200
H	6.74546600	1.81949900	-4.89816800
C	6.13359300	3.05593700	-3.29529100
O	6.94341700	3.50921100	-2.50872700
O	4.89048100	2.64509700	-2.94294700
C	4.47932700	2.71563500	-1.56554600
H	5.21104000	3.29225100	-0.99264000
C	3.11966200	3.42477800	-1.58262200
H	2.49777400	2.98172300	-2.36427800
O	3.36373600	4.81382000	-1.88929000
C	2.67075500	5.39943800	-2.89093200
C	3.12472200	6.81874600	-3.10492700
H	2.50499500	7.28824500	-3.86821700
H	3.06282300	7.37799800	-2.16684100
H	4.17323300	6.82597700	-3.41899000
O	1.79886800	4.84584600	-3.53699700
C	2.40321000	3.32472000	-0.23682900
H	2.92941800	3.87358300	0.54787000
C	2.31094800	1.84586200	0.15056300
H	1.71272900	1.28596000	-0.57961000
O	3.62585900	1.32900600	0.24332000
C	4.34089000	1.28485000	-0.99719100
H	3.80467500	0.66547700	-1.72991700

C	5.65951200	0.62208900	-0.59378100
H	5.45066000	-0.37003100	-0.18428000
H	6.16994500	1.22538200	0.16034900
O	6.48657400	0.50382700	-1.77471800
S	8.05516800	0.08822100	-1.54063900
O	8.26687400	-1.30182100	-1.91864700
O	8.44161400	0.55017700	-0.20722000
N	8.67393400	1.02140600	-2.77718000
H	9.59842900	0.67975700	-3.02785400
H	8.66226000	2.01508600	-2.54365600
O	-6.92609900	0.30384300	-1.95780700
H	-6.34584900	0.56583600	-2.68532700
H	-6.80132800	-0.68128300	-1.91229300
H	-2.85026900	-1.41733600	6.56209400
H	-4.01754500	-5.35963700	3.70279700
H	-8.81042800	-1.34498300	2.55614000
H	-9.34097500	7.99982400	-1.99279500
H	0.74438500	6.91338900	7.33362800
H	-0.20551900	0.89694500	9.29731700
H	2.17849000	-2.53171100	7.25487400
H	12.26299800	-4.33733100	3.39002200
H	-8.88204700	-2.65301000	-1.36835200
O	-0.41016400	3.13317300	-3.40711100
H	0.45997100	3.58906400	-3.45633400
H	0.67769700	-5.97519100	-4.24250200

Coordinates of the optimized ES (ONIOM): *high* and *medium* layer.

C-C_R 65.82287500 44.49137600 20.43882200 M

C-C_R 66.83604900 44.27015600 21.38117900 M

H-H_ 67.28322500 43.27334300 21.48895900 M

C-C_R 67.32498900 45.29973200 22.17469300 M

H-H_ 68.13389800 45.13616200 22.90062700 M

C-C_R 66.77606500 46.58449100 22.03283500 M

O-O_R 67.29685400 47.55913600 22.85164200 M

H-H_ 66.80929200 48.37776500 22.69152000 M

C-C_R 65.75654700 46.82521000 21.10340200 M

H-H_ 65.32078800 47.82776600 21.00272300 M

C-C_R 65.28693100 45.77545700 20.31620400 M

H-H_ 64.49388200 45.98491600 19.58396300 M

C-C_3 65.26238100 57.23015300 19.30932600 M

H-H_ 65.19375300 58.23672400 19.79269300 M

H-H_ 66.07047700 57.28310700 18.53491200 M

C-C_R 65.68012300 56.17134100 20.31082100 M

O-O_R 66.37988900 55.19437800 19.96184700 M

N-N_R 65.30042300 56.30403600 21.61643600 M

H-H_ 64.77909400 57.07786000 21.93715800 M

H-H_ 65.57222100 55.62959900 22.29230300 M

C-C_3 65.47708700 58.73942500 25.15787000 M

H-H_ 66.22080000 59.28022500 25.79991300 M

H-H_ 65.40866300 59.26953300 24.17442500 M

C-C_R 65.95917400 57.32999100 24.87518700 M

O-O_R 65.21106900 56.43300300 24.41756100 M

N-N_R 67.29069700 57.05714000 25.04389300 M

H-H_ 67.86117200 57.65437400 25.59518400 M

H-H_ 67.61791000 56.12984200 24.90721800 M

C-C_3 68.57690300 59.10557600 29.67414800 M

H-H_ 68.08245300 58.39002700 30.37665700 M

H-H_ 69.58961400 59.36582100 30.06109300 M

C-C_R 68.66680800 58.41279000 28.33527400 M

O-O_R 67.62468400 58.05938100 27.72227800 M

N-N_R 69.88880800 58.18780800 27.76925500 M

H-H_ 70.70867300 58.24618700 28.32531400 M

H-H_ 69.94558100 57.63714500 26.94101800 M

C-C_R 66.42118800 54.57346100 28.39028600 H

N-N_R 67.67546900 55.13201900 28.28808300 H

H-H_ 67.91443500 56.02206400 27.86060200 H

C-C_R 68.57841600 54.22699500 28.73517400 H

H-H_ 69.64608100 54.40744000 28.75489900 H

N-N_R 67.95087200 53.11961900 29.11945200 H

C-C_R 66.60168200 53.31714600 28.89761400 H

H-H_ 65.87811000 52.54346000 29.09372300 H

C-C_R 65.68439000 48.41936800 29.03777000 H

N-N_R 66.75356700 47.89513400 28.33778800 H

H-H_ 66.97910300 46.90211000 28.14118000 H

C-C_R 67.70744600 48.85158900 28.29525900 H

H-H_ 68.69638100 48.70929700 27.87539200 H

N-N_R 67.30766800 49.95789100 28.91501800 H

C-C_R 66.04291100 49.69050500 29.40189800 H

H-H_ 65.50938900 50.40907300 30.00077900 H

C-C_3 68.94051900 44.67087800 31.57883800 M

H-H_ 68.99107400 44.02118500 30.67456100 M

H-H_ 69.64544800 44.27920300 32.34779200 M

C-C_R 69.34465900 46.07894500 31.19952000 M

O-O_R 69.80105800 46.87117400 32.07914400 M

O-O_R 69.19268700 46.47402900 29.99556400 M

C-C_R 67.65945300 51.12064200 32.58758100 H

N-N_R 68.48463600 50.73147100 31.54216600 H

C-C_R 68.85805300 49.48298100 31.80887300 H

H-H_ 69.48041600 48.82061400 31.19970200 H

N-N_R 68.31359000 49.06511200 32.96338200 H

H-H_ 68.48496200 48.13094900 33.33134100 H

C-C_R 67.54459500 50.08764400 33.48425900 H

H-H_ 67.00211100 49.97687000 34.40225700 H

C-C_R 78.68079100 62.42349300 29.10160800 M

C-C_R 77.48579800 62.29729400 29.81756500 M

H-H_ 77.18986200 63.06468700 30.54382900 M

C-C_R 76.64438400 61.20627200 29.60751500 M

H-H_ 75.71426400 61.11418900 30.18585400 M

C-C_R 76.98183200 60.23604600 28.66523400 M

H-H_ 76.32194700 59.37291200 28.49727000 M

C-C_R 78.16053300 60.35911200 27.92883700 M

H-H_ 78.41617700 59.60533200 27.15489000 M

C-C_R 79.00255200 61.45043300 28.14992400 M

H-H_ 79.92312000 61.53250900 27.55221300 M

N-N_3 77.49966700 49.46009800 29.89456900 M

H-H_ 78.23427500 49.47852600 29.23252600 M

C-C_3 76.23197300 49.82605900 29.33585500 M

H-H_ 75.56511300 50.20942100 30.16701500 M

C-C_3 76.32660400 50.92939400 28.26858400 M

H-H_ 75.30743200 51.07510300 27.82126200 M

H-H_ 77.00050100 50.59505400 27.43808300 M

C-C_3 76.83582800 52.25151700 28.82358800 M

H-H_ 77.76207900 52.05784100 29.43491900 M

C-C_3 77.21255100 53.17764600 27.68235100 M

H-H_ 76.31542900 53.44019900 27.06887900 M

H-H_ 77.65273500 54.12347000 28.08130200 M

H-H_ 77.96353000 52.69497100 27.01343900 M

C-C_3 75.80561700 52.92215500 29.70918000 M

H-H_ 76.23857500 53.84395900 30.16651200 M

H-H_ 74.90564900 53.21083300 29.11398000 M

H-H_ 75.47846800 52.24554900 30.53403900 M

C-C_R 75.53009600 48.60453200 28.66769000 M

O-O_R 76.19209700 47.66731000 28.17911100 M

N-N_R 74.15210600 48.57036900 28.66895900 M

H-H_ 73.63257300 49.28476600 29.12214500 M

C-C_3 73.40769200 47.47348900 28.11728900 M

H-H_ 74.11369800 46.59272300 27.99337100 M

C-C_3 72.22756200 47.00532700 29.00491600 M

H-H_ 71.57585500 46.32257200 28.38475500 M

C-C_3 72.73226000 46.25600000 30.22145000 M

H-H_ 73.23042500 45.31983000 29.88874300 M

H-H_ 73.46102100 46.87204800 30.79660400 M

H-H_ 71.87165100 46.00725600 30.89343400 M

O-O_3 71.47703200 48.14699000 29.38517400 M

H-H_ 70.73694800 47.81273400 29.92184300 M

C-C_R 72.84884900 47.74142200 26.70031400 M

n

O-O_R 72.76636600 46.80101800 25.88570300 M

N-N_R 72.39217000 49.00894500 26.38761000 M

H-H_ 72.51340800 49.73766500 27.05746900 M

C-C_3 72.34825300 49.45158900 25.01394400 M

H-H_ 72.03806600 48.55103500 24.40249200 M

C-C_3 71.29576800 50.57489000 24.78054400 M

H-H_ 71.58704600 51.15807800 23.85890600 M

C-C_3 69.92138300 49.95954400 24.59548500 M

H-H_ 69.61433300 49.40613900 25.51544500 M

H-H_ 69.17342700 50.76453300 24.39506400 M

H-H_ 69.92466300 49.24954500 23.73428300 M

O-O_3 71.30597300 51.55901300 25.79253500 M

H-H_ 70.90501400 51.19337500 26.59959000 M

C-C_2 73.71125600 49.95350000 24.51078500 M

O-O_2 74.51618600 50.44798100 25.30341500 M

Zn- 68.75399900 51.31710400 29.58047800 H

O-O_3 70.30703400 50.96889300 28.58183100 H

H-H_ 70.96582200 50.39304800 28.99156600 H

C-C_3 68.98327800 53.68834300 24.94138200 H

H-H_ 68.03103400 53.97411100 25.38777400 H

H-H_ 68.85805500 53.51586900 23.86778300 H

H-H_ 69.34566700 52.74839000 25.37978200 H

C-C_R 70.01595500 54.75939900 25.17447200 H

O-O_R 69.81596100 55.82239500 25.74857800 H

O-O_R 71.20138000 54.39410900 24.65763800 H

C-C_3 73.16494600 57.16566700 27.73282200 H

H-H_ 72.79983400 57.96605600 28.37226900 H

H-H_ 74.19924400 56.92539500 27.99524000 H

H-H_ 73.18195300 57.50034800 26.69339800 H

C-C_R 72.29092900 55.94816500 27.93401100 H

O-O_R 71.40575100 55.92444700 28.77213800 H

O-O_R 72.50156000 54.80765100 27.23193700 H

C-C_3 76.79082800 56.54233400 20.90335800 H

H-H_ 76.77015800 55.50283100 20.57383400 H

H-H_ 76.76217500 57.22910100 20.05815700 H

H-H_ 77.72284600 56.71704700 21.46334900 H

C-C_R 75.66629400 56.84902000 21.84459600 H

O-O_R 75.15710200 57.93621100 22.01749700 H

O-O_R 75.27049200 55.72506400 22.52803700 H

C-C_3 74.49357100 55.91508400 23.70874000 H

H-H_ 74.31039900 56.98458800 23.84909500 H

C-C_3 73.16268100 55.17069500 23.53824300 H

H-H_ 73.34307800 54.12611300 23.27166600 H

O-O_R 72.39941300 55.77712000 22.48288500 H

C-C_R 72.03857900 54.99518700 21.42452900 H

C-C_3 71.27609600 55.80388500 20.40767600 H

H-H_ 71.03471700 55.17499700 19.55169700 H

H-H_ 70.35438100 56.18984300 20.85463900 H

H-H_ 71.87122500 56.66556800 20.09251500 H

O-O_R 72.29779700 53.81426600 21.33322200 H

C-C_3 72.35177400 55.24155700 24.82823900 H

H-H_ 72.02315700 56.26370700 25.02836500 H

C-C_3 73.20604500 54.69790100 25.99268200 H

H-H_ 73.32995300 53.61067200 25.87471500 H

O-O_3 74.42674000 55.36347400 26.07941200 H

C-C_3 75.26985200 55.33617500 24.91011800 H

H-H_ 75.56129900 54.30144200 24.69028900 H

C-C_3 76.51797700 56.14679000 25.32863600 H

H-H_ 76.96165900 55.63426000 26.19161500 H

H-H_ 76.19489900 57.14309800 25.64770800 H

O-O_3 77.39435100 56.21208100 24.24647600 H

S-S_R 78.84690100 57.26311400 24.42377600 H

O-O_R 78.60990100 58.38499700 25.37366400 H

O-O_R 79.22841000 57.64639600 23.03707900 H

N-N_R 80.11691400 56.27664200 25.16399600 H

H-H_ 80.82567700 56.95731200 25.43589400 H

H-H_ 80.50518900 55.74503600 24.38611300 H

O-O_3 67.93249300 45.39132600 28.13031000 H

H-H_ 68.84532400 45.67879000 27.97704800 H

H-H_ 67.91600700 45.13967300 29.06754900 H

p

Paper I

Promiscuous Ability of Human Carbonic Anhydrase: QM and QM/MM Investigation of Carbon Dioxide and Carbodiimide Hydration

Paolo Piazzetta, Tiziana Marino, and Nino Russo

Inorg. Chem. **2014**, *53*, 3488–3493
DOI: 10.1021/ic402932y

Inorganic Chemistry

Article

pubs.acs.org/IC

Promiscuous Ability of Human Carbonic Anhydrase: QM and QM/MM Investigation of Carbon Dioxide and Carbodiimide Hydration

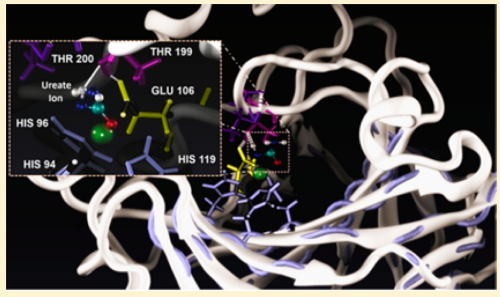
Paolo Piazzetta,[†] Tiziana Marino,*[,†] and Nino Russo[†,‡]

[†]Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Rende CS, Italy

[‡]Departamento de Quimica, Division de Ciencias Basicas e Ingenieria, Universidad, Autonoma Metropolitana-Iztapalapa, Av. San Rafael Atlixco No. 186, Col. Vicentina, CP 09340 Mexico D.F., Mexico

Supporting Information

ABSTRACT: The hydration of carbodiimide, isoelectronic with carbon dioxide, mediated by human carbonic anhydrase (EC 4.2.1.1) was studied at theoretical level in comparison with the native substrate. Quantum-mechanical (QM) and combined quantum-mechanics/molecular-mechanics (QM/MM) approaches indicate that human carbonic anhydrase is able to catalyze also the hydration of carbodiimide to urea with an energy barrier higher than that required by the native CO_2. The rate-determining step for both substrates is represented by the nucleophilic addition to the double bond, showing the final product ureate more strongly anchored to the enzyme active site than the hydrogen carbonate. The reduced catalytic activity for the carbodiimide substrate can be ascribed not only to the highest barrier but also to the difficulty in releasing the product in favor of the water molecule, delaying the catalytic turnover as indicated by QM and QM/MM analyses.



INTRODUCTION

Carbonic anhydrase (CAs) enzymes (14 iso-enzymes) are one of the most ancient and interesting enzymatic families involved in different biological processes.[1,2] Between the CAs, the human Carbonic Anhydrase II (hCAII), which catalyzes a simple but very important physiological reaction, the reversible interconversion of CO_2 to HCO_3^-, is one of the most widely studied enzyme from both experimental and theoretical points of view. In fact, the presence in the literature of thousands of scientific papers devoted to this system accounts for its importance in chemistry, biology, and other fields.[2,3] Despite this impressive amount of interest, hCAII continues to reserve surprises to the researchers since this system can be considered one of the first examples of enzymes showing broad promiscuous catalytic activity.[4−7] In fact, besides the native CO_2, it is able to process substrates with different structures (substrate promiscuity), to catalyze distinct chemical reactions (catalytic promiscuity), and to be still efficient when the native Zn^{2+} metal center is replaced with other metal ions.[8−10] As mentioned above, whereas the mechanism of CO_2 transformation in bicarbonate was extensively studied, it is still not fully clear how the enzyme catalyzes the hydration of carbodiimide (and its tautomer cyanamide) to urea. Carbodiimide is a noncompetitive inhibitor toward the esterase activity of hCAII but, at the same time, can be catalyzed by hCAII, giving rise to the urea product.[11,12] It is interesting to note that at the end of this reaction the catalytic cycle is sensibly turned down. X-ray crystallographic and spectroscopic studies[13] reveal the presence of the reaction product (ureate) in the catalytic site and the binding mode of the carbodiimide within the hCAII active site. A qualitative reaction mechanism for the hydration reaction of carbodiimide was proposed,[13] but the energetic details are still lacking; thus, the reduction of the catalytic efficiency with respect to the native substrate remains an open problem. To the best of our knowledge, while the hydration of CO_2 by hCAII was studied at the quantum-chemical level by employing a cluster model,[14−17] theoretical studies on the cognate substrates do not exist in literature. Some previous works were devoted to the hydration of carbodiimide itself in the gas phase and in solution.[18,19]

In the present work, detailed potential energy surfaces (PESs) for the reaction catalyzed by hCAII toward carbodiimide were obtained by using both a quantum-mechanical (QM) cluster model[20−24] and quantum-mechanics/molecular-mechanics (QM/MM)[25−28] approaches. Because of the lack in the literature of a PES at the QM/MM level for the hydration of native CO_2 substrate, we have studied also this reaction.

COMPUTATIONAL METHODS

For both QM and QM/MM our own *n*-layered integrated molecular orbital and molecular mechanics (ONIOM) calculations, the Gaussian03 package was used.[29] Geometry optimizations were

Received: November 26, 2013
Published: March 17, 2014

dx.doi.org/10.1021/ic402932y | *Inorg. Chem.* 2014, 53, 3488−3493

performed using the hybrid density functional B3LYP[30,31] associated with the 6-31+G (d,p) basis set. More accurate energies were then obtained by single-point 6-311+G(2d,2p) calculations carried out on the previously optimized geometries. The use of the B3LYP functional is largely established in the literature.[20−24,32] In QM calculations the effects of the protein surroundings were mediated by using the popular polarizable continuum model (PCM)[33,34] previously tested in a huge number of papers.[20−24,32,35] As requested by one of the referees, to further verify the reliability of the B3LYP functional, we redid the computation of the rate-determining barrier by using the PBE0 exchange-correlation one.[36] Results for both of the considered substrates give barriers of 4.5 and 17.7 kcal/mol for CO_2 and carbodiimide, respectively. For the rate-determining barrier universal solvation method density (SMD)[37] method was also used. Comparison shows very similar results (19.3 versus 20.8 kcal/mol at PCM and SMD levels, respectively). In both PCM and SMD, the suggested value of 4 was assumed for the dielectric constant of the protein environment.[32] In the case of ONIOM, the remaining portion of the protein, where the chemical events (formation and breaking of bonds) do not take place, is explicitly included in the MM region. To verify the nature of the minimum or maximum along the PES and to obtain the thermochemical analysis, frequency calculations were accomplished at the same level of theory used for optimizations. With the aim of assessing whether the localized transition states (TSs) correctly connect to the corresponding minima along the imaginary mode of vibration, the intrinsic reaction coordinate (IRC)[38] method was used. For TSs the vibrational mode associated with the imaginary frequency was checked. Natural bond order (NBO) analysis was performed on every stationary point of the PESs.[39]

To restore the active enzyme, water exchange with both the bicarbonate and ureate products was also investigated at the QM level. For this purpose, relaxed PES scans (with geometry optimization at each point), using 10 steps with a step size value of 0.1 Å on the Zn−O_w distance, were performed. Results are reported in Supporting Information.

For both QM and QM/MM calculations the structure of the enzyme was retrieved from the crystallographic structure.[40] The cluster for QM study was built considering the amino acid residues directly and indirectly involved in the catalytic mechanism. More specifically, it includes the Zn^{2+} cation with its four primary ligands of a water molecule and three histidine residues (His94, His96, and His119) and the hydrophilic binding cavity formed by Glu106, Thr199, and Thr200 residues belonging to the outer coordination sphere. Their presence in the catalytic pocket is strictly related to the management of the substrate in a way that is correct and suitable for establishing favorable interactions with the environment surrounding the active site. Another water molecule (so-called deep water) was incorporated into the cluster. Following a consolidating procedure,[20−24,32,35] in the QM cluster, all the amino acid residues were truncated to reduce the size of the model, and an atom for each residue was fixed during the geometry optimizations (circled in Scheme 1).[20−24,32]

To further verify the goodness of this approximation, we also optimized the first intermediate (INT1) of the carbodiimmide reaction by using the entire amino acids. Comparison between the two structures shows good agreement in the main structural parameters (see Supporting Information, Figure S1).

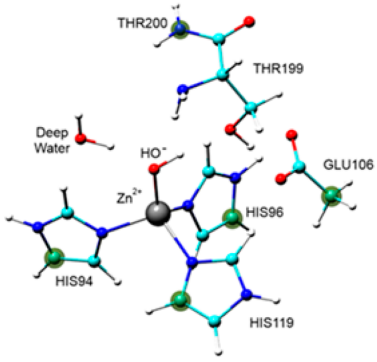
In the QM/MM computations, starting from the crystallographic structure,[40] the QM region includes the first and second coordination shells around the zinc cation (His94, His96, His119, Glu106, Thr199, and Thr200) and a deep water molecule. The remaining portion of the enzyme, along with 219 water molecules, is described at the MM level.

RESULTS AND DISCUSSION

The comparison of the PESs of the reaction mechanism for two considered substrates by means of the combined use of QM and QM/MM approaches can allow the rationalization of the "promiscuous" activity of this enzyme and could evidence some reasons for which the enzymatic turnover is reduced in the case of carbodiimide hydration. The carbodiimide tautomer was chosen as initial substrate since previous studies in the gas phase have demonstrated it to be the more stable one.[19]

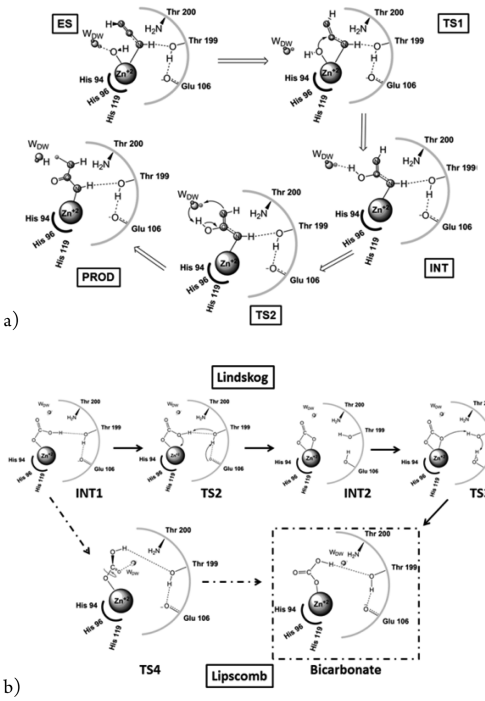
The reaction mechanism derived by the available experimental data[2,41−44] and by previous theoretical[14−16,45−48] studies on the native substrate was followed in the carbodiimide substrate (Scheme 2). In this mechanism, the carbodiimide

Scheme 1. Catalytic Site Model Used in the QM calculations[a]



[a]The circled atoms were frozen during the optimization procedure.

Scheme 2[a]



[a](a) Proposed reaction mechanism for addition of water to the carbodiimide; (b) Lipscomb and Lindskog mechanisms followed in the case of CO_2 substrate.

dx.doi.org/10.1021/ic402932y | Inorg. Chem. 2014, 53, 3488−3493

along with the hydroxyl group (OH^- arising from the water ligand) is coordinated to the zinc ion. In such a way, the nucleophile species (OH^-) is ready to perform the nucleophilic attack on the carbodiimide substrate, generating a pentacoordinated adduct. The formed intermediate suffers a proton shift at the expense of a water molecule that gives rise to the ureate anion as final product. For both substrates the activated water (Zn−OH) was used as nucleophile agent responsible for the attack on the substrate and not the neutral water as suggested by Thoms.[44] In fact, recent calculations[16] based on the use of zinc-bound water rather than zinc-bound hydroxide did not support Thoms' hypothesis, because the present hydrogen bond network is not sufficient to enhance either the electrophilic character of the carbon on CO_2 or the nucleophilic character of the metal-bound water. In the case of CO_2 hydration, both the Lipscomb mechanism, favoring a proton shift between the two oxygens,[45] and the Lindskog one,[46] favoring a rotation step, were taken into account (see Scheme 2b). In Figure 1 are shown the PESs obtained for both studied substrates at QM (solid line) and QM/MM (dotted line) theoretical levels.

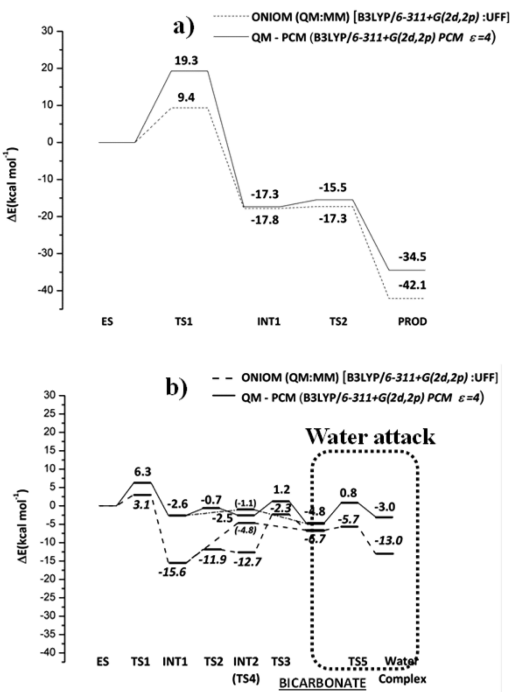


Figure 1. PESs for the hydrolysis of carbodiimide (a) and carbon dioxide (b) by hCAII at QM and QM/MM levels.

Enzyme−Substrate Complexes. In both QM and QM/MM structures of the Michaelis−Menten complex (ES), the carbodiimide substrate is coordinated to the zinc ion by nitrogen atom N1 (the Zn−N1 distance = 2.759 Å), giving a pentacoordinated complex (Figure 2). In the case of natural substrate (CO_2) the ES structure is different, since the CO_2 lies at a Zn−O1 distance of 3.63 Å. This finding agrees with the cryocrystallographic study[13] where the carbodiimide, in spite of

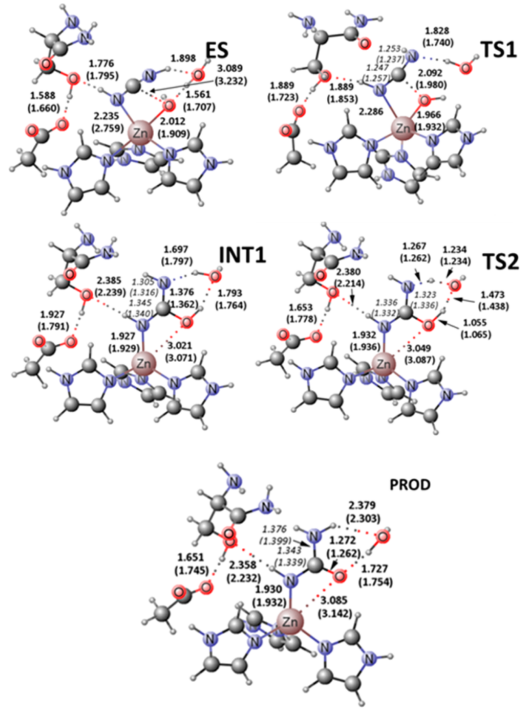


Figure 2. Optimized structures of the stationary points along the PES of carbodiimide. All bond lengths are given in Å. Values refer to the QM and the QM/MM approach, in parentheses.

owing isoelectronic and isostructural with the native substrate, binds with the active site in a different manner from CO_2. In addition, in the case of carbodiimide, the ES is stabilized by a network of hydrogen bonds with the surrounding groups (Thr199, Thr200, and the water molecule, Figure 2).

Nucleophilic Attack. The subsequent stationary point on the reaction path (Figures 1 and 2) is the transition state (TS1) for the nucleophilic attack on the carbon of two substrates. The distance of the nucleophilic OH group to the substrate carbon atom is 1.980 Å for carbodiimide, indicating that the C−O bond is starting to occur. The corresponding distance in CO_2 is 1.750 Å. The occurred activation of the linear geometries characterizing the two considered substrates is also indicated by the valence angle value that is 152.2° in carbodiimide and 146.5° in CO_2, at QM/MM level. The imaginary frequency of 161.5*i* and 321.1*i* obtained in CO_2 and carbodiimide, respectively, is clearly associated with the stretching of the C−OH moiety describing the nucleophile attack by OH group on the electrophile atom of the two considered substrates (O═C═O and HN═C═NH).

In the TS1 of carbodiimide, the Zn−N1 (nitrogen atom of the substrate) is 3.825 Å, while the corresponding value in ES is 2.286 Å. This is because the carbodiimide is rotated (about 60°) with respect to the position assumed in the ES. The hydrogen bond network present in the ES is retained, although with longer distances, and contributes to stabilize the TS1. Furthermore, in this species is observed an elongation of the two C−N bonds (C−N1 1.257 Å and C−N2 1.237 Å) (Figure

dx.doi.org/10.1021/ic402932y | Inorg. Chem. 2014, 53, 3488−3493

2). The same behavior is observed in the corresponding C−O bonds in CO_2 activation.

From the PES in Figure 1, the activation barrier (TS1) with respect to the energy of the reactants (9.4(3.1) kcal/mol at the QM/MM level and 19.3(6.3) kcal/mol at the QM level for carbodiimide (CO_2)) represents the rate-determining step (RDS) of the reaction at both levels of theory used and for both investigated substrates. The calculated barrier value for processing the carbodiimide (9.4 kcal/mol) shows that hCAII is able to hydrolyze the substrate with acceptable kinetics for an enzyme-catalyzed reaction.

Referring to CO_2, the corresponding barrier proposed by Bottoni et al.[14] in his QM cluster calculations was 0.6 kcal/mol, which became 7.6 kcal/mol in a most recent paper,[16] but in both works the TS is characterized by a longer distance (2.191 and 2.030 Å) between the oxygen hydroxyl group and carbon atom with respect to our TS geometries (1.850 and 1.750 Å at QM and QM/MM levels, respectively). For CO_2 substrate a comparison with the experimental value is possible since it is possible to extract from the measured available kinetic data (turnover number, 10^6 s^{-1}) the corresponding activation barrier (4.02 kcal/mol). If we consider the corresponding barrier computed at QM and QM/MM levels (6.3 and 3.1 kcal/mol, respectively), we can conclude that the QM/MM value is closer to the experimental data. In any case the obtained value for QM barrier gives deviations of about 2 kcal/mol, which falls in the range of error previously established for QM cluster calculations.[49]

We note that the barrier obtained at QM level is 9.9 kcal/mol higher than the corresponding QM/MM one. If we consider that in the case of natural substrate the QM barrier (TS1) is overestimated and the QM/MM one is underestimated with respect to experimental counterpart, the difference between the energy barriers computed employing the two computational protocols should decrease. Furthermore, we underline that this difference was found also in other enzymes,[28] and in our case it is essentially due to the more extended H-bond network present in QM/MM that contributes to stabilize the TS1. On the basis of our two results it seems that QM/MM gives better results, but this conclusion cannot be generalized without additional evidence that is consistent with these findings.

Lipscomb and Lindskog Mechanisms. The next stationary point suggested by the intrinsic reaction coordinate procedure is INT1, where the C−O bond is completely formed. NBO analysis shows that this bond is covalent in nature with the carbon and oxygen atoms overlapping their s(25%) p(75%) and s(32%) p(68%) orbitals for both substrates. The same analysis for carbodiimide reveals as charge values on the nucleophilic species (OH^-) suffer a drop in going from the starting complex (−1.310 |e|) to the INT1 (−0.808 |e|) as generally occurs in a nucleophilic attack. The energy stabilization of INT1 related to the Michaelis complex is 17.8 kcal/mol (Figure 1). After the INT1 species, the catalytic reaction of hCAII or the Lipscomb or Lindskog mechanism (see Scheme 2b) can follow, but in the case of HN═C═NH the Lindskog mechanism must be ruled out to make possible the ureate formation. Consequently in carbodiimide after INT1 formation the catalytic process evolves toward the ureate (PROD) throughout the Lipscomb-like TS (TS2, see Scheme 2a). In this step the assistance of a deep water molecule is indispensable for acting as proton acceptor and proton donor contemporaneously (see TS2 Figure 2). The proton shift concerning the NH group is characterized by the distances of 1.262 Å (N···H) and 1.234 Å (H···O), while the other one presents a pronounced shortening of the O(water)−H (1.438 Å compared to that present in INT1 (1.764 Å). The analysis of the calculated imaginary frequency (1097.2i cm^{-1}) properly depicts the reorganization just described. The TS2 energy (−17.3 kcal/mol) lies below the energy of reactants, which underlines that the interconversion between the two ureate tautomer forms, present in the INT1 and PROD species, can occur without a great energetic effort. The computed PESs at the QM and QM/MM levels clearly indicate the reaction is highly exothermic (−42.1 kcal/mol at the QM/MM level and −34.5 kcal/mol at the QM one), meaning that the ureate product is sensibly stabilized in the active site. Referring to the geometrical parameters, the ureate in PROD shows coordination to the Zn^{2+} by deprotonated nitrogen (N1) (1.932 Å) and by carbonyl oxygen (3.142 Å). The NBO charge values of the nitrogen (N1) show its negative increase on every point of the PES. The N1 of ureate remains engaged in hydrogen bonds with the side chain of Thr199, while the N2 with the side chain of Thr200, water, and the oxygen atom of ureate is mainly implicated in the hydrogen bond with the water molecule. This arrangement agrees well with that observed in the experimental counterpart.[11,13]

As mentioned above, the mechanism proposed for HN═C═NH differs from the CO_2 one in the number of elementary steps. In fact, in the case of CO_2, starting from INT1, the Lipscomb or Lindskog mechanism can occur with activation barriers lying below the energy of reactants (see Scheme 2b). The one-step Lindskog mechanism involves a rotation around the Zn−O1 bond with simultaneous breaking of the Zn−OH bond (TS4) giving rise to the rearranged bicarbonate fragment in the active site. The PES in this region is flat enough due to the low barriers found in both theoretical levels used (−6.7 and −4.8 kcal/mol at QM/MM and QM level, respectively, related to the ES). In the Lipscomb mechanism the rearrangement of bicarbonate takes place via a double proton transfer mechanism involving the Thr199 that acts as a proton shuttle. In particular, the TS2 describes a proton transferred from the original hydroxide oxygen to the oxygen of the OH group of the Thr199 that in turn transfers its proton to the Glu106. The energy barrier for this step is low at both levels of theory (3.7 and 1.9 kcal/mol at QM/MM and QM level, respectively, obtained with respect to the INT1 species).

The INT2-TS3-bicarbonate step describes the double proton transfer: from Glu106 to Thr199 and from Thr199 to the bicarbonate moiety. The barrier for the second proton transfer (TS3) is 10.4 and 3.7 kcal/mol at QM/MM and QM levels, respectively. On the basis of the calculated TS4 (Lindskog) and TS2/TS3 (Lipscomb) activation barriers, it is evinced that the competition between the two possible rearrangements is still feasible, contrary to what was proposed by a recent QM work on the catalytic mechanism of hCAII.[16]

Since all the calculations indicate a great stability of the ureate with respect to HCO_3^- on the active site of the enzyme, we have evaluated the BEs (enthalpy values) at the QM level. Results show that the BE for ureate is 62.9 kcal/mol, while the corresponding one for HCO_3^- is 42.6 kcal/mol. These data suggest that the release of the product is more difficult when hCAII explicates its promiscuous activity catalyzing the carbodiimide. This is not an innocent consequence because of the difficulty in restoring the enzymatic turnover.

Attack of the Water Molecule. For both native and promiscuous substrates the last step of the reaction closing the

dx.doi.org/10.1021/ic402932y | Inorg. Chem. 2014, 53, 3488−3493

catalytic cycle and describing the displacement of the products (ureate and bicarbonate anions), respectively, by water molecule was investigated. For CO_2 the PES concerning this step is evidenced in Figure 1b. In the 4-coordinated carbonate complex at both QM (QM/MM) levels the entering water lies at 1.967 (1.949) Å from the oxygen of the bicarbonate moiety and at 3.934 (3.706) Å from the Zn ion. The water attack leads to the formation of the pentacoordinated water complex throughout the TS5 located at 0.8 (−5.7) kcal/mol relative to the reactant. This TS is characterized by a water oxygen−Zn^{2+} distance of 2.608 (2.806) Å, and the HCO_3^- is now at 1.986 (1.958) Å from the Zn^{2+}; the H-bond network involving the water and neighboring atoms both of the carbonate 1.945 (1.841) and of Thr199 1.862 (2.771) is more appropriate (Figure 3). In the obtained pentacoordinated zinc complex

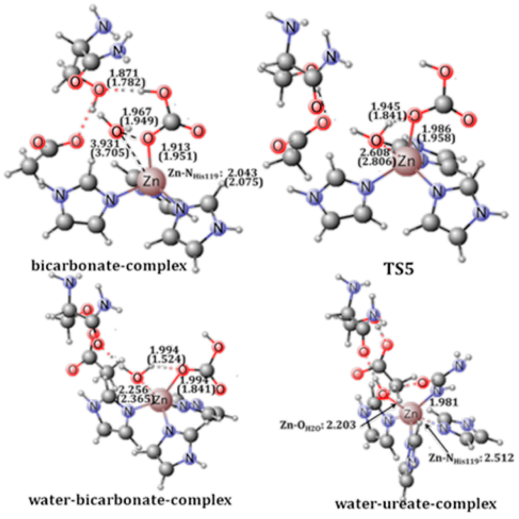


Figure 3. Optimized structures of the stationary points along the PES of CO_2 during the attack of the water molecule. Values refer to the QM and the QM/MM approach, in parentheses. Structure of the "water−ureate complex" obtained from the scan on the Zn−O_w bond at QM level. All bond lengths are given in Å.

(water complex) both the water and the bicarbonate are firmly bonded to the metal (the two O_w−Zn and O−Zn bonds are 2.256 (2.365) and 2.043 (2.011) Å, respectively (Figure 3). This intermediate lies at only 3.0 (13.0) kcal/mol above the reactants; thus, the reaction can proceed, releasing the bicarbonate and consequently restoring the catalytic cycle. A different behavior was observed in the carbodiimide. In fact, all the attempts to localize a water−ureate final complex failed at both the QM and QM/MM levels. As previously noted this intermediate is strongly due to the presence of Zn−N covalent bond as indicated by the NBO analysis. The potential energy scan on the Zn−O_w bond at the QM level for both substrates is provided in Supporting Information, Figures S2 and S3. The scan for carbodiimide reveals that as the water approaches the Zn^{2+}, the His119 is detached by the zinc coordination sphere. Figure 3 shows that at a distance of 2.203 Å (similar to that found in the case of carbonate) the Zn−NHis119 one becomes 2.512 Å. In the case of the pentacoordinated complex with the bicarbonate the corresponding Zn−NHis119 distance is 2.043 (2.075) Å at QM (QM/MM) levels. These findings suggest that the release of the reaction product in the carbodiimide reaction is more difficult. Considering that the barriers for the RDS for the two reactions (9.4 and 3.1 kcal/mol for carbodiimide and CO_2, respectively at the QM/MM level) are not so large to justify the different catalytic efficiency between the two substrates, we think that the QM and QM/MM results indicate that the release of the product should be considered for a more convincing explanation of the reduced turnover for the carbodiimide substrate.

On the other hand, as previously suggested,[50,51] enzymes reach maximum efficiency when the barrier in the RDS is lower, and when they do not reach maximum efficiency, they allow a higher stabilization of the intermediates. Our study also underlines that the stabilization of the final product should be an important step in the energetic behavior of the catalyzed reaction and more specifically in the restoring of the catalytic cycle.

CONCLUSIONS

QM and QM/MM methods were employed in studying, at atomistic level, the promiscuous enzymatic activity of hCAII toward the natural substrate (carbon dioxide) and the carbodiimide one. The comparative analysis of the PESs reveals interesting features that contribute to highlight the promiscuous activity of this enzyme. The following conclusions can be outlined:

All calculations show that, for both substrates, the rate-determining step is the nucleophilic attack, with a higher energy barrier in the case of the carbodiimide substrate. In particular, the QM/MM barrier is found to be 9.4 and 3.1 kcal/mol for carbodiimide and carbon dioxide, respectively. This result suggests a lower kinetic constant for carbodiimide catalysis.

Both levels of theory evidence a great thermodynamic stability of ureate product in the carbodiimide PES. The great difference in BE value (62.9 kcal/mol vs for 42.6 kcal/mol for HCO_3^-), obtained at the QM level, suggest that the release of the ureate product cannot easily occur.

Contrary to the case of bicarbonate product formation, for which we were able to ascertain substitution with a new water molecule, indispensable for the restoring of the catalytic cycle, for the case of ureate formation we fall in the determination of the pentacoordinated complex that is the crucial intermediate for the coordination of water and the release of the reaction product. This suggests that the release of the ureate should be more difficult, causing reduced turnover when hCAII carries on its promiscuous activity toward the carbodiimide substrate.

In addition, our QM and QM/MM comparison of both computed reaction PESs evidence quantitative differences in these two cases. In particular we underline that the QM/MM energy barrier for the rate-determining step is about half that of the corresponding QM value.

ASSOCIATED CONTENT

Supporting Information

Cartesian coordinates of all the stationary points on the PES of carbodiimide and CO_2 hydration at QM level are reported in Tables S1 and S2, respectively. The optimized QM/MM Cartesian coordinates are available upon request. The INT1 species optimized at the QM level by taking into account different truncations of the amino acid residues are displayed in Figure S1. The results of NBO calculations performed on the

dx.doi.org/10.1021/ic402932y | Inorg. Chem. 2014, 53, 3488−3493

bicarbonate and ureate products are shown in Figure S2. The relaxed scans on the Zn^{2+}–O_w distance of the bicarbonate and ureate species are illustrated in Figures S3 and S4, respectively. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tmarino@unical.it.

Author Contributions

All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was financially supported through Department of Chemistry and Chemical Technologies of the Università della Calabria (Italy). P.P. gratefully acknowledges Commissione Europea, Fondo Sociale Europeo, and Regione Calabria for the financial support.

REFERENCES

(1) Sly, W. S.; Hu, P. Y. *Annu. Rev. Biochem.* **1995**, *64*, 375–401.
(2) Christianson, D. W.; Fierke, C. A. *Acc. Chem. Res.* **1996**, *29*, 331–339.
(3) Krishnamurthy, V. M.; K.Kaufman, G.; Urbach, A. R.; Gitlin, I.; Gudisen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chem. Rev.* **2008**, *108*, 946–1051.
(4) Babtie, A. C.; Bandyopadhyay, S.; Olguin, L. F.; Hollfelder, F. *Angew. Chem., Int. Ed.* **2009**, *48*, 3692–3694.
(5) O'Brien, P. J.; Herschlag, D. *Chem. Biol.* **1999**, *6*, R91–R105.
(6) Khersonsky, O.; Roodveldt, C.; Tawfik, D. S. *Curr. Opin. Chem. Biol.* **2006**, *10*, 498–508.
(7) Olguin, L. F.; Askew, S. E.; O'Donoghue, A. C.; Hollfelder, F. *J. Am. Chem. Soc.* **2008**, *130*, 16547–16555.
(8) Kaiser, E. T.; Lo, K. W. *J. Am. Chem. Soc.* **1969**, *91*, 4912–4918.
(9) Pocker, Y.; Sarkanen, S. *Biochemistry* **1978**, *17*, 1110–1118.
(10) Cu, T.; Thomas, H. G.; Wynns, G. C.; Silverman, D. N. *J. Biol. Chem.* **1986**, *261*, 10100–10103.
(11) Briganti, F.; Mangani, S.; Scozzafava, A.; Vernaglione, G.; Supuran, C. T. *J. Biol. Inorg. Chem.* **1999**, *4*, 528–536.
(12) Scolnick, L. R.; Christianson, D. W. *Biochemistry* **1996**, *35*, 16429–16434.
(13) Guerri, A.; Briganti, F.; Scozzafava, A.; Supuran, C. T.; Magani, S. *Biochemistry* **2000**, *39*, 12391–12397.
(14) Bottoni, A.; Lanza, C. Z.; Miscione, G. P.; Spinelli, D. *J. Am. Chem. Soc.* **2004**, *126*, 1542–1550.
(15) Mauksch, M.; Bräuer, M.; Weston, J.; Anders, E. *ChemBioChem* **2001**, *2*, 190–198.
(16) Miscione, G. P.; Stenta, M.; Spinelli, D.; Anders, E.; Bottoni, A. *Theor. Chem. Acc.* **2007**, *118*, 193–201.
(17) Amata, O.; Marino, T.; Russo, N.; Toscano, M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 3468–3477.
(18) Tordini, F.; Bencini, A.; Bruschi, M.; De Gioia, L.; Zampella, G.; Fantucci, P. *J. Phys. Chem.* **2003**, *107*, 1188–1196.
(19) Nguyen, M. T.; Raspoet, G.; Vanquickenborne, L. G. *J. Chem. Soc., Perkin Trans.* **1999**, *2*, 813–820.
(20) Amata, O.; Marino, T.; Russo, N.; Toscano, M. *J. Am. Chem. Soc.* **2009**, *131*, 14804–14811.
(21) Amata, O.; Marino, T.; Russo, N.; Toscano, M. *J. Am. Chem. Soc.* **2011**, *133*, 17824–17831.
(22) Marino, T.; Russo, N.; Toscano, M. *Chem.—Eur. J.* **2013**, *19*, 2185–2192.
(23) Liao, R.-Z.; Yu, J.-G.; Himo, F. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (52), 22523–22527.
(24) Ramos, M. J.; Fernandes, P. A. *Acc. Chem. Res.* **2008**, *41* (6), 689–698.
(25) van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708–2728.
(26) Ke, Z.; Abe, S.; Ueno, T.; Morokuma, K. *J. Am. Chem. Soc.* **2012**, *134*, 15418–15429.
(27) Martí, S.; Andrés, J.; Silla, E.; Moliner, V.; Tuñón, I.; Bertrán, J. *Angew. Chem., Int. Ed.* **2007**, *46*, 286–290.
(28) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
(29) Frisch, M. J. et al. Gaussian 03, Gaussian, Inc.: Wallingford, CT, 2003.
(30) Becke, A. D. J. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(31) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(32) Siegbahn, P. E. M.; Himo, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 323–336.
(33) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
(34) Lipparini, F.; Barone, V. *J. Chem. Theory Comput.* **2011**, *7*, 3711–3724.
(35) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
(36) Van Voorhis, T.; Scuseria, G. E. *J. Chem. Phys.* **1998**, *109*, 400–410.
(37) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
(38) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154–2161.
(39) Glendening, E. D.; Reed, A. E.; Carpenter J. E.; Weinhold, F. *NBO*, Version 3.1; Gaussian, Inc.: Wallingford, CT, 2003
(40) Hakansson, K.; Wehnert, A. *J. Mol. Biol.* **1992**, *228*, 1212–1218.
(41) Bertini, I.; Luchinat, C. *Acc. Chem. Res.* **1983**, *16*, 272–279.
(42) Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30–36.
(43) Hakansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. *J. Mol. Biol.* **1992**, *227*, 1192–1204 44.
(44) Thoms, S. *J. Theor. Biol.* **2002**, *215*, 399–404.
(45) Lindskog, S. In *Zinc Enzymes;* Spiro, T. G., Ed.; Wiley: New York, 1983; pp 77–121.
(46) Lipscomb, W. N. *Annu. Rev. Biochem.* **1983**, *52*, 17–34.
(47) Merz, K. M.; Banci, L. *J. Am. Chem. Soc.* **1997**, *119*, 863–871.
(48) Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 1071–1078.
(49) Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2006**, *11*, 695–701.
(50) Lüdtke, S.; Neumann, P.; Erixon, K. M.; Leeper, F.; Kluger, R.; Ficner, R.; Tittman, K. *Nat. Chem.* **2013**, *5*, 762–767.
(51) Albery, W. J.; Knowles, J. R. *Angew. Chem., Int. Ed. Engl.* **1977**, *16*, 285–293.

Promiscuous ability of human carbonic anhydrase: QM and QM/MM investigation of carbon dioxide and carbodiimide hydration

Paolo Piazzetta, Tiziana Marino, and Nino Russo

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Rende (CS), (Italy)

Table S1. Optimized cartesian coordinates of minima and maxima points on the QM PES of carbodiimide hydration.

ES Carbodiimide

C	4.7921	-1.18225	-2.21001
N	4.81815	-2.18607	-1.25894
C	3.65668	-0.46467	-1.95682
C	3.72828	-2.04881	-0.46628
N	3.009	-1.00972	-0.86651
C	0.568469	3.10379	-2.467
N	-0.55963	2.32441	-2.62822
C	1.415519	2.359651	-1.68385
C	-0.38315	1.17995	-1.93378
N	0.809329	1.15514	-1.35397
C	-2.096	4.504449	-0.17801
C	-2.31986	2.996319	-0.16175
O	-1.80962	2.36987	0.839569
O	-2.95675	2.476859	-1.1043
C	3.139879	1.411131	2.132529
N	1.928169	1.295941	1.470319
C	3.097279	2.507901	2.952519
C	1.157499	2.29487	1.879919
N	1.835999	3.051631	2.779959
N	-4.53059	-2.08309	1.842999
C	-4.82009	-1.28511	0.636749
C	-4.39375	-2.00721	-0.65932
O	-4.58997	-1.50414	-1.76383
C	-4.18239	0.106169	0.682109
O	-2.79703	-0.03144	1.013209
N	-3.79316	-3.20584	-0.47024
Zn	1.3481	-0.30876	0.199059
N	-0.18507	-3.70469	-0.50345
C	-0.44274	-2.53564	-0.28879
N	-0.60961	-1.30838	-0.20252
O	1.71846	-1.64804	1.653939
O	1.834371	-4.21092	1.427609
H	3.25837	0.389161	-2.48179
H	3.483631	-2.69805	0.361989
H	2.401599	2.615871	-1.32574
H	-1.14834	0.42727	-1.84152
H	-2.49235	4.954739	-1.09113
H	-1.02838	4.73478	-0.08536
H	-2.5971	4.958729	0.684809
H	3.823999	2.936801	3.624359
H	0.129629	2.48083	1.561619
H	-5.3007	-2.04674	2.502269
H	-3.70632	-1.7006	2.302589
H	-5.90395	-1.15062	0.539829
H	-4.29827	0.612459	-0.28022
H	-2.38953	0.896349	0.959489
H	-3.75326	-3.5781	0.469759
H	-1.49756	-0.93847	0.211769
H	1.34772	-1.34272	2.490179
H	1.849421	-4.66865	2.275439
H	1.765811	-3.20835	1.620099
H	0.385411	-4.20199	0.202199
H	1.463539	3.862311	3.252799
H	-1.48265	2.63172	-2.91759
H	5.511721	-2.91506	-1.17641
H	-3.49925	-3.74902	-1.26798
H	-4.68281	0.700429	1.461859
H	5.55541	-1.0834	-2.96562
H	3.93314	0.694901	1.985709
H	0.667948	4.08213	-2.90993

TS1 Carbodiimide

C	-4.45669	-1.28282	2.67239
N	-4.29241	-2.43603	1.93204
C	-3.64203	-0.35348	2.0878
C	-3.40659	-2.18297	0.93864
N	-2.99762	-0.9221	1.00614
C	-0.078	2.84952	2.51055
N	1.056921	2.13358	2.21998
C	-1.11059	2.21124	1.86792
C	0.700891	1.11896	1.40947
N	-0.60994	1.12042	1.17312
C	2.081141	4.492219	-0.12389
C	2.988651	3.286179	0.13886
O	3.245211	2.545169	-0.85704
O	3.356271	3.067219	1.33216
C	-3.67851	1.83504	-1.56866
N	-2.34792	1.52465	-1.36551
C	-3.76855	2.89349	-2.4363
C	-1.64039	2.36965	-2.09641
N	-2.46439	3.21798	-2.76007
N	5.53806	-1.64951	-0.84487
C	4.66144	-1.77592	0.32979
C	3.51172	-2.77142	0.07396
O	2.70857	-3.07673	0.96517
C	4.03811	-0.42766	0.73149
O	3.27588	0.038689	-0.37185
N	3.426	-3.21894	-1.19324
Zn	-1.50787	-0.10995	-0.30875
N	0.26526	-3.36161	-0.88019
C	0.26172	-2.11004	-0.94216
N	0.5757	-0.91129	-0.80206
O	-1.63244	-1.55607	-1.6357
O	-2.31232	-4.45123	-1.02442
H	-3.46685	0.66797	2.38529
H	-3.08971	-2.93256	0.22389
H	-2.15624	2.47953	1.84032
H	1.41498	0.439	0.97762
H	2.362981	5.341159	0.50542
H	1.051751	4.21393	0.13738
H	2.105691	4.779729	-1.17809
H	-4.61401	3.42547	-2.84276
H	-0.56184	2.39097	-2.1682
H	6.50821	-1.54003	-0.56986
H	5.26409	-0.82256	-1.37395
H	5.23065	-2.17378	1.17742
H	3.40792	-0.57259	1.61969
H	3.372601	1.043339	-0.51723
H	4.1784	-2.98079	-1.82757
H	1.57889	-0.65667	-0.82665
H	-2.31963	-2.23516	-1.63209
H	-1.36045	-4.17817	-1.05785
H	-2.45064	-5.03746	-1.77717
H	0.83118	-3.65091	-0.06991
H	-2.1681	3.95265	-3.38642
H	2.037961	2.499089	2.22473
H	-4.72948	-3.32969	2.10455
H	2.66817	-3.82382	-1.4744
H	4.82884	0.289779	0.98256
H	-5.11107	-1.23199	3.52846
H	-4.46959	1.27747	-1.09191
H	-0.0596	3.73628	3.12472

INT1 Carbodiimide

C	-5.19323	0.674066	-2.30661
N	-4.9455	1.911636	-1.74089
C	-4.18725	-0.139	-1.85922
C	-3.82535	1.822187	-0.98641
N	-3.34504	0.586457	-1.03527
C	-0.35201	-3.73023	-1.75132
N	0.702652	-2.89735	-2.06684
C	-1.36245	-2.90802	-1.32527
C	0.334071	-1.63417	-1.78649
N	-0.92185	-1.59282	-1.34835
C	2.339974	-4.43186	0.81026
C	2.448612	-2.99159	0.33329
O	1.888702	-2.11548	1.07739
O	3.046282	-2.76865	-0.75013
C	-3.0894	-1.18711	2.33183
N	-1.95802	-1.06046	1.547
C	-2.78502	-1.9204	3.44771
C	-0.97514	-1.70372	2.17105
N	-1.44485	-2.23702	3.3237
N	4.857828	2.342174	1.0316
C	4.919279	1.360064	-0.06764
C	4.242478	1.887984	-1.34664
O	4.236539	1.236683	-2.38988
C	4.27518	0.023344	0.30215
O	2.93733	0.294142	0.7099
N	3.640987	3.094883	-1.20713
Zn	-1.53989	0.005979	-0.19667
N	0.541767	3.4744	0.87286
C	-0.33434	2.55343	0.57778
N	-0.16315	1.33377	0.03764
O	-1.65399	2.840259	0.8414
O	-1.08225	5.350209	1.75852
H	-4.00427	-1.17801	-2.08546
H	-3.36571	2.630547	-0.43393
H	-2.35462	-3.17109	-0.9902
H	1.011021	-0.79612	-1.86304
H	2.774984	-5.12752	0.08869
H	1.290894	-4.6937	0.99121
H	2.864534	-4.54292	1.7665
H	-3.37914	-2.23024	4.2927
H	0.052911	-1.8219	1.82276
H	5.704738	2.316675	1.58981
H	4.076628	2.101503	1.63998
H	5.966749	1.181715	-0.33689
H	4.295651	-0.66129	-0.55096
H	2.49539	-0.58631	0.85577
H	3.852027	3.630543	-0.3741
H	0.823069	1.097851	-0.0478
H	-1.68449	3.737989	1.25492
H	-0.24671	4.85092	1.47643
H	-0.97838	5.560549	2.69399
H	1.484697	3.184031	0.6297
H	-0.8877	-2.76508	3.98009
H	1.697793	-3.1457	-2.09867
H	-5.49071	2.750955	-1.87646
H	3.244047	3.537613	-2.02258
H	4.84651	-0.42838	1.13029
H	-6.03202	0.492455	-2.9595
H	-4.03052	-0.7402	2.04941
H	-0.29784	-4.80105	-1.86719

TS2 Carbodiimide

C	-5.1713	0.869749	-2.22928
N	-4.9138	2.074199	-1.59948
C	-4.18366	0.02078	-1.80976
C	-3.8055	1.93238	-0.83614
N	-3.34239	0.69296	-0.94046
C	-0.36466	-3.59513	-1.90588
N	0.66678	-2.72994	-2.20974
C	-1.37923	-2.81356	-1.41936
C	0.283029	-1.48887	-1.86109
N	-0.96237	-1.49109	-1.39176
C	2.2915	-4.48114	0.63588
C	2.39404	-3.03655	0.17521
O	1.81715	-2.17238	0.9201
O	3.00737	-2.79683	-0.89646
C	-3.13575	-1.30462	2.30215
N	-1.99246	-1.12944	1.54308
C	-2.84964	-2.11651	3.36696
C	-1.02145	-1.82165	2.13246
N	-1.50958	-2.43193	3.23829
N	5.055029	2.113862	1.13519
C	4.955449	1.249802	-0.05624
C	4.186769	1.939902	-1.19749
O	4.038119	1.406662	-2.29439
C	4.274689	-0.08734	0.23339
O	2.976239	0.200681	0.74344
N	3.663728	3.151862	-0.87033
Zn	-1.56127	0.02927	-0.13274
N	0.633218	3.445641	0.86123
C	-0.29622	2.532311	0.63542
N	-0.13076	1.295391	0.15883
O	-1.56445	2.89171	0.91672
O	-0.90618	5.19918	1.54728
H	-4.01136	-1.00818	-2.08491
H	-3.33692	2.70709	-0.24357
H	-2.35861	-3.10994	-1.07464
H	0.940539	-0.63385	-1.9156
H	2.72113	-5.16743	-0.09759
H	1.24469	-4.74683	0.82362
H	2.82581	-4.60182	1.58567
H	-3.45637	-2.48108	4.18054
H	0.00871	-1.923	1.78605
H	5.932899	1.959392	1.62037
H	4.300959	1.873472	1.77717
H	5.961779	1.056272	-0.44531
H	4.215539	-0.69352	-0.67509
H	2.496259	-0.66812	0.82381
H	4.037768	3.598712	-0.04055
H	0.851249	1.039671	0.07528
H	1.576018	3.197721	0.58489
H	-0.96544	-3.01012	3.8625
H	1.66461	-2.96714	-2.25772
H	-5.44359	2.928089	-1.70133
H	3.262678	3.709961	-1.61083
H	4.872139	-0.63559	0.98072
H	-6.00394	0.732959	-2.90073
H	-4.07019	-0.83206	2.04017
H	-0.29345	-4.65851	-2.07144
H	0.078408	4.511331	1.26399
H	-1.51247	3.8864	1.26412
H	-0.88673	5.48783	2.46776

PROD Carbodiimide

C	-5.22576	0.677858	-2.13734
N	-4.97548	1.887778	-1.51373
C	-4.21513	-0.1533	-1.73811
C	-3.84876	1.768408	-0.77385
N	-3.36729	0.535958	-0.88882
C	-0.23499	-3.58998	-1.99146
N	0.78033	-2.68878	-2.23909
C	-1.27506	-2.85404	-1.48759
C	0.360439	-1.47235	-1.84595
N	-0.8907	-1.52545	-1.39715
C	2.49364	-4.42533	0.48997
C	2.55688	-2.94605	0.14092
O	1.988	-2.15342	0.96793
O	3.13341	-2.61237	-0.92554
C	-3.03392	-1.51766	2.31021
N	-1.90705	-1.26761	1.54801
C	-2.7001	-2.33952	3.35348
C	-0.89894	-1.92493	2.11408
N	-1.34722	-2.58504	3.20844
N	4.916338	2.348891	1.20501
C	4.897219	1.464091	0.02484
C	4.122978	2.09734	-1.14456
O	4.030799	1.5494	-2.23781
C	4.296159	0.089431	0.31584
O	2.976389	0.29627	0.80811
N	3.527388	3.29186	-0.8482
Zn	-1.55462	-0.04775	-0.11146
N	0.596258	3.38182	0.8101
C	-0.44357	2.509789	0.58307
N	-0.16461	1.257689	0.18531
O	-1.63679	2.905009	0.77732
O	-1.66511	5.545329	1.38226
H	-4.03291	-1.18014	-2.01501
H	-3.36298	2.546348	-0.19359
H	-2.25159	-3.18834	-1.17059
H	0.996709	-0.59984	-1.85245
H	2.91631	-5.04313	-0.30576
H	1.45784	-4.72714	0.68408
H	3.05625	-4.60795	1.41315
H	-3.28182	-2.7528	4.16204
H	0.132979	-1.96981	1.76201
H	5.782808	2.238251	1.72188
H	4.152888	2.080111	1.82476
H	5.922909	1.328631	-0.33791
H	4.286479	-0.52524	-0.58903
H	2.553899	-0.60441	0.88326
H	3.903388	3.77544	-0.03821
H	0.829049	1.03478	0.17157
H	1.507108	3.17819	0.4168
H	-0.76972	-3.15071	3.81387
H	1.78329	-2.89362	-2.30026
H	-5.52155	2.732398	-1.60471
H	3.179418	3.83113	-1.62978
H	4.918009	-0.41424	1.07473
H	-6.07074	0.525298	-2.78979
H	-3.99235	-1.08536	2.06573
H	-0.13617	-4.64315	-2.20204
H	0.345608	4.354289	0.9368
H	-1.75873	4.584819	1.1592
H	-1.80345	5.604699	2.33455

Table S2. Optimized cartesian coordinates of minima and maxima points on the QM PES of CO_2 hydration.

ES Carbon Dioxide

C	4.87878	-0.99327	-2.18473
N	4.20477	-2.19746	-2.26937
C	4.12442	-0.20896	-1.36102
C	3.07981	-2.11653	-1.52061
N	3.00906	-0.9169	-0.95552
C	0.90997	3.70225	-1.86615
N	-0.27649	3.08147	-2.20458
C	1.70054	2.72267	-1.32713
C	-0.19586	1.79079	-1.82844
N	0.99704	1.52857	-1.30626
C	-1.97189	4.78072	0.40784
C	-2.25774	3.34214	-0.02252
O	-1.87482	2.42746	0.79057
O	-2.77457	3.15627	-1.14937
C	2.85088	0.78555	2.46609
N	1.76029	0.86874	1.62063
C	2.64181	1.60139	3.54652
C	0.89638	1.7173	2.16976
N	1.40275	2.1799	3.33789
N	-3.15229	-2.8765	0.73847
C	-4.17818	-1.82857	0.60778
C	-4.76176	-1.75288	-0.81945
O	-5.67401	-0.97577	-1.09065
C	-3.71034	-0.43557	1.05311
O	-2.52137	-0.09583	0.35472
N	-4.20819	-2.60997	-1.7132
Zn	1.30535	-0.13495	-0.12935
O	-0.05314	-1.42688	0.02462
O	0.49013	-3.6646	-1.21474
H	5.81066	-0.81466	-2.69738
H	4.30493	0.80626	-1.04421
H	2.33852	-2.90647	-1.43326
H	1.08873	4.74957	-2.05183
H	2.70726	2.80122	-0.9463
H	-1.02997	1.10738	-1.88993
H	-2.46008	5.49635	-0.25769
H	-0.8886	4.95563	0.39125
H	-2.30827	4.94995	1.43633
H	3.68759	0.13793	2.25302
H	3.23978	1.81166	4.41905
H	-0.07104	2.01433	1.75889
H	-3.22114	-3.34917	1.63336
H	-2.21779	-2.47891	0.68462
H	-5.03495	-2.09925	1.23728
H	-4.5042	0.28928	0.84027
H	-3.52854	-0.45606	2.14046
H	-2.337	0.88414	0.46728
H	-4.58403	-2.65294	-2.64797
H	-3.54484	-3.29238	-1.37206
H	-0.92788	-0.99775	0.11827
H	-0.21468	-4.32016	-1.17645
H	0.13574	-2.8247	-0.78997
H	-1.19238	3.50987	-2.35888
H	4.48469	-3.00392	-2.80873
H	0.933	2.8303	3.95145
C	0.86311	-3.18875	2.05019
O	1.9486	-2.76847	1.92062
O	-0.19131	-3.65385	2.2467

TS1 Carbon Dioxide

C	4.86418	-1.3075	-2.08019
N	4.5213	-2.50431	-1.47461
C	3.9553	-0.39573	-1.63064
C	3.43774	-2.29487	-0.69483
N	3.07755	-1.02125	-0.76386
C	0.90634	3.40787	-2.11487
N	-0.25178	2.73486	-2.45083
C	1.68968	2.49795	-1.45968
C	-0.16598	1.4835	-1.96253
N	1.00797	1.2934	-1.37119
C	-1.93874	4.68936	0.09939
C	-2.16039	3.21608	-0.23346
O	-1.69532	2.37416	0.61767
O	-2.71505	2.92891	-1.31862
C	2.91065	0.87837	2.42677
N	1.7915	0.91812	1.61696
C	2.75632	1.78363	3.44218
C	0.96444	1.82947	2.12114
N	1.52233	2.37228	3.22878
N	-3.69661	-2.68378	1.28708
C	-4.41979	-1.53588	0.72213
C	-4.70548	-1.70844	-0.78587
O	-5.34606	-0.87159	-1.41614
C	-3.77199	-0.17297	0.9833
O	-2.44266	-0.14256	0.4687
N	-4.21626	-2.85361	-1.3292
Zn	1.36792	-0.20581	-0.03502
O	-0.04298	-1.49241	0.23889
O	0.64314	-3.33057	-1.73658
H	5.6973	-1.222	-2.75946
H	3.86157	0.64758	-1.88526
H	2.93001	-3.05102	-0.11572
H	1.07176	4.44062	-2.37823
H	2.67145	2.63376	-1.03182
H	-0.98335	0.77893	-2.00116
H	-2.46531	5.33598	-0.606
H	-0.86545	4.9147	0.06341
H	-2.27765	4.90825	1.11777
H	3.72492	0.19617	2.23787
H	3.38771	2.0481	4.2754
H	-0.00504	2.11591	1.70626
H	-4.05263	-2.92701	2.2051
H	-2.70148	-2.49933	1.39725
H	-5.41746	-1.49713	1.17822
H	-4.37459	0.60258	0.49795
H	-3.76146	0.0071	2.06972
H	-2.1602	0.82182	0.4522
H	-4.47677	-3.08797	-2.27519
H	-3.84562	-3.54788	-0.69139
H	-0.954	-1.12279	0.34653
H	0.24009	-2.71804	-1.08275
H	-0.08833	-3.83398	-2.11167
H	-1.16393	3.13898	-2.668
H	4.96968	-3.39764	-1.61741
H	1.08904	3.08056	3.80374
C	0.36955	-2.5226	1.72246
O	-0.66643	-2.86327	2.21101
O	1.57404	-2.53658	1.73373

INT1 Carbon Dioxide

C	5.02803	-2.25681	-1.5559
N	4.47161	-3.2757	-0.80486
C	4.29707	-1.14124	-1.26733
C	3.43535	-2.76417	-0.10005
N	3.31287	-1.47027	-0.35918
C	1.31278	2.50812	-2.73625
N	0.07607	1.8981	-2.82314
C	2.02842	1.75992	-1.84214
C	0.05844	0.84418	-1.98695
N	1.23295	0.71937	-1.37888
C	-1.36137	4.50239	-0.86247
C	-1.81444	3.07054	-0.58306
O	-1.54458	2.61699	0.58902
O	-2.3954	2.4425	-1.49688
C	3.3925	1.16492	2.25507
N	2.19792	1.01373	1.56757
C	3.3127	2.28613	3.0376
C	1.40631	2.02214	1.91917
N	2.05326	2.81062	2.80976
N	-4.80225	-1.49388	1.93494
C	-4.90531	-0.82877	0.62957
C	-4.35538	-1.72815	-0.49663
O	-4.21074	-1.29505	-1.64731
C	-4.24984	0.55887	0.57292
O	-2.92979	0.47061	1.10949
N	-4.0748	-2.99694	-0.14178
Zn	1.61086	-0.45283	0.23902
O	-0.85911	-0.91822	0.44413
O	-1.55211	-2.05788	-2.3118
H	5.86961	-2.41852	-2.21083
H	4.40233	-0.14376	-1.66349
H	2.78094	-3.32153	0.55562
H	1.56598	3.39221	-3.29929
H	3.03717	1.91254	-1.48995
H	-0.80507	0.21584	-1.84526
H	-1.61465	4.80848	-1.88042
H	-0.28032	4.5979	-0.70599
H	-1.8457	5.18593	-0.15596
H	4.20269	0.46008	2.14897
H	4.01411	2.74209	3.71816
H	0.39879	2.21618	1.54834
H	-5.50569	-1.14216	2.57656
H	-3.88831	-1.30242	2.34431
H	-5.96905	-0.69924	0.39058
H	-4.22101	0.92599	-0.45777
H	-4.8427	1.25199	1.1882
H	-2.42431	1.32862	0.90671
H	-3.62252	-3.60621	-0.80799
H	-4.13107	-3.24329	0.83911
H	-1.76051	-0.54048	0.70701
H	-2.49653	-1.8046	-2.32836
H	-1.28613	-1.89846	-1.39319
H	-0.78601	2.28464	-3.19596
H	4.76562	-4.24187	-0.79844
H	1.66093	3.6374	3.23753
C	-0.39735	-1.89122	1.31691
O	0.89515	-2.00759	1.1941
O	-1.14661	-2.51247	2.05269

TS2 Carbon Dioxide

C	4.95856	-2.15566	-1.60481
N	4.48692	-3.16426	-0.78248
C	4.12051	-1.09393	-1.39795
C	3.40064	-2.69856	-0.11821
N	3.16034	-1.44409	-0.4697
C	1.25642	2.64843	-2.6607
N	0.26078	1.80941	-3.11769
C	1.82524	1.99944	-1.60007
C	0.22314	0.71494	-2.32929
N	1.17163	0.79379	-1.40201
C	-1.47447	4.54568	-0.76808
C	-1.78311	3.0777	-0.52364
O	-1.58983	2.67996	0.70734
O	-2.20334	2.36516	-1.44169
C	3.20898	1.08016	2.31629
N	2.02793	0.98514	1.59823
C	3.13208	2.14702	3.17116
C	1.2455	1.97754	2.00635
N	1.88337	2.701	2.95808
N	-4.19263	-1.81593	1.92224
C	-4.5219	-0.93566	0.80328
C	-4.26746	-1.64499	-0.54194
O	-4.1981	-1.01969	-1.60746
C	-3.90604	0.47411	0.81503
O	-2.51345	0.48415	1.15622
N	-4.17533	-2.98674	-0.46912
Zn	1.40432	-0.47375	0.21829
O	-0.74799	-0.95339	0.38311
O	-1.55444	-1.80422	-2.32599
H	5.81665	-2.28664	-2.24493
H	4.13291	-0.11923	-1.86129
H	2.79654	-3.26104	0.58081
H	1.46514	3.60339	-3.11627
H	2.63566	2.32063	-0.96364
H	-0.48747	-0.0986	-2.44992
H	-1.50127	4.77387	-1.83532
H	-0.49956	4.81604	-0.3494
H	-2.22531	5.1594	-0.25701
H	4.01348	0.37408	2.17864
H	3.82743	2.54553	3.89278
H	0.25128	2.21302	1.63277
H	-4.5655	-1.45239	2.79353
H	-3.17886	-1.93305	2.02914
H	-5.61063	-0.77438	0.81259
H	-4.05136	0.93966	-0.1653
H	-4.41801	1.06948	1.58128
H	-2.07882	1.51759	0.92178
H	-3.92617	-3.5066	-1.2978
H	-4.1326	-3.4139	0.44867
H	-1.74544	-0.28943	0.72482
H	-2.48424	-1.50782	-2.35961
H	-1.32849	-1.73623	-1.37957
H	-0.46418	2.04499	-3.78079
H	4.86651	-4.09663	-0.7035
H	1.4912	3.50004	3.43531
C	-0.3971	-1.95181	1.26127
O	0.89792	-2.10316	1.29581
O	-1.24269	-2.57932	1.90124

INT2 Carbon Dioxide

C	5.02729	-2.04365	-1.62168
N	4.56402	-3.06413	-0.80826
C	4.15517	-1.00462	-1.43114
C	3.45205	-2.62717	-0.16582
N	3.18459	-1.37933	-0.52226
C	1.30481	2.75807	-2.61504
N	0.34776	1.90528	-3.12715
C	1.83342	2.10721	-1.53363
C	0.30056	0.79606	-2.35555
N	1.19901	0.88715	-1.38166
C	-1.47958	4.57812	-0.72456
C	-1.89289	3.14401	-0.49464
O	-1.74063	2.76677	0.77855
O	-2.32243	2.42083	-1.37826
C	3.17206	1.06462	2.35367
N	1.9839	0.95578	1.65281
C	3.05479	2.04772	3.29996
C	1.15707	1.85706	2.1635
N	1.77026	2.53909	3.16201
N	-4.108	-2.0748	1.8149
C	-4.51616	-1.10058	0.80863
C	-4.25358	-1.644	-0.611
O	-4.2121	-0.89891	-1.59843
C	-3.97828	0.33629	0.95399
O	-2.5855	0.4306	1.2743
N	-4.11966	-2.98167	-0.69339
Zn	1.381	-0.4827	0.20534
O	-0.6434	-1.0203	0.2893
O	-1.53781	-1.62316	-2.35561
H	5.90138	-2.15173	-2.24424
H	4.15222	-0.02885	-1.89282
H	2.84632	-3.20103	0.5238
H	1.5214	3.71864	-3.05541
H	2.60663	2.4354	-0.85537
H	-0.39179	-0.0304	-2.51363
H	-1.62196	4.84441	-1.77235
H	-0.43094	4.72018	-0.44413
H	-2.07828	5.24317	-0.09336
H	4.0137	0.42337	2.1406
H	3.74299	2.42469	4.03976
H	0.14475	2.05568	1.83681
H	-4.44612	-1.81504	2.73597
H	-3.08591	-2.181	1.85884
H	-5.61195	-1.00206	0.85688
H	-4.19147	0.88303	0.02807
H	-4.50004	0.82	1.78778
H	-2.09181	1.80189	0.92864
H	-3.81491	-3.38255	-1.56879
H	-4.01807	-3.50077	0.17112
H	-1.97498	-0.21589	0.79697
H	-2.46088	-1.30831	-2.36743
H	-1.29006	-1.58987	-1.40649
H	-0.30867	2.11094	-3.86665
H	4.96538	-3.98645	-0.72028
H	1.34308	3.26684	3.71662
C	-0.32742	-2.04717	1.12615
O	0.96919	-2.16176	1.2886
O	-1.19719	-2.75366	1.65444

TS3 Carbon Dioxide

C	5.070627	-1.87843	-1.83425
N	4.91076	-2.89281	-0.90427
C	4.036286	-1.0163	-1.62303
C	3.805555	-2.62298	-0.16746
N	3.263177	-1.48953	-0.58149
C	1.293244	2.959941	-2.30498
N	0.207349	2.205725	-2.69407
C	2.090823	2.124208	-1.57438
C	0.349138	0.967004	-2.18307
N	1.490445	0.878707	-1.50141
C	-1.49185	4.54902	-0.21754
C	-1.94872	3.10049	-0.3951
O	-1.85573	2.375722	0.691153
O	-2.35381	2.702658	-1.49194
C	3.220521	0.789982	2.450991
N	2.118992	0.764504	1.612456
C	2.981695	1.664217	3.478041
C	1.224082	1.602088	2.119966
N	1.714905	2.167624	3.248674
N	-4.42868	-2.35445	1.631951
C	-4.79165	-1.30261	0.671579
C	-4.4255	-1.71218	-0.76919
O	-4.47829	-0.90149	-1.69934
C	-4.23861	0.106057	0.969543
O	-2.82491	0.162569	0.866734
N	-4.05829	-3.00215	-0.93502
Zn	1.607633	-0.52449	0.062098
O	-0.26196	-0.9898	0.093064
O	-1.65911	-1.36773	-2.34923
H	5.879382	-1.87572	-2.5475
H	3.775865	-0.11221	-2.14875
H	3.418826	-3.24358	0.62873
H	1.399967	3.998542	-2.57512
H	3.030153	2.33814	-1.08631
H	-0.39547	0.183506	-2.30461
H	-1.59364	5.097873	-1.15529
H	-0.45037	4.580498	0.120962
H	-2.09691	5.034736	0.555524
H	4.089806	0.177428	2.267751
H	3.575834	1.958967	4.32837
H	0.241134	1.816192	1.711451
H	-4.96053	-2.25853	2.490962
H	-3.43686	-2.31994	1.876169
H	-5.88715	-1.21483	0.678947
H	-4.65843	0.810203	0.243611
H	-4.56986	0.396612	1.979977
H	-2.34454	1.274599	0.698659
H	-3.6753	-3.28798	-1.82439
H	-3.99334	-3.59423	-0.11761
H	-2.17831	-0.68491	1.369481
H	-2.57452	-1.03461	-2.35185
H	-1.38567	-1.35183	-1.41023
H	-0.68637	2.559075	-3.02198
H	5.498653	-3.70776	-0.80295
H	1.215575	2.823678	3.831727
C	-0.31598	-1.58077	1.281752
O	0.755485	-2.0469	1.771813
O	-1.46611	-1.56891	1.886041

TS4 Carbon Dioxide

C	-4.71153	-2.84429	1.389829
N	-3.58813	-3.64711	1.483588
C	-4.33179	-1.78555	0.611968
C	-2.57813	-3.07439	0.78854
N	-3.00556	-1.93879	0.24263
C	-1.13878	1.928388	2.925375
N	0.149457	1.55936	2.621887
C	-1.95607	1.212704	2.090234
C	0.095846	0.669141	1.616835
N	-1.16834	0.426744	1.264869
C	1.019522	4.508567	1.102294
C	2.058738	3.381552	1.054648
O	2.474267	3.038018	-0.0893
O	2.378766	2.836275	2.156848
C	-4.01477	1.413269	-1.81867
N	-2.79147	1.204019	-1.21934
C	-4.03144	2.648803	-2.41182
C	-2.06299	2.288369	-1.45616
N	-2.78401	3.187983	-2.16506
N	5.879278	-0.61364	-0.88743
C	4.946204	-1.06537	0.156284
C	4.084815	-2.25537	-0.30856
O	3.484825	-2.96721	0.502086
C	4.013784	0.074265	0.604697
O	3.456133	0.678979	-0.55703
N	4.004353	-2.39353	-1.64979
Zn	-1.75536	-0.48266	-0.45621
O	1.55658	-0.16071	-2.30424
O	0.476766	-3.73106	0.508497
H	-5.64749	-3.09638	1.861811
H	-4.91596	-0.93259	0.302578
H	-1.56882	-3.48009	0.724917
H	-1.36605	2.646398	3.698537
H	-3.0335	1.214085	2.022733
H	0.958	0.228688	1.142434
H	1.024675	5.08754	0.175036
H	1.20088	5.168526	1.955916
H	0.022925	4.068202	1.235143
H	-4.7952	0.668653	-1.80184
H	-4.78946	3.172158	-2.97238
H	-1.03972	2.45574	-1.14617
H	6.833096	-0.56525	-0.54888
H	5.598556	0.304488	-1.22459
H	5.488724	-1.42464	1.03726
H	3.236859	-0.33274	1.263102
H	4.588582	0.819291	1.16872
H	3.128118	1.611528	-0.33196
H	3.459526	-3.14783	-2.0408
H	4.662764	-1.87575	-2.21804
H	2.345368	-0.05102	-1.70338
H	1.371234	-4.00895	0.757201
H	0.636306	-2.84169	0.126426
H	1.040703	2.087598	2.802676
H	-3.51359	-4.51697	1.99197
H	-2.44537	4.086717	-2.47844
C	0.556623	-0.65805	-1.58435
O	0.754334	-1.18572	-0.4597
O	-0.63038	-0.51026	-2.07216

BICARBONATE

C	4.927313	-1.96829	-1.85702
N	4.716969	-2.97284	-0.92551
C	3.991147	-1.01585	-1.59033
C	3.677751	-2.61116	-0.13531
N	3.227117	-1.4265	-0.51552
C	1.085681	2.797886	-2.50997
N	-0.04342	2.034826	-2.72706
C	1.930333	2.027092	-1.76584
C	0.120998	0.857398	-2.09927
N	1.315587	0.811018	-1.51428
C	-1.58401	4.549121	-0.40162
C	-1.85246	3.039918	-0.42829
O	-1.46276	2.375146	0.595842
O	-2.42893	2.566155	-1.43774
C	3.312214	1.049857	2.288946
N	2.14525	0.930501	1.553297
C	3.173813	2.084964	3.174188
C	1.308745	1.875044	1.975643
N	1.904617	2.590806	2.957705
N	-5.19386	-1.53611	1.602962
C	-4.9703	-0.84456	0.321218
C	-4.30434	-1.77083	-0.71378
O	-3.97195	-1.34994	-1.82816
C	-4.15496	0.444216	0.452594
O	-2.91372	0.14578	1.093876
N	-4.12146	-3.04538	-0.31797
Zn	1.607611	-0.38881	0.088541
O	-1.32685	-1.97643	2.123884
O	-1.22182	-2.02317	-2.10242
H	5.698264	-2.03405	-2.60821
H	3.79297	-0.09088	-2.10693
H	3.262422	-3.20222	0.669987
H	1.184217	3.800835	-2.89475
H	2.910823	2.267996	-1.38296
H	-0.61581	0.063994	-2.08547
H	-1.80905	5.008784	-1.36684
H	-0.54396	4.752715	-0.12214
H	-2.21814	5.013859	0.36272
H	4.148023	0.383164	2.143755
H	3.839037	2.48975	3.920194
H	0.308068	2.075945	1.58395
H	-6.06512	-1.23993	2.029962
H	-4.44271	-1.30854	2.251389
H	-5.94311	-0.57996	-0.11176
H	-3.97396	0.882016	-0.53346
H	-4.7296	1.159984	1.061657
H	-2.29826	0.907434	0.89577
H	-3.64787	-3.69556	-0.92764
H	-4.41365	-3.3041	0.616219
H	-1.91282	-1.32937	1.660928
H	-2.17516	-1.81703	-2.1647
H	-0.98959	-1.89653	-1.16552
H	-0.97672	2.38077	-2.95247
H	5.228774	-3.84103	-0.85873
H	1.472462	3.35881	3.451347
C	-0.10766	-1.92498	1.554368
O	0.864912	-2.48622	2.061595
O	-0.07627	-1.22454	0.441733

TS5 Carbon Dioxide

C	4.70339	-0.96345	-2.78247
N	5.04971	-1.7719	-1.71493
C	3.4999	-0.40885	-2.45176
C	4.07385	-1.68886	-0.78057
N	3.12431	-0.86453	-1.20308
C	0.26267	3.25232	-2.14358
N	-0.89478	2.49771	-2.17974
C	1.22855	2.42339	-1.63279
C	-0.62251	1.28169	-1.6713
N	0.66219	1.18765	-1.3415
C	-2.2644	4.26931	0.5432
C	-2.60901	2.82196	0.20707
O	-2.17761	1.93319	1.00806
O	-3.2368	2.61511	-0.87093
C	3.35488	1.13618	1.97986
N	2.12628	1.00547	1.35445
C	3.22802	1.99945	3.03589
C	1.26538	1.76591	2.0192
N	1.89962	2.38629	3.0438
N	-5.81574	-2.20942	0.77349
C	-5.13265	-1.47998	-0.31221
C	-3.89415	-2.26225	-0.78638
O	-3.05464	-1.76162	-1.54832
C	-4.75146	-0.04992	0.07488
O	-3.96382	-0.10855	1.25361
N	-3.83101	-3.53377	-0.33982
Zn	1.45034	-0.29583	-0.1831
O	0.997	-2.96879	2.89065
O	-0.24219	-1.87645	-1.38346
H	5.3222	-0.87075	-3.66091
H	2.87836	0.26703	-3.01834
H	4.04907	-2.20212	0.17371
H	0.30063	4.27442	-2.48536
H	2.27201	2.62786	-1.44588
H	-1.37568	0.52538	-1.51575
H	-2.33668	4.44751	1.6208
H	-2.91404	4.96602	0.0074
H	-1.22566	4.46901	0.24904
H	4.22215	0.58769	1.64862
H	3.93658	2.35411	3.76746
H	0.20644	1.87959	1.79826
H	-6.80579	-1.98338	0.78105
H	-5.42637	-1.88538	1.65884
H	-5.80316	-1.44301	-1.1818
H	-4.21503	0.4465	-0.73712
H	-5.68117	0.51487	0.25779
H	-3.27016	0.59654	1.19558
H	-3.08149	-4.13922	-0.63848
H	-4.57039	-3.84893	0.27812
H	0.04458	-2.82475	2.78268
H	-1.20552	-1.855	-1.56831
H	-0.1678	-2.29036	-0.50841
H	-1.86935	2.82566	-2.19673
H	5.87886	-2.34348	-1.6424
H	1.4602	3.00823	3.707
C	1.62289	-2.27656	1.89432
O	2.85301	-2.26651	1.86457
O	0.80389	-1.69559	1.06796

W-COMPLEX Carbon Dioxide

C	4.44711	-0.99773	-3.11875
N	5.01894	-1.63454	-2.03289
C	3.23826	-0.53472	-2.68117
C	4.16729	-1.54187	-0.98472
N	3.08048	-0.87528	-1.35144
C	-0.09073	3.04971	-2.15593
N	-1.23936	2.34623	-1.86383
C	0.96403	2.23571	-1.81343
C	-0.86318	1.16274	-1.3519
N	0.46779	1.04568	-1.30062
C	-2.35515	4.07502	0.75275
C	-3.39746	3.02812	0.32549
O	-4.05004	2.44952	1.22129
O	-3.52007	2.81943	-0.93561
C	3.52814	1.24894	1.67861
N	2.28659	1.21996	1.07478
C	3.62532	2.36718	2.46473
C	1.6373	2.29708	1.48772
N	2.41398	3.02082	2.33061
N	-5.99512	-2.60684	0.47911
C	-5.05322	-2.0135	-0.48422
C	-3.66184	-2.66235	-0.34419
O	-2.66699	-2.17178	-0.89989
C	-4.97579	-0.48341	-0.38912
O	-4.73729	-0.15443	0.96276
N	-3.64026	-3.81865	0.35195
Zn	1.48067	-0.35792	-0.17267
O	2.01415	-2.56545	3.3564
O	0.05596	-2.00267	-0.76932
H	4.93377	-0.94457	-4.07975
H	2.47564	0.00442	-3.22163
H	4.3313	-1.93533	0.01252
H	-0.11444	4.04192	-2.58002
H	2.02505	2.42416	-1.88601
H	-1.56831	0.41844	-1.00849
H	-2.66213	5.0637	0.39261
H	-1.38145	3.85974	0.29946
H	-2.26307	4.10396	1.84076
H	4.23424	0.44395	1.55014
H	4.41574	2.73982	3.0967
H	0.63452	2.58051	1.20567
H	-6.93997	-2.59374	0.10742
H	-5.99109	-2.02178	1.31509
H	-5.37762	-2.277	-1.5002
H	-4.18765	-0.10754	-1.0513
H	-5.93473	-0.0676	-0.74007
H	-4.56007	0.81987	1.03959
H	-2.77752	-4.32634	0.47355
H	-4.50672	-4.10719	0.79358
H	1.04934	-2.50459	3.41821
H	-0.92099	-2.04587	-0.89682
H	0.22876	-2.4694	0.06434
H	-2.27004	2.70168	-1.64662
H	5.91443	-2.10057	-2.01631
H	2.14093	3.87776	2.78968
C	2.37566	-1.99452	2.16652
O	3.5778	-1.91668	1.90401
O	1.37479	-1.60054	1.44534

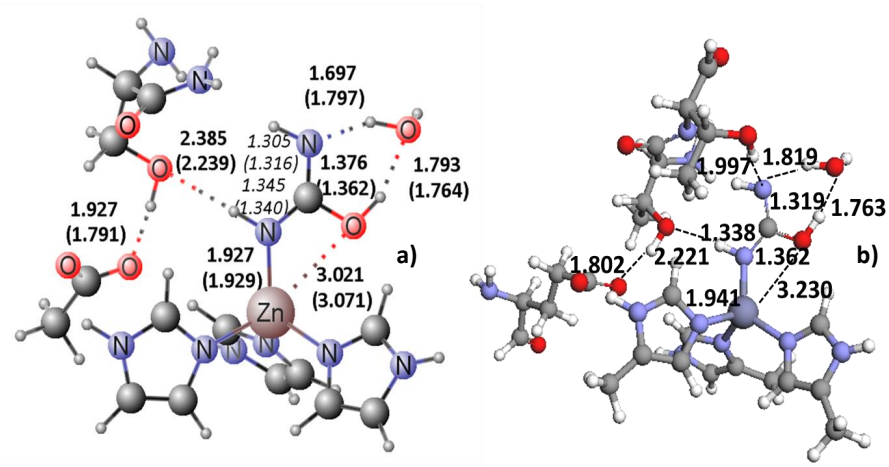
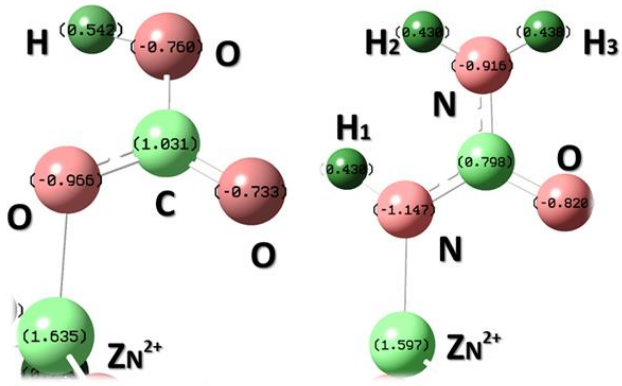


Figure S1 INT1 species (a) optimized at QM level compared with INT1 (b) optimized at the same level including entire amino acid residues. Main reported distances are in Å.



BICARBONATE	
Zn^{2+}	1.635
Zn^{2+} - **O** - C	-0.966
O - **C** = O	1.031
O - C = **O**	-0.733
O - C - **O**H	-0.760
O - C - O**H**	0.542

UREATE (PROD)	
Zn^{2+}	1.597
Zn^{2+} - **N***H1* - C	-1.147
Zn^{2+} - N***H1*** - C	0.430
NH*1* - **C** = O	0.798
NH*1* - C = **O**	-0.820
NH*1* - C = H*2* **N**H*3*	-0.916
NH*1* - C = ***H 2*** NH*3*	0.430
NH*1* - C = H*2* N***H3***	0.438

Figure S2 NBO atomic charge values in |e| for the bicarbonate and ureate species**.**

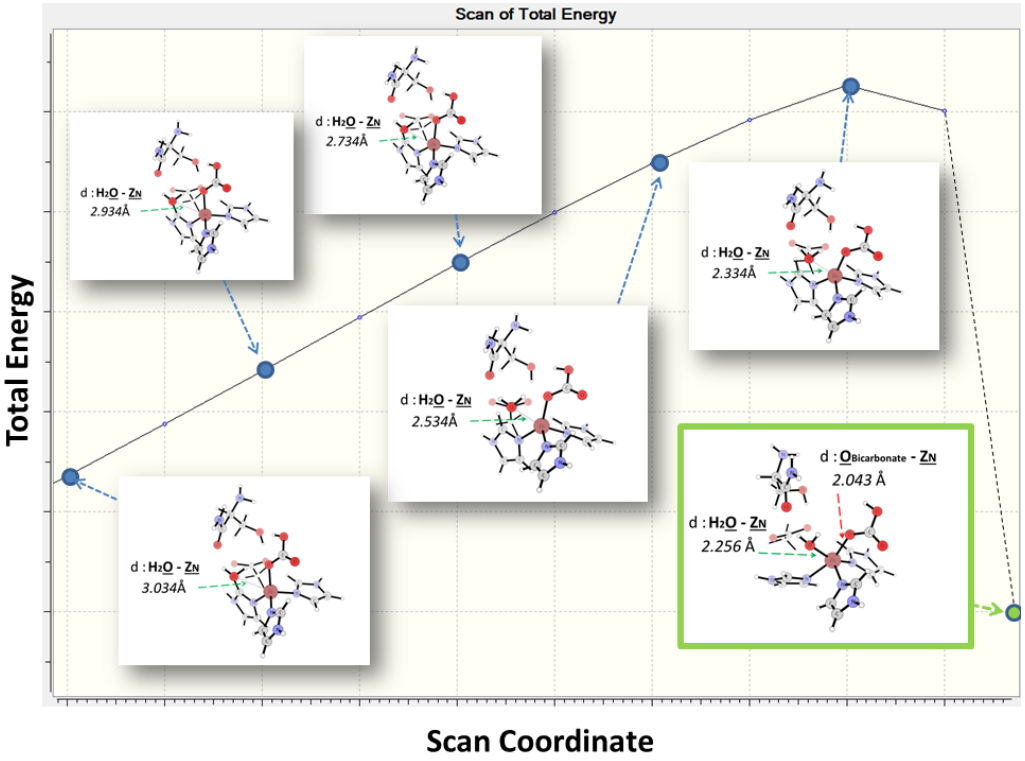


Figure S3 Relaxed potential energy scan on the Zn-Ow bond and relative geometries of the last six points of the 10 analyzed as far as the CO_2 is concerned. Distance are in Å.

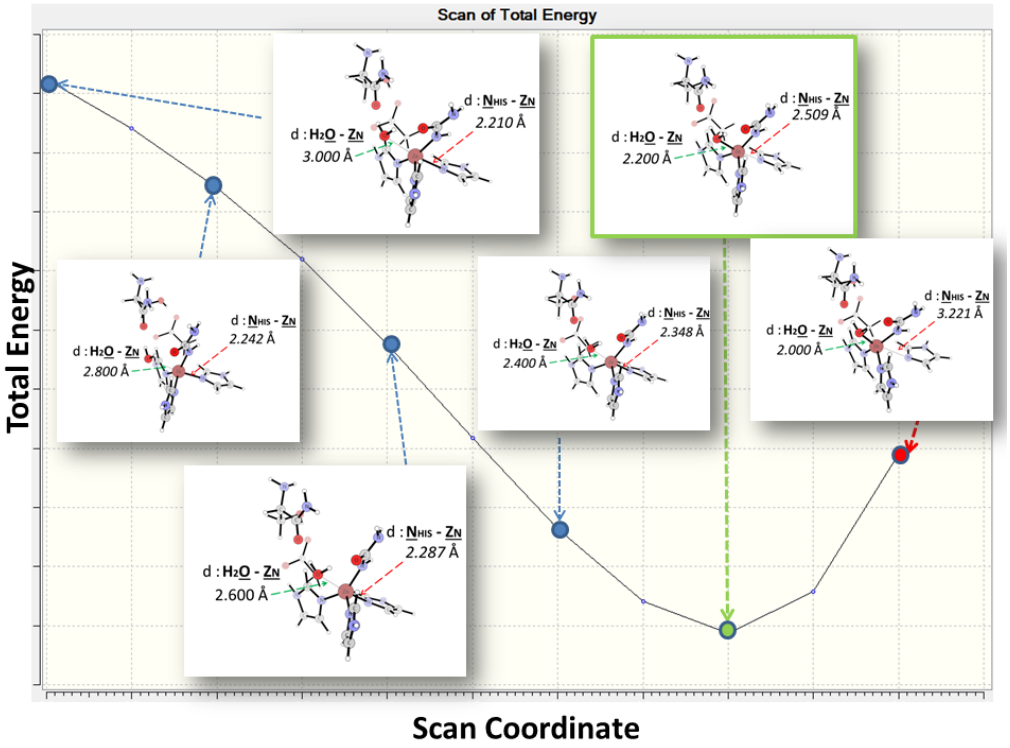


Figure S4 Relaxed potential energy scan on the Zn-Ow bond and relative geometries of the last six points of the 10 analyzed as far as the carbodiimide is concerned. Distance are in Å.

Paper II

Insight into the promiscuous activity of human carbonic anhydrase against the cyanic acid substrate from a combined QM and QM/MM investigation

Paolo Piazzetta, Tiziana Marino, and Nino Russo

Phys. Chem.Chem.Phys. **2014**, *16*, 16671-16676
DOI: 10.1039/C4CP02363C

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2014, **16**, 16671

Insight into the promiscuous activity of human carbonic anhydrase against the cyanic acid substrate from a combined QM and QM/MM investigation†

P. Piazzetta, T. Marino* and N. Russo

Received 29th May 2014,
Accepted 24th June 2014

DOI: 10.1039/c4cp02363c

www.rsc.org/pccp

The promiscuous activity of human carbonic anhydrase (hCAII) against a non-physiological cyanic acid substrate has been investigated by using a combined QM and QM/MM level of theory. Results show that the hCAII is able to hydrate the cyanic acid by a reaction mechanism similar to that of the CO_2 native substrate. The energy barrier for the nucleophilic attack is found to be 15.6 and 4.3 kcal mol^{-1} at QM and QM/MM levels, respectively. This result underlines the importance of taking into account the surrounding residues around the active site in the presence of the substrate having small molecular sizes. The carbamate is strongly stabilized with respect to the bicarbonate of the native substrate indicating a more difficult release of the reaction product.

A Introduction

Enzymes are essential for life in both vegetal and animal kingdoms. After three billion years of evolution these extraordinary chemical factories are able to catalyze a myriad of reactions involved in cell metabolism with high selectivity and efficiency. Starting from the pioneering work of Emil Fischer in 1894, many enzyme properties and structures have been accumulated in the framework of a reductionist approach to enzymology.[1] The discovery of the enzyme promiscuity evidence as the pluralism of causes and effects in enzymology requires a more holistic vision through quantitative measures, rigorous knowledge of the mechanism at the atomistic level and by data integration. Many recent studies underlined that enzyme promiscuity occurs in a consistent number of enzymes demonstrating that the paradigm for which one enzyme catalyzes one specific reaction is not entirely sufficient to explain the complex chemistry of these particular proteins.[2–7] For a long time, the enzymes' function was essentially concentrated on their ability to stabilize low energy transition states that allow us to explain the catalytic power of natural enzymes. Very recently, other important effects such as preorganization, energetic destabilization of reaction intermediates, dynamics and the role of product complex in the catalytic turnover have emerged as key factors to explain some behaviours of the complex enzyme chemistry.[8]

Due to its unique ability to reduce the thermodynamically stable carbon dioxide, human carbonic anhydrase II (hCAII) is of continuous interest and a huge number of experimental and theoretical studies have been devoted to its chemistry and biology.[9–18] In spite of this impressive amount of data, hCAII continues to attract the researchers' attention since it is able to show substrate and catalytic promiscuity and to substitute the native Zn ion with other metal ions maintaining its catalytic power.[19–23] These properties are still far from their complete knowledge and more work is necessary to obtain explanation and rationalization at the atomistic level. In this work, we present the result of a careful computational work performed by using quantum mechanical (QM) and hybrid quantum mechanical/molecular mechanics (QM/MM) tools, on the promiscuous ability of hCAII to also catalyze the hydration of cyanic acid. Previous experimental studies evidenced that hCAII catalyzes the hydration of cyanic acid to carbamic acid but the physiological relevance of this reaction is still unknown.[23–26] Experimental information based on NMR measurements shows that cyanic acid binds to the active sites of human and bovine carbonic anhydrase[27,28] and X-ray structural data evidence that this substrate coordinates to the active site of the enzyme in the same location and orientation as the carbon dioxide native substrate.[24,27,29,30] Molecular dynamics and gas-phase *ab initio* computations on a very small cluster model considering the Zn ion coordinated to three histidine residues have been previously performed in order to obtain information only on the

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Cubo 14/C, Piano 7°, Arcavacata di Rende (CS), Italy. E-mail: tmarino@unical.it; Fax: +39-0984-492085; Tel: +39-0984-493390

† Electronic supplementary information (ESI) available: Information on the MD procedure, force field parameters and their derivation, and figures containing optimized geometries at the QM level for both considered approaches. See DOI: 10.1039/c4cp02363c

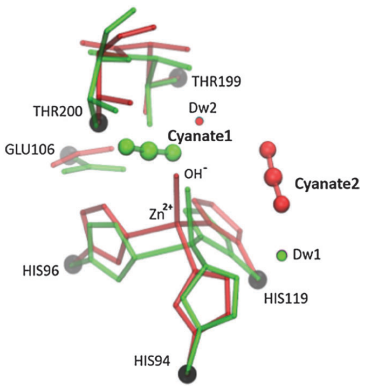
enzyme substrate complex.[28] To our knowledge no computational studies on the entire reaction mechanism are available in the literature.

A possible reaction scheme, similar to that for the native substrate, has been proposed[29] but the energetics and the atomistic details on the reaction mechanism are still unknown. A detailed knowledge of the potential energy surface can give new insights on the reaction path and can also contribute to elucidate why this substrate acts as a competitive inhibitor of hCAII and, at the same time, can undergo catalysis.

B Computational details

Our starting geometries are based on the X-ray structures of the free hCAII (PDB code 2CBA).[31] With the aim of obtaining a substrate interacting with enzyme in a fashion similar to that deduced by the experimental suggestions,[27,30] the QM cluster (see Scheme 1) has been built up performing docking simulation by using the Autodoc code.[32] Two different substrate approaches to the enzyme active site, namely carbodiimide- (approach **1**) and CO_2-like (approach **2**), have been considered (see Scheme 1).

The QM cluster includes, other than the substrate, the His94, His96, His119 fragments and an OH group coordinated to the metal center as the nucleophilic species. The second coordination shell is formed by Glu106, Thr199, Thr200 residues and one water molecule. The total number of atoms in the obtained cluster is 83. Furthermore, the choice of retaining the Thr199 and Glu106 H-bond network in the active site is well supported by the evidence of the observed reduction of catalytic activity when mutants at 199 and 106 positions have been taken into account.[33] QM/MM computations have been performed considering the ONIOM procedure[34–38] where the QM region, previously described, has been completed retaining the whole chain for every retained amino acid. The surrounding MM region includes the remaining protein. The UFF force field[39] previously employed in a QM/MM study on the natural substrate[40] for which an available kinetic study allowed us to verify the reliability of the theoretical results on the rate determining energetic barrier, has been used. Furthermore, in the MM region all the 220 water molecules present in the deposited crystallographic structure have been preserved.[31] So, our QM/MM system includes 115 atoms in the QM region and 4582 in the MM one. QM atoms have been represented by the 6-31+G(d,p) basis set concerning the C, N, O and H atoms while the SDD pseudopotential for the Zn ion[41] during the geometry optimization, whereas for the potential energy surfaces (PESs) the more extended 6-311+G(2d,2p) basis set has been employed. In the QM computations, as usual in the cluster model computations,[40,42–49] the considered residues have been truncated at the Cα atom and, in order to avoid an unrealistic cluster expansion, the terminal carbon atom for all the amino acids has been fixed in their X-ray crystal position (see Scheme 1). These constraints have been removed in the QM/MM computations. Transition states and intermediates have been subjected to the vibrational analysis in order to check the nature of their stationary states on the considered PES and to obtain the zero point corrections. In the QM computations the environment effects have been taken into account by means of single point calculations on the optimized geometries by using the polarizable continuum model[50,51] assuming a dielectric constant of 4 as suggested by different authors.[43,44,46–49] The optimizations were performed by employing the hybrid B3LYP exchange–correlation functional[52,53] but the PES's have been also obtained using the following other exchange correlation functionals: wB97XD,[54,55] M062x[56] and MPWB1K.[57] All the computations have been carried out by using Gaussian 03 code.[58] NBO analysis[59] was performed on every stationary point of the PESs.

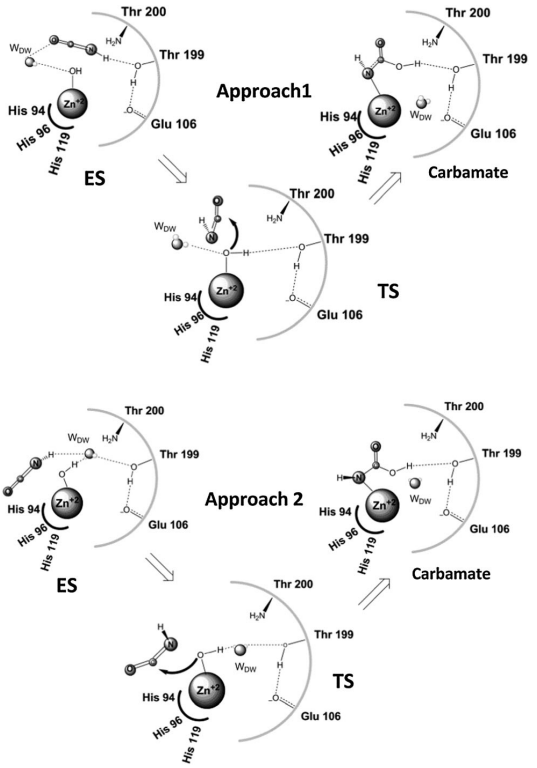


Scheme 1 Superposition of the amino acid residues treated at the QM level for the carbodiimide- (**1**) and CO_2-like (**2**) approaches.

C Results and discussion

The followed reaction mechanism is depicted in Scheme 2. By using this guideline we have been able to characterize all the stationary points along the considered reaction coordinate.

As a first step of the work we performed QM computations on the selected cluster, considering both the substrate approaches to the enzyme active site described above, by using different exchange–correlation functionals. In particular, we also used other than the B3LYP, previously employed in the computational studies of different kinds of enzymatic reactions,[42–45] the below-mentioned exchange–correlation functionals. In particular, the M062x and wB97XD take into account, in different ways, the contribution of the weak interactions that can often play an important role in the energetics of biological reaction processes. The results are collected in Table 1. From this table it is evident that the result for the barrier related to nucleophilic attack does not differ considerably. In contrast, the computed overall hydration energies noticeably differ even if all the employed exchange–correlation functionals give exothermic values. In particular, MPWB1K stabilizes the reaction product (-14.5 kcal mol^{-1}) while wB97XD gives the highest value (-4.3 kcal mol^{-1}). B3LYP value (-7.9 kcal mol^{-1}) lies above that MPWB1K and M062X by 6.6 and



Scheme 2 Proposed mechanisms for carbodiimide-like (**1**) and CO_2-like (**2**) approaches.

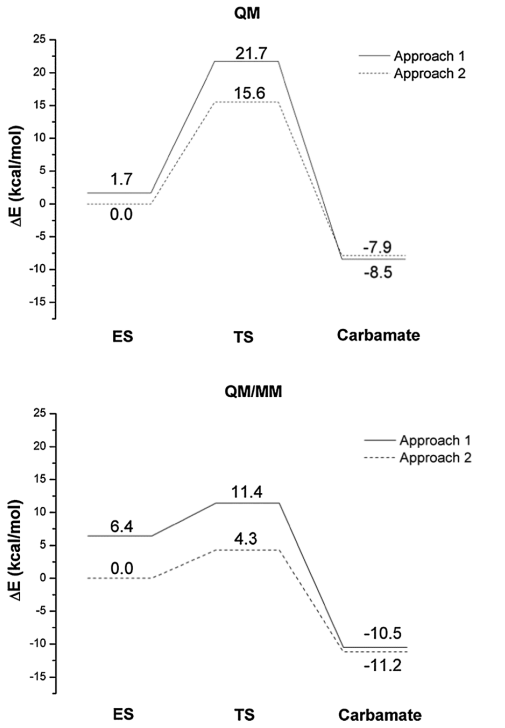


Fig. 1 QM and QM/MM potential energy surfaces for the hydration of cyanic acid by hCAII following carbodiimide-like (1) and CO_2-like (2) approaches. QM results are referred to the B3LYP/6-311+G(2d,2p) zero point corrected energies including solvent effects (PCM, ε = 4).

Table 1 Zero point corrected relative energy values for all stationary points of the PES following the carbodiimide-like (**1**) and CO_2-like (**2**) approaches obtained with different functionals used in QM calculations

Species	B3LYP		wB97XD		MPWB1K		M062X	
	1	**2**	**1**	**2**	**1**	**2**	**1**	**2**
ES	1.7	0.0	4.5	0.0	1.5	0.0	3.1	0.0
TS	21.7	15.5	23.4	18.2	20.3	15.0	21.4	15.3
Carbamate	−8.5	−7.9	−5.9	−4.3	−17.8	−14.5	−15.8	−12.4

4.5 kcal mol^{-1}, respectively. On the basis of these results and in spite of the possibility of performing comparisons with previous published reports on other hCAII substrates,[40] we have chosen to employ the B3LYP functional for our QM/MM investigation.

In Fig. 1 are depicted the QM and QM/MM PESs for the two considered approaches (**1** and **2**), while in Fig. 2 are shown the optimized structures of the species present on the PES with main geometrical distances obtained at QM/MM level for both considered approaches. The Cartesian coordinates of all the QM optimized species are available in the ESI.† As far as the QM cluster calculations are concerned, in the Michaelis–Menten complex, ES, the cyanic acid binds in the same fashion of the natural CO_2 substrate (approach **2**) and is only 1.7 kcal mol^{-1} more stable than the corresponding complex obtained through approach **1** (see Fig. 1). The TS relative to the nucleophilic attack lies at 21.7 and 15.6 kcal mol^{-1} for the approaches **1** and **2** respectively, whereas the products lie at 8.5 (approach **1**) and 7.9 (approach **2**) kcal mol^{-1} below the ES complex.

The obtained QM/MM PES (see Fig. 1) shows remarkable differences with respect to the QM one. The ES obtained for the approach **1** lies at 6.4 kcal mol^{-1} above that obtained with the CO_2-like orientation. The barrier for the carbodiimide-like attack is now 11.4 kcal mol^{-1} while that for the CO_2-like approach is only 4.3 kcal mol^{-1}. Furthermore, in this case the carbamate product obtained from the approach **2** is now slightly stable (11.2 *versus* 10.5 kcal mol^{-1}). Our findings for the ES and product agree well with previous HF and MP2 results performed only on these two stationary points and by using a minimal cluster.[28] Looking at the various geometries (Fig. 2) the following considerations can be drawn. The ES in both the approaches shows a tetrahedral geometry around the metal site; the spectator water molecule plays different role: in the approach **1**, water establishes hydrogen bond with the OH nucleophile and Thr200 (1.642 Å and 2.024 Å, respectively) but not with the substrate, while in the CO_2-like orientation the water molecule stabilizes the substrate with a hydrogen bond, 1.644 Å (1.656 Å at the QM level).

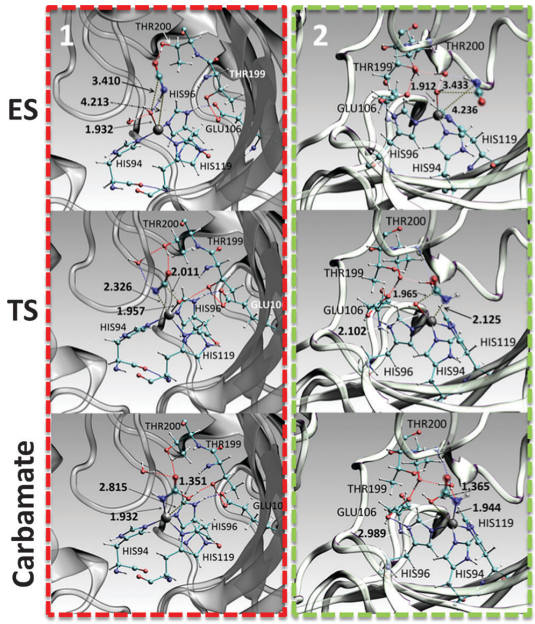


Fig. 2 B3LYP/6-31+G(d,p):UFF optimized geometries for the stationary points along the PES for carbodiimide-like (**1**) and CO_2-like (**2**) approaches.

This well accounts for the different stabilities of the two obtained enzyme–substrate complexes. Comparison of the two TS structures reveals that, also in this case, the presence of a hydrogen bond between the nitrogen atom of the substrate and the hydrogen of the water molecule (2.313 Å at QM/MM level), contributes to the stabilization of the TS structure only in the CO_2-like PES. The product structures have different orientations and, again, in approach **2** water molecules contribute to stabilize the product with a hydrogen bond with the NH group of the formed carbamate (2.116 Å).

The analyses of the geometries of the stationary points can help to elucidate the difference in the obtained PESs at the QM and QM/MM levels of theory. We will discuss only the CO_2-like approach but the same consideration can be drawn for the other pathway. Owing to the lack of the environment residues around the active site, the substrate lies farther from the nucleophilic OH group in QM approach than in QM/MM one. In fact, the distance involving the carbon atom of the cyanic acid and the oxygen of the OH group is 3.605 and 3.433 Å at QM and QM/MM levels, respectively.

TS structure, at the QM/MM level, is also stabilized by the H-bonds network created by the water that connects the substrate with the Thr199 and Thr200 residues. Going to discuss the structural features of the product topology, we underline that the most evident difference between the two structures concerns the position of the N–H bond of the carbamic acid that in the case of QM/MM is oriented in order to optimize the formation of a hydrogen bond with the oxygen atom of Thr200 residue that, in turn, is linked to both Glu106 and water with two hydrogen bonds.

Because of the consistent difference in the energy barrier obtained using the QM and QM/MM approaches, we have considered other QM clusters characterized by different sizes. The first one composed of 59 atoms and the same residues differently truncated gave a barrier of 15.6 kcal mol^{-1}, while the larger one with 102 atoms and the considered residues without truncation generated 13.7 kcal mol^{-1}. From these results it is evident that the cluster size does not play a significant role. On the other hand this means that the protein environment plays an important role. A comparison with QM and QM/MM PESs previously obtained for native CO_2 and promiscuous carbodiimide substrates[40] also shows that in these cases the barrier obtained at the QM level is about twice greater than the corresponding QM/MM value. As for the barrier's height found for CO_2, carbodiimide and cyanic acid, we note that these three isoelectronic substrates differ in agreement with their different polarities as evidenced by NBO analysis.

The different behaviour of hCAII towards three substrates does not depend on the nucleophile species, represented by the OH^- portion of the active site, or on the NBO charge on Zn^{2+} ion but on the different electrophilic nature of the carbon in the three examined substrates, as can be evinced by the trend of NBO charge depicted in Fig. 3. In fact, in the native substrate the carbon atom assumes more positive values in the ES, TS1 and INT species than found in the promiscuous substrates, underlining its more favourable nature for the nucleophilic addition.

With the aim of restoring the catalytic cycle in the final ligand-exchange step carbamate must be displaced with water. This step is required for the formation of the Zn–OH moiety that, consequently, hydrates the next substrate molecule. Owing to the importance of this process we have investigated at the QM level the trend of the energy when the water molecule approaches the metal center coordinated with the carbamate product. The resulting scan of the energy as a function of the approaching distance ($Zn–OH_2$) is depicted in Fig. 4. As this distance decreases, the energy increases and at a distance of 2.0 Å reaches

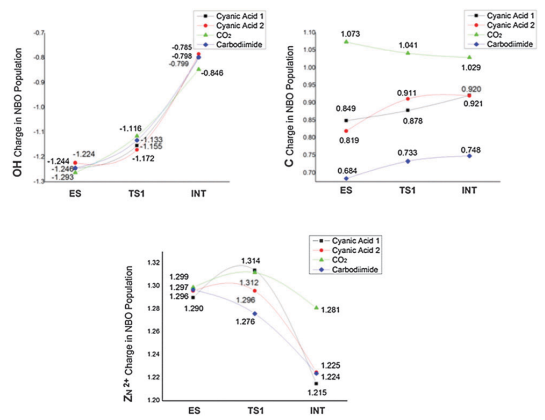


Fig. 3 NBO charges of the atoms involved in the reaction for the CO_2, carbodiimide and cyanic acid considered substrates.

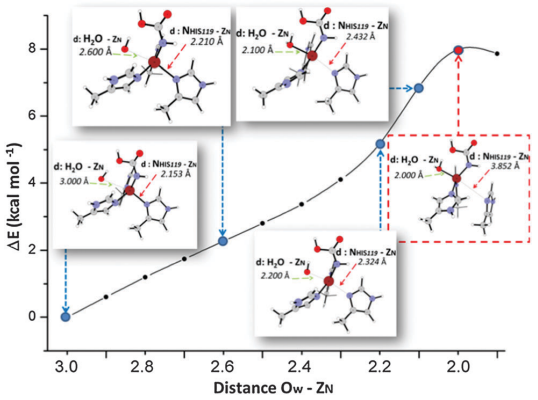


Fig. 4 Potential energy scan for the approach of H_2O molecule to the carbamate–hCAII complex as a function of H_2O–Zn distance.

its maximum value. The inspection of the geometrical parameters during the energy scan shows that the carbamate does not leave the metal but the His119 residue, linked to the Zn ion, starts to leave the metal coordination sphere; as a consequence the penta-coordinated complex is not formed. In fact, at a Zn–OH_2 distance of 2.0 Å, this histidine is practically detached from the metal center (Zn–N_{His119} = 3.852 Å), similar behaviour has been recently obtained for the carbodiimide product.[40] In the case of the native substrate, at the QM level the same energy scan shows that the bicarbonate replacement by water requires 5.6 kcal mol^{-1} giving rise to a penta-coordinated complex.[40] This behaviour was found similar to that previously studied by Bottoni on a CO_2 substrate at the QM level.[10] This information coupled with a consistent stabilization of the carbamate product (−8.2 kcal mol^{-1} at the QM level) indicate that the restoring of the catalytic cycle for the two substrates has different behaviours.

In order to have another insight on this important step, molecular dynamics (MD) investigations, whose computational procedures are described in the ESI,† have been performed on the bicarbonate- and carbamate–enzyme complexes. From the analysis of the conformational behaviour during the simulation time (10 ns), it emerges that MD simulation concerning the native product yields a water molecule approaching the metal center of the active site unlike what occurs in the carbamate product. In Fig. 5 is reported the trajectory followed by the water molecule for coordinating to the zinc ion in the case of the bicarbonate–hCAII complex. This fact corroborates the existence of the penta-coordinated complex during the restoration of the catalytic cycle, also observed by studies carried out at higher levels of theory.[10,40]

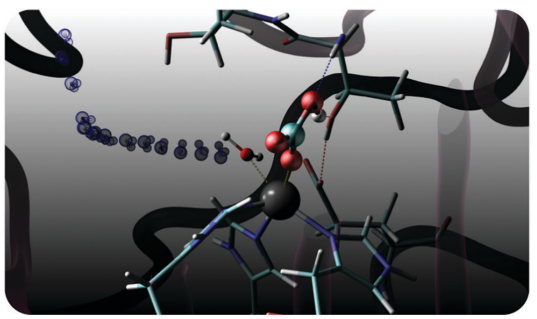


Fig. 5 Snapshot from MD simulation showing the trajectory of the water molecule to zinc ion in the case of bicarbonate–hCAII complex.

D Conclusions

In this paper we have presented a combined QM and QM/MM study on the hydration of a non-physiological substrate (cyanic acid) of the hCAII. Results show the promiscuous activity of human carbonic anhydrase against this substrate and reveal the following new features:

– hCAII is able to hydrate the cyanic acid substrate although with a slower kinetics with respect to that with the native CO_2 substrate;

– The cyanic acid prefers to approach the catalytic center in a fashion similar to that of the CO_2 molecule;

– The barrier for the nucleophilic attack is higher than in the CO_2 natural substrate;

– The carbamate formation is an exothermic process;

– The release of the product at the expense of water for restoring the catalytic cycle seems to be the critical step with respect to the catalysed natural substrate reaction.

Acknowledgements

This work was financially supported by Department of Chemistry and Chemical Technologies of the Università della Calabria (Italy). P. Piazzetta gratefully acknowledges Commissione Europea, Fondo Sociale Europeo, and Regione Calabria for the financial support.

Notes and references

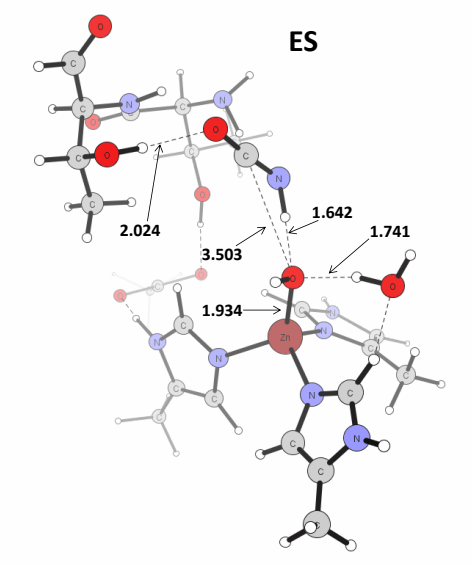
1 E. Fischer, *Ber. Dtsch. Chem. Ges.*, 1894, **27**, 2985–2993.
2 P. J. O'Brien and D. Herschlag, *Chem. Biol.*, 1999, **6**, R91–R105.
3 T. M. Penning and J. M. Jez, *Chem. Rev.*, 2001, **101**, 3027–3046.
4 S. D. Copley, *Curr. Opin. Chem. Biol.*, 2003, **7**, 265–272.
5 L. C. James and D. S. Tawfik, *Trends Biochem. Sci.*, 2003, **28**, 361–368.
6 O. Khersonsky, C. Roodveldt and D. S. Tawfik, *Curr. Opin. Chem. Biol.*, 2006, **10**, 498–508.
7 I. Catrina, P. J. O'Brien, J. Purcell, I. Nikolic-Hughes, J. G. Zalatan, A. C. Hengge and D. Herschlag, *J. Am. Chem. Soc.*, 2007, **129**(17), 5760–5765.
8 W. J. Albery and J. R. Knowles, *Angew. Chem., Int. Ed. Engl.*, 1977, **16**, 285–293.
9 M. Mauksch, M. Bräuer, J. Weston and E. Anders, *ChemBioChem*, 2001, **2**, 190–198.
10 G. P. Miscione, M. Stenta, D. Spinelli, E. Anders and A. Bottoni, *Theor. Chem. Acc.*, 2007, **118**, 193–201.

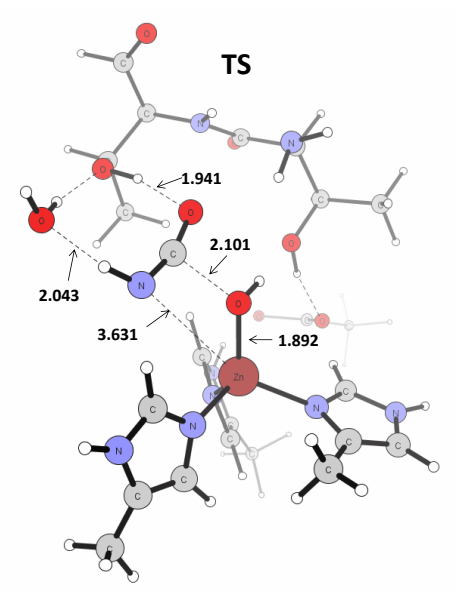
11 O. Amata, T. Marino, N. Russo and M. Toscano, *Phys. Chem. Chem. Phys.*, 2011, **13**, 3468–3477.
12 T. Marino, N. Russo and M. Toscano, *J. Am. Chem. Soc.*, 2005, **127**, 4242–4253.
13 I. Bertini and C. Luchinat, *Acc. Chem. Res.*, 1983, **16**, 272–279.
14 D. N. Silverman and S. Lindskog, *Acc. Chem. Res.*, 1988, **21**, 30–36.
15 K. Hakansson, M. Carlsson, L. A. Svensson and A. Liljas, *J. Mol. Biol.*, 1992, **227**, 1192–1204.
16 D. W. Christianson and C. A. Fierke, *Acc. Chem. Res.*, 1996, **29**, 331–339.
17 V. M. Krishnamurthy, G. K. Kaufman, A. R. Urbach, I. Gitlin, K. L. Gudisen, B. D. Weibel and G. M. Whitesides, *Chem. Rev.*, 2008, **108**, 946–1051.
18 D. R. Garmer, *J. Phys. Chem. B*, 1997, **101**(15), 2945–2953.
19 E. T. Kaiser and K. W. Lo, *J. Am. Chem. Soc.*, 1969, **91**, 4912–4918.
20 Y. Pocker and D. R. Storm, *Biochemistry*, 1968, **7**, 1202–1214.
21 Y. Pocker and S. Sarkanen, *Biochemistry*, 1978, **17**, 1110–1118.
22 C. Tu, H. G. Thomas, G. C. Wynns and D. N. Silverman, *J. Biol. Chem.*, 1986, **261**, 10100–10103.
23 M. Lopez, H. Vu, C. K. Wang, M. G. Wolf, G. Groenhof, A. Innocenti, C. T. Supuran and S.-A. Poulsen, *J. Am. Chem. Soc.*, 2011, **133**, 18452–18462.
24 C. T. Supuran, in *Carbonic Anhydrase Its Inhibitors and Activators*, ed. C. T. Supuran, A. Scozzafava and J. Conway, CRC Press LLC, Boca Raton, FL, 2004, pp. 1–23.
25 F. Briganti, S. Mangani, A. Scozzafava, G. Vernaglione and C. T. Supuran, *JBIC, J. Biol. Inorg. Chem.*, 1999, **4**(5), 528–536.
26 Y. Pocker and J. T. Stone, *Biochemistry*, 1968, **7**, 3021–3031.
27 I. Bertini, C. Luchinat, R. Pierattelli and A. J. Vila, *Inorg. Chem.*, 1992, **31**, 3975–3979 and references therein.
28 Z. Peng, K. M. Merz and L. Banci, *Proteins: Struct., Funct., Genet.*, 1993, **17**, 203–216.
29 C. T. Supuran, C. W. Conroy and T. H. Maren, *Proteins: Struct., Funct., Genet.*, 1997, **27**, 272–278.
30 M. Lindahl, L. A. Svensson and A. Liljas, *Proteins: Struct., Funct., Genet.*, 1993, **15**, 177–182.
31 K. Hakansson, M. Carlsson, L. A. Svensson and A. Liljas, *J. Mol. Biol.*, 1992, **227**, 1192–1204.
32 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.
33 J. F. Krebs, F. Rana, R. A. Dluhy and C. A. Fierke, *Biochemistry*, 1993, **32**, 4496–4505.
34 A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227–249.
35 Z. Ke, S. Abe, T. Ueno and K. Morokuma, *J. Am. Chem. Soc.*, 2012, **134**, 15418–15429.
36 S. Martí, J. Andrés, E. Silla, V. Moliner, I. Tuñón and J. Bertrán, *Angew. Chem., Int. Ed.*, 2007, **46**, 286–290.
37 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198–1229.
38 M. W. van der Kamp and A. J. Mulholland, *Biochemistry*, 2013, **52**, 2708–2728.
39 K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard III and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.
40 P. Piazzetta, T. Marino and N. Russo, *Inorg. Chem.*, 2014, **53**, 3488–3493.
41 D. Andrae, U. Haussermann, M. Dolg, H. Stoll and H. Preuss, *Theor. Chim. Acta*, 1990, **77**, 123–141.
42 O. Amata, T. Marino, N. Russo and M. Toscano, *J. Am. Chem. Soc.*, 2011, **133**, 17824–17831.
43 R.-Z. Liao, J.-G. Yu and F. Himo, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**(52), 22523–22527.
44 M. J. Ramos and P. A. Fernandes, *Acc. Chem. Res.*, 2008, **41**(6), 689–698.
45 M. E. Alberto, T. Marino, M. J. Ramos and N. Russo, *J. Chem. Theory Comput.*, 2010, **6**(8), 2424–2433.
46 L. Noodleman, T. Lovell, W.-G. Han, J. Li and F. Himo, *Chem. Rev.*, 2004, **104**, 459–508.
47 L. Noodleman and W.-G. Han, *JBIC, J. Biol. Inorg. Chem.*, 2006, **11**, 674–694.
48 R. A. Torres, T. Lovell, L. Noodleman and D. A. Case, *J. Am. Chem. Soc.*, 2003, **125**, 1923–1936.
49 P. E. M. Siegbahn and M. R. A. Blomberg, *Chem. Rev.*, 2000, **100**, 421–437.
50 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
51 M. Cossi, N. Rega and G. Scalmani, *J. Comput. Chem.*, 2003, **24**, 669–681.
52 A. D. J. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
53 C. T. Lee, W. T. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.
54 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.
55 J. D. Chai and M. Head-Gordon, *J. Chem. Phys.*, 2008, **128**, 84106–84155.
56 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
57 (*a*) A. D. Becke, *J. Chem. Phys.*, 1996, **104**, 1040–1046; (*b*) C. Adamo and V. Barone, *J. Chem. Phys.*, 1998, **108**, 664–675; (*c*) Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 6908–6918.
58 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian 03, Revision A.1*, Gaussian, Inc., Pittsburgh PA, 2003.
59 E. D. Glendening, A. E. Reed, J. E. Carpenter and F. Weinhold, NBO, version 3.1.

Insight on the promiscuous activity of human carbonic anhydrase against the cyanic acid substrate from a combined QM and QM/MM investigation

P. Piazzetta,[a] T. Marino[a] and N. Russo[a]

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, P Bucci,I-87036 Arcavacata di Rende, Italy. E-mail: tmarino@unical.it.





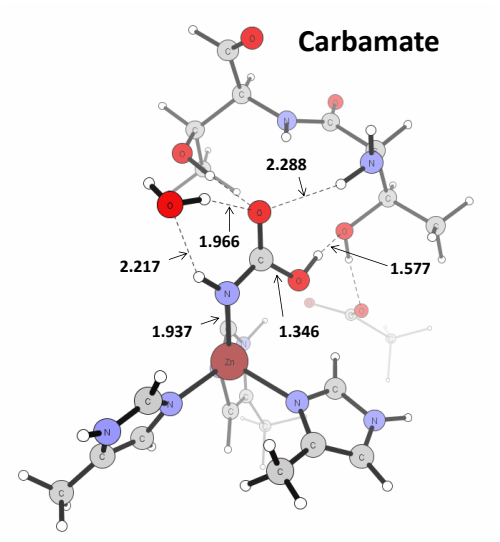


Figure S1. QM optimized structures of the stationary points along the PES of cyanic acid for approach **1**. All bond lengths are given in Å

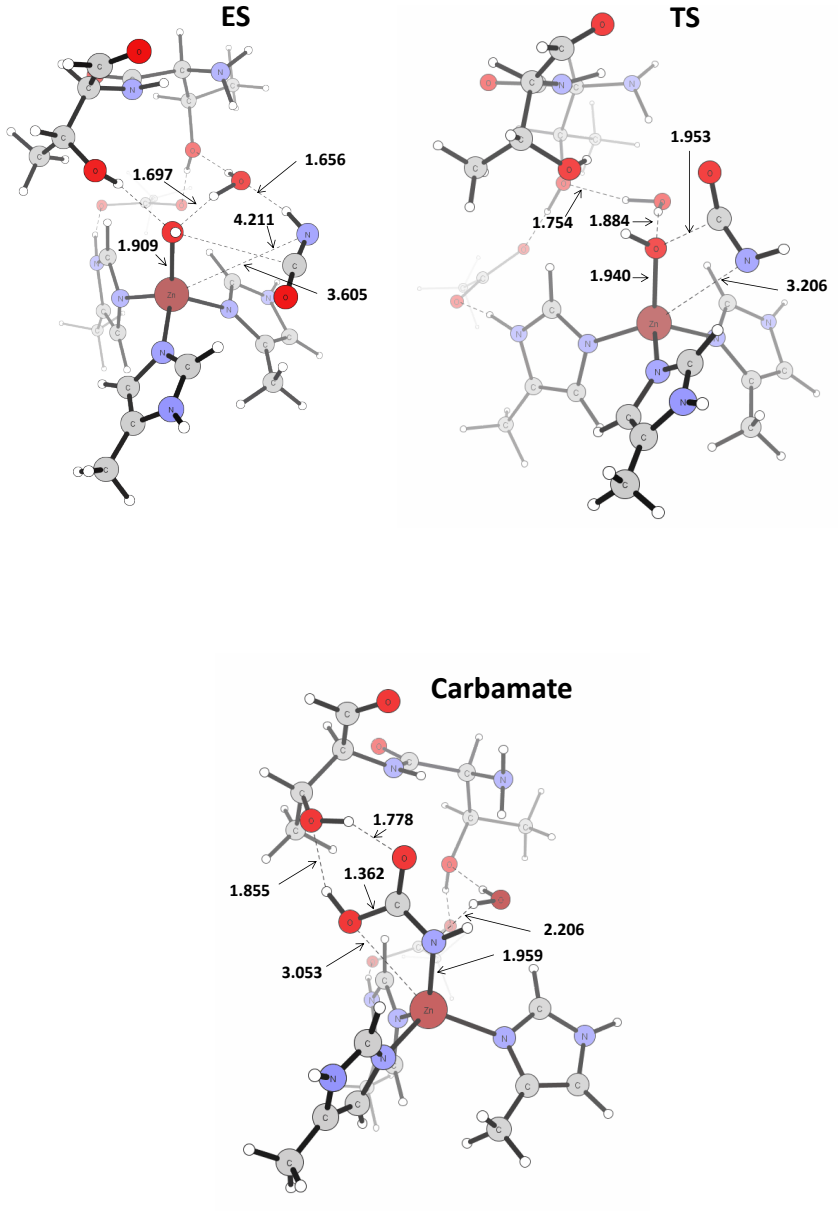


Figure S2. QM optimized structures of the stationary points along the PES of cyanic acid for approach **2**. All bond lengths are given in Å

MD Procedure

MD simulations were carried out with NAMD[1] using CHARMM[2] force field. The crystal structure of HCO_3^--hCAII complex was available (PDB Code 2VVB) and was used as starting structure, while due to the lacking of crystal structure of the analogue carbamate-hCAII, the minimized energy structure of carbamate was docked into the hCAII (PDB code 2CBA) using Autodock[3]. All ligands required custom parameters for the force field. They were obtained from parameters by similar models through analogies (carbon dioxide) and parameterized in order to create the topology and input files required by the CHARMM calculations starting from the Gaussian output files (carbamate). The charges for the ligands atoms, for the metal core and for the coordinated histidines were set using the NBO charges from DFT calculations.
A combination of bonded and non-bonded description was used in order to reproduce the natural flexibility of the enzymatic metal core and to preserve the zinc coordination: harmonic constraints were introduced using the Colvars module of NAMD with QM/MM values regarding bonds, angles and dihedral.[see Table S1] [4]

A water layer of 16 Å was built around each studied enzyme-product complex creating a water box of 81 x 77 x 87 $Å^3$ for hCAII-carbamate and 80 x 78 x 89 $Å^3$ for hCAII-HCO_3^-. The resulting number of atoms corresponded to 51053 for hCAII-carbamate complex and 52101 for the hCAII-HCO_3^- one. Due to the neutral charge of the studied systems no further neutralization was needed. SHAKE [5] algorithm was practiced in order to constrain all bonds involving hydrogen atoms. Periodic boundary conditions were used. The long-range electrostatic interactions were evaluated by Particle Mesh Ewald (PME) method [6]. The supramolecular systems obtained including protein, waters and the ligand (HCO_3^- and carbamate) were minimized using the conjugate gradient algorithm: in first instance, minimizing the water molecules keeping the protein and ligand fixed, and then minimizing the entire system.

A time step of $\Delta t = 2$ fs was used for both systems. A cutoff of 12 Å for non-bonded interactions was applied and a switching scheme was used. Scaled 1–4 parameters were enabled for 1–4 interactions. A first equilibration of 60 ps was performed on the minimized models by increasing temperature until to 300 K in increments of 20 K. Then a simulation time equal to 10 ns was performed in the NPT ensemble with a Langevin dampening coefficient of 5 ps^{-1} controlling pressure with the Langevin piston pressure control.

Table S1 : Parameters for the zinc coordination in the bicarbonate–hCAII complex.

Atoms			Force Constant	Equilibrium
Bonds			k	r_e
Zn		N_{119}	30 kcal/mol/ Å^2	2.07 Å
Zn		N_{94}	30 kcal/mol/ Å^2	2.01 Å
Zn		N_{96}	30 kcal/mol/ Å^2	2.01 Å
Angles			k	θ_e
N_{119}	Zn	N_{96}	23 kcal/mol/ rad^2	96.9°
N_{96}	Zn	N_{94}	23 kcal/mol/ rad^2	107.2°
N_{94}	Zn	N_{119}	23 kcal/mol/ rad^2	118.2°
$O_{bicarbonate}$	Zn	N_{119}	0.05 kcal/mol/ rad^2	104°
$O_{bicarbonate}$	Zn	N_{94}	0.05 kcal/mol/ rad^2	108°
$O_{bicarbonate}$	Zn	N_{96}	0.05 kcal/mol/ rad^2	120°
Improper Dihedrals			k	ϕ_0
Zn	N_{119}	N_{96} N_{94}	25 kcal/mol/ rad^2	0

Table S2 : Parameters for the zinc coordination in the carbamate –hCAII complex.

Atoms			Force Constant	Equilibrium
Bonds			k	r_e
Zn		N_{119}	30 kcal/mol/ Å^2	2.13 Å
Zn		N_{94}	30 kcal/mol/ Å^2	2.01 Å
Zn		N_{96}	30 kcal/mol/ Å^2	2.01 Å
Angles			k	θ_e
N_{119}	Zn	N_{96}	23 kcal/mol/ rad^2	94.5°
N_{96}	Zn	N_{94}	23 kcal/mol/ rad^2	103.4°
N_{94}	Zn	N_{119}	23 kcal/mol/ rad^2	110.1°
$O_{carbamate}$	Zn	N_{119}	0.05 kcal/mol/ rad^2	109.7°
$O_{carbamate}$	Zn	N_{94}	0.05 kcal/mol/ rad^2	123.7°
$O_{carbamate}$	Zn	N_{96}	0.05 kcal/mol/ rad^2	113.6°
Improper Dihedrals			k	ϕ_0
Zn	N_{119}	N_{96} N_{94}	25 kcal/mol/ rad^2	0

References

[1] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten. J. Comp. Chem. 2005, 26, 1781.
[2] K. Vanommeslaeghe, et al J. Comp. Chem. 2010, 31, 671.
[3] O. Trott, A. J. Olson, J. Comput. Chem. 2010, 31, 455.
[4] M. Schmid, E. S. Nogueira, F. W. Monnard, T. R. Ward and M. Meuwly Chem. Sci., 2012, **3**, 690.
[5] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, J. Comput. Phys. 1977,**23**, 327.
[6] T., Darden, D. York, L. Pedersen. J. Chem. Phys. 1993, **98**, 10089.

Paper III

The working mechanism of the β-carbonic anhydrase degrading carbonyl sulfide (COSase): a theoretical study

Paolo Piazzetta, Tiziana Marino, and Nino Russo

Submitted manuscript

The working mechanism of the β-carbonic anhydrase degrading carbonyl sulfide (COSase): a theoretical study

Paolo Piazzetta,[a] Tiziana Marino[a]* and Nino Russo[a,b]

Abstract: In order to give insights on the working mechanism of the novel characterized enzyme carbonyl sulfide hydrolase (COSase), that efficiently converts COS in H_2S and CO_2, we have performed a detailed theoretical investigation in the framework of density functional theory (B3LYP and M06 exchange-correlation functionals) by using the cluster model approach. The proposed mechanism only partially follows that accepted for human carbonic anhydrase. In the final part of the reaction the metal ion is unable to form a penta-coordinated species. The B3LYP-D3 and M06 potential energy surfaces show a very similar shape. The elucidation of the catalytic reduction of COS assumes also importance in view of the fact that carbonyl sulfide contributes to the global climate change.

Introduction

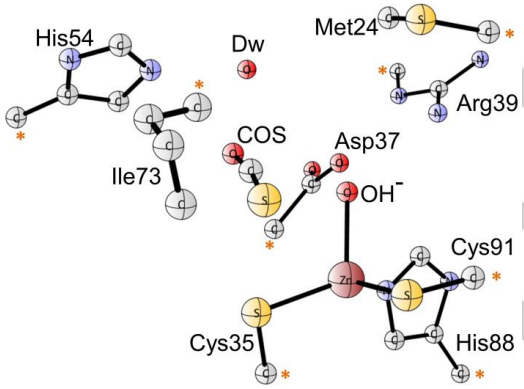
The consistent climate change caused by the release in the atmosphere of different pollutants produced by anthropogenic activities constitutes one of the major effort for the scientists today. Other than the reduction of the pollutant emissions, technologies aim at the capture and utilization also via transformation of the dangerous compounds. Carbonyl sulfide (COS) is released in the environment both by natural and anthropogenic sources and reaches in the troposphere the concentration of about 500 parts per trillion by volume[1] contributing to influence the Earth's radiation balance, to increase the greenhouse effect and to affect the ozone concentration.[2,3] In this scenario, the COS degradation via chemical or biological ways is a whishing effort to contribute at the decrease of sulfate aerosol formation. Nature provides a guideline in this direction since a series of enzymes are able to degrade the COS although this is not their natural substrate. [4] In particular, carbon disulfide hydrolase extracted from *Acidianus sp.* Strain A1-3 is also able to catalyze the hydrolysis of COS other than the CS_2 with an efficient catalytic turnover but the presence of this thermophilic archaeon is limited to solfataric springs.[5,6] Very recently, Katayama et al. have isolated, characterized and determined the structure of carbonyl sulfide hydrolase (COSase) from *Thiobacillus Thioparus* bacterium that efficiently convert COS in H_2S and CO_2.[7] This enzyme belongs to the carbonic anhydrase enzyme family and is homologous to the β-class Carbonic Anhydrases (CA).[8] The *T. thioparus* is commonly found in soil and freshwater and is used for the wastewater treatment as thiocyanate degrading microorganism.[9] COSase participates to the SCN^- degradation cycle that occurs in different steps: the first is initiated by the cobalt-containing enzyme thiocyanate hydrolase that converts the substrate in ammonia and COS,[10] while in the second step, COS is converted in H_2S and CO_2 by COSase and finally the hydrogen sulfite is oxidized in sulfate.[11] Although the knowledge of the COSase structure that gives insights on the active site, the catalytic process is not well characterized and only one hypothesis on the reaction mechanism is present in the literature.[7] In this work, starting from the proposed mechanism, we have performed a detailed quantum chemical investigation on the reduction of COS by COSase building up the entire potential energy surface (PES) for the reaction by using a cluster of 86 atoms representing the active site of the enzyme. This type of approach has been previously and successfully employed by different authors in order to highlight the reaction mechanism of a series of different classes of enzyme.[12-21] In particular, we mention as the cluster approach gives reliable results especially in metal containing enzymes if, other than the residues of the first coordination shell, also that indirectly implicated in the chemical process or in the stabilization of the transition state are considered. This approach has been used in a huge number of different types of enzymatic reactions and its validity has been evidenced in the recent reviews of Siegbahn and Himo[13, 22,23] and Ramos groups[15] on the basis of comparison with experimental data and QM/MM calculations.

Active site model. Although the X-ray structure of carbonyl sulfide hydrolase (COSase) from *T. thioparus* has been recently published,[7] we have employed as starting structure that coming from β-CA of *Myc. Tubercolosis,*[8] for the following reasons: a) in the deposited COSase structure (PDB=3VQJ some amino acid residues of the active site are missing since, as authors state, "the structures of COSase and the COSase/thiocyanate complex were solved by a molecular replacement method with the program Molrep using the partial structures of β-CA from *Myc. Tubercolosis* (PDB code 1YLK, residues 23-99)"; b) in their work T. Ogawa *et al.* report a superimposition of 146 Cα of two structures evidencing a high similarity with a RMSD value of 1.2. [7]

The $Zn(Cys)_2HisOH$ primary coordination shell of the carbonic anhydrase along with amino acid residues belonging to the outer one has been taken into account for building the cluster. The obtained cluster for mimicking the active site is more extended than that used in our previous work on the enzyme member of the ζ-class carbonic anhydrases.[16] In particular, amino acids involved in the network of H-bond present in the active site and essential to warranty the structural flexibility useful during the

[a] Dr. P. Piazzetta, Dr. T. Marino, Prof. N. Russo
Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Cubo 14C, Via P. Bucci 87036 Arcavacata di Rende (CS) Italy Phone: +39-0984-492085, fax: +39-0984-492044
E-mail: tmarino@unical.it
[b] Prof. N. Russo
Departamento de Química, División de Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana-Iztapalapa, Av. San Rafael Atlixco No. 186, Col. Vicentina C.P. 09340, México D. F.
Supporting information for this article is given via a link at the end of the document.

working mechanism of the enzyme have been retained: histidine54 (His54), isoleucine73 (Ile73), methionine24 (Met24). Amino acids surrounding the metal center are aspartate37 (Asp37) and arginine39 (Arg39). Ile73 residue comes into play in the approaching phase of the substrate because it marks the area surrounding the active center where the COS can be activated by the nucleophile group. The dyad Asp37-Arg39 forms the salt bridge observed in the active site of β-carbonic anhydrases arising from other microorganisms.[17,18] The obtained cluster consists of 86 atoms (about 1000 and 1500 basis functions for optimization and single point calculations, respectively) with a total charge of -1. We remark that due to the hydrophobic nature of catalytic cavity and to the long distance of other residues, our cluster contains all the amino acid residues significant for the reaction mechanism as well as for taking into account the stabilization of the species along the reaction path. As usual in the cluster modeling,[16,19-21, 24-26] the retained amino acids have been truncated at the Cα atom and, in order to avoid an unrealistic cluster expansion, the terminal carbon atom for all the amino acids has been fixed in their X-ray crystal position (see Scheme 1).



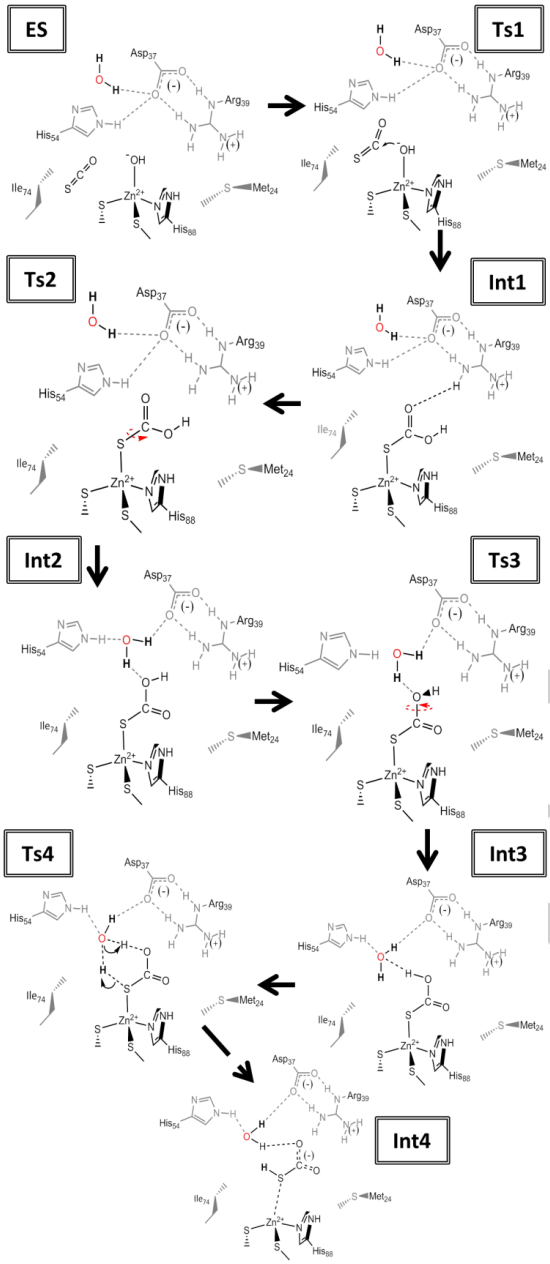
Scheme 1. Catalytic site model used in the calculations. The stars indicate the atoms kept frozen during the optimization procedure.

Results and Discussion

On the basis of the hypothesis available in literature[7] and of the consolidated mechanism proposed for the CO_2 substrate, [27-31] we have considered two possible mechanisms that essentially differ from the position of the deep water molecule interacting with the salt bridge Asp-Arg. One of these mechanisms (Scheme S1 in supporting information) shows a PES with higher energetic values while the other one (lower energy path) is depicted in Scheme 2 and the corresponding PES is reported in Figure 1. As shown in Scheme 2 and Figure 1, the production of H_2S and CO_2 requires different steps characterized by energetics lying under the formation of the Michaelis-Menten complex (ES) except for the first transition state (TS1). The approaching of the COS substrate gives rise to the enzyme-substrate complex (ES), depicted in Figure 2. The optimized structure of ES reveals a COS oriented in a similar fashion to thiocarbonate observed in crystallographic structure of (SCN^-) COSase complex[7] with the sulphur atom pointing toward the zinc ion with the distance (3.972 Å versus 3.928 Å in X-ray structure) that indicates to not be coordinated to the metal ion. The hydroxide group shows a distance from COS carbon of 3.313 Å, that means that the nucleophilic attack on the substrate is feasible. A thick network of H-bonds involving deep water molecule and the most hydrophilic portion of active site (His and Asp-Arg dyad) is established. The observed arrangement of the COS in the ES complex is different from that of CO_2 natural substrate in the previously investigated carbonic anhydrase cambialistic enzyme[16] where the zinc ion has a similar coordination environment. Different factors can influence the changed behavior and can be related to the chosen model for simulating the active site (in the present study the cluster contains also His, Met and Ile residues) as well as to the different nature of ion-dipole interaction.

The nucleophilic attack leads to the formation of the first intermediate (INT1) throughout the transition state (TS1), see Scheme 2. The optimized geometry of TS1 is shown in Figure 3 where is evident from the value of the OH-COS distance (2.403 Å) that the substrate is approaching to the hydroxyl group to undergo the nucleophilic attack. At the same time the S-Zn distance suffers an elongation (4.411 Å) with respect to that in the Michaelis-Menten complex, while the OH-Zn bond is a little bit longer (2.001 Å versus 1.994 Å in ES). A hydrogen-bonding network similar to that found in the enzyme-substrate complex is preserved and contributes to stabilize the TS1 structure. The vibrational analysis clearly indicated only one imaginary frequency (88.4*i* cm^{-1}) related to the stretching of the incoming OH-COS bond. The TS1 energy lies at 1.4 and 1.3 kcal/mol above the starting Michaelis-Menten complex (see Figure 1) at B3LYP-D3 and M06 level, respectively. The INT1 species is characterized by a hydroxide oxygen- Zn^{2+} distance of 3.459 Å and a sulphur–Zn^{2+} length is 2.537 Å indicating that the formed hydroxythiocarbonate acts as monodentate ligand toward the metal ion center. The deep water molecule is located close to the hydrophilic region of the catalytic pocket characterized by the Asp-Arg dyad and His. INT1's energy is calculated to be 24.5 (-26.8) kcal/mol lower than the ES complex at B3LYP-D3 (M06) level. From INT1 intermediate the reaction proceeds with a rotation of the hydroxythiocarbonate in a Lindskog-like fashion giving rise to a new intermediate (INT2) reached bypassing the TS2. The hydroxythiocarbonate in TS2 must suffer a rotation around the S-C bond to give the isomeric form with the OH group oriented in a more suitable way for the next steps of the reaction. The 59.79*i* cm^{-1} found value of the imaginary frequency accounts for the occurring rotation around the mentioned bond.

The torsion angle (Zn-S-C=O) measures -114.03° indicating that this moiety is not planar. The TS2 lies at 6.4 (B3LYP-D3) and 6.3 (M06) kcal/mol above the INT1.



Scheme 2. Proposed reaction mechanism for the conversion of COS in H_2S and CO_2

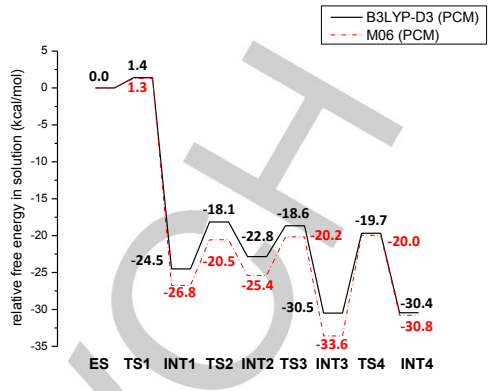


Figure 1. Potential energy surfaces obtained at B3LYP/6-311+G(2d,2p) (PCM, ε=4) and M06/6-311+G(2d,2p) (SMD, ε=4) levels.

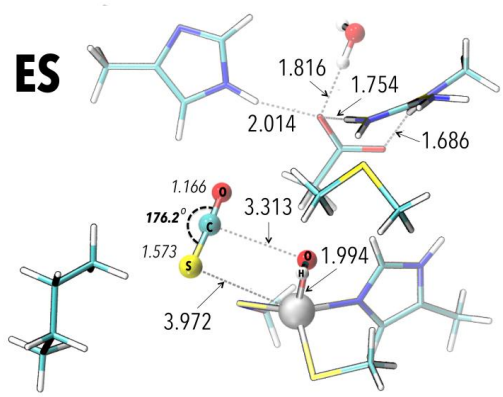


Figure 2. B3LYP/6-31+G(d) optimized geometry of the Michaelis-Menten complex (ES). All bond lengths are given in Å.

In INT2 is possible to observe that the OH group is pointed towards the hydrophilic region establishing a series of hydrogen bonds where the water molecule bridges this group (1.958 Å) to the His54 (1.777 Å) and Asp37 (1.767 Å) residues. The value of the torsion angle (Zn-S-C=O) is now 26.12° accounting for the occurred rotation. The Zn-S distance suffered a shortening (2.512 Å) compared to the previous minimum (2.537 Å) while the oxygen of the carbonyl group lies at 3.762 Å from the zinc ion. The next transition state (TS3) connects INT2 with the new intermediate, INT3, in which the hydroxyl group assumes a different orientation. The barrier for this process is found to be 4.2 and 5.2 kcal/mol at B3LYP-D3 and M06, respectively.

In TS3 the occurring rotation concerns the S-C-O-H torsion angle, that assumes the value of 98.38° and the corresponding found imaginary frequency is 498*i* cm^{-1}. This rotation generates a new arrangement of hydrogen bonds network, in which the water molecule is trapped in four H-bonds with His (1.797 Å),

Asp (1.695 Å) and with S (2.292 Å) and OH (1.785 Å) moieties of hydroxythiocarbonate. The presence of this more extended network makes INT3 more stable than INT2 by 7.7 (B3LYP-D3) and 8.2 (M06) kcal/mol.

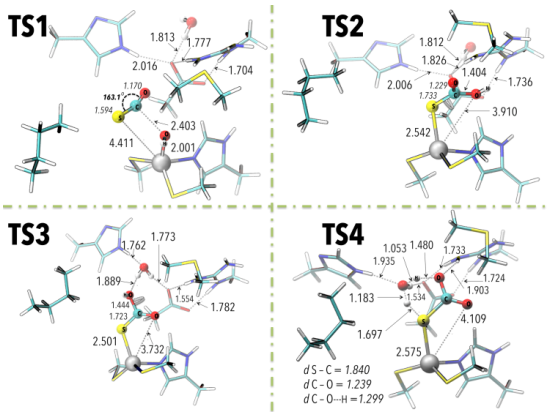


Figure 3. B3LYP/6-31+g(d,p) optimized geometries for the transition states along the explored PES. All bond lengths are given in Å.

The next step of the mechanism requires a proton transfer from the OH to sulphur atom of the hydroxythiocarbonate fragment. This H-shift is mediated by the deep water molecule that simultaneously acts as hydrogen acceptor and donor. The transition state located for this process lies at 10.8 (B3LYP-D3) and 13.6 (M06) kcal/mol above the INT3 (Figure 1). The imaginary frequency of 886.62*i* cm^{-1} obtained for this TS well accounts for the two concerted hydrogen transfers. Looking at the TS4 geometry (see Figure 3) we note that the O-H bond is almost broken (1.480 Å) while the S-H one is forming (1.697 Å). In INT4 topology is possible to note that the HS-C(=O)-O species is already formed and lies very far from the metal center (4.297 Å) (see INT4, Figure S2) The final products of the reaction (H_2S and CO_2) can be reached by a proton transfer coming from deep water or from the solution buffer to the sulphur center. We have explored the first possibility that other than to transfer the proton can also restore the nucleophilic OH species. All attempts to find a transition state for this process are failed. From the potential energy scan on the Zn-O_{water} reaction coordinate results that a pentacoordinated complex around the zinc ion center is not possible since by approaching the water to the metal ion, the HS-C(=O)-O species is pushed out from the inner coordination shell of the Zn^{2+}. (see INT5 of Figure S2) Furthermore, starting from the formed Zn-H_2O complex (see INT5 of Figure S2) we have verified the possibility that one hydrogen of the zinc ion coordinated water can be transferred to the sulphur center. Also in this case the potential energy scan and the optimization procedure didn't give positive answers. So, we think that C-S bond cleavage of HS-C(=O)-O should be caused by a proton arising from buffer. We have considered this possibility by approaching a H_3O^+ to the sulphur atom obtaining the S-C breaking without energy barrier. A comparison between the experimental derived energy barrier ($\Delta G^{\#}$=6.5 kcal/mol) and the highest barrier of our PES (INT3-TS4-INT4) shows as M06 considerably overestimates the required energy, while B3LYP-D3 gives a value more close to the experimental counterpart. The difference of about 4 kcal/mol can be explained by the different experimental and theoretical conditions. Although COSase belongs to the β-CA superfamily, it shows lower CO_2 hydration activity with respect to β-CA.[23] Its reactivity against CO_2 substrate has been studied and resulted to be poorly efficient.[7] Comparison between the present and previous theoretical study on CA modelled by very similar cluster,[16] evidences as the degradation of COS by COSase has a PES that lies below the Michaelis-Menten complex energy, while the corresponding PES of CO_2 hydration by β-CA gives an energetic profile above the initial complex. This means that the first reaction is thermodynamically favorable.

Conclusions

On the basis of our computational study we propose a reaction mechanism for the degradation of COS in H_2S+CO_2 products.

The found highest barrier for the different considered steps is that related to the proton shift process.

The restoring of the catalytic cycle cannot be done by the proton transfer from deep water to the sulphur center of the HS-C(=O)-O moiety (intermediate) probably since a penta-coordinated intermediate is not found.

The data obtained by using the B3LYP-D3 exchange-correlation functional seem to give results more close to the experimental kinetic study.

Computational Section

The optimizations have been performed by employing the hybrid B3LYP exchange-correlation functional,[33,34] coupled with the 6-31+G(d,p) basis set for the C, N, O and H atoms while the SDD pseudo-potential for the Zn ion.[35] In order to have a more realistic energetic path, the reported potential energy surfaces include the more extended 6-311+G(2d,2p) basis set, the dispersion contributions as suggested by Grimme et al.[36,37] and the enzyme environment by using the polarizable continuum model.[38,39] The dielectric constant value of 4 was chosen by taking into account the coupled effect of the protein itself and the water medium surrounding the protein according to previous numerous suggestions.[40-43,13,15-24] Furthermore, the PES has been also computed by using the M06[44] exchange-correlation functional coupled with the PCM solvation procedure and the same basis set as above mentioned. Transition states (only one imaginary frequency) and intermediates (no imaginary

frequency) have been subjected to the vibrational analysis in order to check their nature of stationary points on the considered PES and to obtain the zero point corrections. Furthermore, every transition state has been confirmed to connect reactants and products by the intrinsic reaction coordinate (IRC)[45] calculations. All the computations have been carried out by using Gaussian 09 code.[46]

Acknowledgements

This work was financially supported through Department of Chemistry and Chemical Technologies of the Università della Calabria (Italy). P. Piazzetta gratefully acknowledges Commissione Europea, Fondo Sociale Europeo, and Regione Calabria for the financial support.

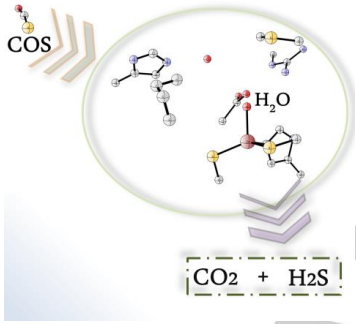
Keywords: carbonic anhydrase •carbonyl sulphide hydrolase • carbonyl sulfide • potential energy surface • density functional theory

[1] M. Chin, D. D. Davis, *J. Geophys. Res.: Atmos.* **1995**, *100*, 8993−9005.
[2] R. P. Turco, R. C. Whitten, O. B. Toon, J. B. Pollack, P. Hamill, *Nature* **1980**, *283*, 283−286.
[3] S. Solomon, R. W. Sanders, R. R. Garcia, J. G. Keys, *Nature* **1993**, *363*, 245−248.
[4] V. Alterio, A. Di Fiore, K. D'Ambrosio, C. T. Supuran, G. De Simone, *Chem. Rev.* **2012**, *112*, 4421−4468 and references therein.
[5] M. J.Smeulders,; T. R. M. Barends, A. Pol, A. Scherer, M. H. Zandvoort, A. Udvarhelyi, A. F. Khadem, A. Menzel, J. Hermans, R. L. Shoeman, H. J. C. T. Wessels, L. P. van den Heuvel, L. Russ, I. Schlichting, M. S. M. Jetten, H. J. M. Op den Camp, *Nature* **2011**, *478*, 412−416.
[6] H. Huber, K. O. Stetter, Genus II. Acidianus Segerer, Neuner, Kristjansson and Stetter 1986, 561VP. In Bergey's *Manual of Systematic Bacteriology*, 2nd ed.; Garrity, G. M., Boone, D. R., Castenholz, R. W., Eds.; Springer: New York, 2001; Vol.1, pp 202−204.
[7] T. Ogawa, K. Noguchi, M. Saito, Y. Nagahata, H. Kato, A. Ohtaki, H. Nakayama, N. Dohmae, Y. Matsushita, M. Odaka, M. Yohda, H. Nyunoya, and Y. Katayama, *J. Am. Chem. Soc.* **2013**, *135*, 3818−3825.
[8] A. S. Covarrubias, A. M. Larsson, M. Högbom, J. Lindberg, T. Bergfors, C. Björkelid, S. L. Mowbray, T. Unge, T. A. Jones, *J. Biol. Chem.* **2005**, *280*, 18782−18789.
[9] D. P. Kelly, A. P. Wood, E. Stackebrandt, Genus II. Thiobacillus Beijerinck 1904b, 597AL. In *Bergey's Manual of Systematic Bacteriology*,2nd ed.; G. M., Garrity, D. J., Brenner, N. R., Krieg, J. T., Staley, Eds.; Springer: New York, **2005**; *Vol. 2*, Part C, pp 764−769.
[10] Y. Katayama, Y. Narahara, Y. Inoue, F. Amano, T. Kanagawa, H. Kuraishi, *J. Biol. Chem.* **1992**, *267*, 9170−9175.
[11] S.-J. Kim, Y. Katayama, *Water Res.* **2000**, *34*, 2887−2894.
[12] R.-Z. Liao, J.-G. Yu, F. Himo, *PNAS*, **2010**, *107*(52), 22523-22527.
[13] M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, P. E. M. Siegbahn *Chem. Rev.* **2014**, *114*, 3601−3658
[14] M. Bruschi, M. Tiberti, A. Guerra, L. De Gioia *J. Am. Chem. Soc*., **2014**, *136* (5), 1803–1814.
[15] M. J. Ramos, P. A. Fernandes, *Acc. Chem. Res.* **2008**, *41*(6), 689-698.
[16] O. Amata; T. Marino; N. Russo; M. Toscano *Phys. Chem. Chem. Phys.*, **2011**, *13*, 3468–3477.
[17] Y. Xu, L. Feng, P. D. Jeffrey, Y. Shi, F. M. M. Morel, *Nature*, **2008**, *452*, 56–61.
[18] M. S. Kimber and E. F. Pai, *EMBO J.*, **2000**, *19*, 1407–1418.
[19] P. Piazzetta, T. Marino, N. Russo *Inorg. Chem.* **2014**, *53*, 3488−3493
[20] O. Amata, T. Marino, N. Russo, M. Toscano, *J. Am. Chem. Soc.* **2009**, *131*, 14804–14811.
[21] O. Amata, T. Marino, N. Russo, M. Toscano, *J. Am. Chem. Soc.* **2011**, *133*, 17824-17831.
[22] F. Himo, , P. E. M. Siegbahn *J. Biol. Inorg. Chem.* **2009**, *14*, 643-651.
[23] P. E. M. Siegbahn, F. Himo, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 323-336.
[24] T. Marino, N. Russo, M. Toscano *J. Am. Chem. Soc.* **2005**, *127*, 4242-4253.
[25] P. Piazzetta, T. Marino, N. Russo *Phys. Chem. Chem. Phys.*, **2014**, *16,* 16671-16676.
[26] T. Marino, N. Russo, M. Toscano, *Chem. - A Eur. J.* **2013**, *19,* 2185-2192
[27] I. Bertini, C. Luchinat, *Acc. Chem. Res.* **1983**, *16*, 272−279.
[28] D. N. Silverman, S. Lindskog, *Acc. Chem. Res.* **1988**, *21*, 30−36.
[29] T.; Minakuchi, I. Nishimori, D. Vullo, A. Scozzafava, C. T. Supuran, *J. Med. Chem.* **2009**, *52*, 2226−2232.
[30] A. Bottoni, C. Z. Lanza, G. P. Miscione, D. Spinelli, *J. Am. Chem. Soc.* **2004**, *126*, 1542.
[31] G. P. Miscione, M. Stenta, D. Spinelli, E. Anders, A. Bottoni, *Theor. Chem. Acc.* **2007**, *118*, 193−201.
[32] M. J. Smeulders, T. R. M. Barends, A. Pol, A. Scherer, M. H. Zandvoort, A. Udvarhelyi, A. F. Khadem, A. Menzel, J. Hermans, R. L. Shoeman, H. J. C. T. Wessels, L. P. van den Heuvel, L. Russ, I. Schlichting, M. S. M. Jetten, H. J. M. Op den Camp, *Nature* **2011**, *478*, 412−416.
[33] A. D. J. Becke, *J. Chem. Phys.* **1993**, *98*, 5648-5652.
[34] C. T. Lee, W. T. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785-789.
[35] D. Andrae, U. Haussermann, M. Dolg, H. Stoll, H. Preuss, *Theor. Chim. Acta* **1990**, *77*, 123-141.
[36] S. Grimme, J. Antony, S. Ehrlich, H. Krieg *J. Chem. Phys.* **2010**, *132*, 154104.
[37] S. Grimme, S. Ehrlich, L. Goerigk *J. Comput. Chem.* **2011**, *32*, 1456-1465.
[38] V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
[39] M. Cossi, N. Rega, G. Scalmani, *J. Comput. Chem.* **2003**, *24*, 669–681.
[40] R. A. Torres, T. Lovell, L. Noodleman, D. A. Case, *J. Am. Chem. Soc.* **2003**, *125*, 1923-1936.
[41] P. E. M. Siegbahn, M. R. A. Blomberg, *Chem. Rev.* **2000**, *100*, 421-437.
[42] P. E. M. Siegbahn *J. Comput. Chem.* **2001**, *22*, 1634-1645.
[43] A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley: New York, 1991.
[44] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.*, **2008**, *120*, 215-41.
[45] (a) K. J. Fukui, *Phys. Chem.* **1970**, *74*, 4161-4163. (b) C. Gonzalez, H. B. Schlegel *J. Chem. Phys.* **1989**, *90*, 2154-2161.
[46] Gaussian 09, Revision A.1, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

Entry for the Table of Contents

FULL PAPER

The working mechanism of the novel characterized enzyme carbonyl sulfide hydrolase (COSase), that efficiently converts COS in H_2S and CO_2, has been investigated at density functional theory level. The B3LYP-D3 and M06 potential energy surfaces show a very similar shape. Comparison with CO_2 hydration process catalysed by another carbonic anhydrase having a similar active site has been done.



Paolo Piazzetta, Tiziana Marino and Nino Russo, Corresponding Author**

Page No. – Page No.

The working mechanism of the β-carbonic anhydrase degrading carbonyl sulfide (COSase): a theoretical study

The working mechanism of the β-carbonic anhydrase degrading carbonyl sulfide (COSase): a theoretical study

Paolo Piazzetta, Tiziana Marino and Nino Russo

Table of contents

S-1

Cartesian coordinates (Å) of optimized structures of stationary points for conversion of COS in H_2S and CO_2 by COSase taking into account the X-ray position of water molecule.

ES

C -0.26374 0.38799 3.78708
S 0.96173 0.66025 5.12631
C 1.06081 -1.01934 5.84850
C -1.90782 6.45792 -3.28464
C -0.73398 5.79704 -2.61613
N 0.35990 6.52946 -2.19122
C -0.54143 4.46190 -2.32072
C 1.18772 5.64759 -1.65375
N 0.69117 4.38099 -1.70523
C -8.51357 1.62020 2.03876
C -9.19856 0.27082 1.75840
C -7.00561 1.63560 1.75549
C -9.19758 -0.14557 0.28105
C -0.48610 -2.36338 -4.24065
S -1.62008 -1.74486 -2.91343
C 1.87715 0.38780 -2.04416
C 2.86963 1.08378 -1.13627
O 3.94514 0.50177 -0.81113
O 2.54670 2.24481 -0.67492
C 6.53480 1.58832 1.73263
N 5.11470 1.39850 1.46680
C 4.15850 1.46629 2.38629
N 4.46172 1.42775 3.71940
N 2.88747 1.62325 2.01788
C 4.24010 -5.12692 -2.27190
C 3.16423 -4.23889 -1.75302
N 3.39251 -3.15604 -0.91276
C 1.80225 -4.27829 -1.91136
C 2.18792 -2.59568 -0.60976
N 1.21616 -3.25500 -1.20210
C -0.25719 -4.98231 1.23594
S -1.68482 -4.60993 0.11889
S -3.58926 0.08441 -0.00099
C -2.30379 0.92000 -0.35428
Zn -0.84166 -2.64539 -0.88202
O -0.33571 -1.34483 0.54269
H -0.48778 1.36835 3.35922
H -1.18589 -0.02200 4.20690
H 1.77367 -0.97520 6.67563
H 0.08418 -1.32063 6.23658
H 1.40320 -1.74943 5.11059
H -2.36714 7.20478 -2.62573
H -1.59869 6.97711 -4.20001
H -1.15474 3.58740 -2.47495
H 1.16436 3.53871 -1.36212
H 2.14872 5.87088 -1.21012
H -9.00661 2.40247 1.44293
H -6.78219 1.45086 0.70041
H -6.48563 0.86718 2.33975
H -6.56442 2.60414 2.01588
H -10.23709 0.32200 2.11434
H -8.70842 -0.51060 2.35636
H -9.74351 -1.08467 0.13603
H -9.67763 0.61880 -0.34321
H -8.18317 -0.29688 -0.10087
H -0.57753 -3.44663 -4.37729
H 0.56091 -2.12587 -4.02549
H 1.01159 0.06118 -1.45602
H 2.32497 -0.47713 -2.53425
H 6.68740 2.39191 2.46006
H 7.02055 0.67156 2.09641
H 4.77474 1.18196 0.48987
H 2.14775 1.39305 2.66599
H 2.65854 1.88585 1.03848
H 3.67498 1.37044 4.35736
H 5.28562 0.91558 3.99762
H 3.79918 -5.90116 -2.90563
H 4.97592 -4.58261 -2.87910
H 4.28651 -2.75344 -0.67150
H 1.20811 -4.98383 -2.47424
H 2.03892 -1.74124 0.03325
H -0.56867 -5.72314 1.98144
H 0.59145 -5.39365 0.67843
O 4.71190 4.01724 -0.52749
H -0.92875 -1.60051 1.26140
O -1.39556 1.60209 -0.61832
H 4.79857 4.35806 -1.42531
H 3.95483 3.38737 -0.56957
H 0.08197 -4.08234 1.75892
H -0.76447 -1.87517 -5.18018
H 4.78631 -5.63192 -1.46380
H 7.00955 1.88395 0.79564
H 1.49621 1.08603 -2.79563
H -2.67184 5.72051 -3.55097
H -8.68293 1.88766 3.09110
H 0.10945 -0.26616 2.99246

Ts1

C -0.54000 0.56500 3.88300
S 0.76800 0.74100 5.15600
C 0.81000 -0.96900 5.80300
C -2.05000 6.33100 -3.50700
C -0.86900 5.70700 -2.81500
N 0.18600 6.47700 -2.36200
C -0.63100 4.37700 -2.52400
C 1.03500 5.62400 -1.81100
N 0.59100 4.34000 -1.88200
C -8.71300 1.48500 1.54600
C -9.41600 0.13500 1.32100
C -7.18400 1.43200 1.42100
C -9.28100 -0.42100 -0.10300
C -0.37200 -2.46000 -4.31500
S -1.36500 -1.67200 -2.96300
C 1.84700 0.37500 -2.07500
C 2.83200 1.10600 -1.18700
O 3.88300 0.51800 -0.79400
O 2.52700 2.30000 -0.80800
C 6.34600 1.73700 1.83700
N 4.94400 1.51600 1.50600
C 3.93700 1.57900 2.37100
N 4.16400 1.55300 3.71900
N 2.68700 1.71800 1.92800
C 4.35700 -5.07800 -2.16200
C 3.28700 -4.17900 -1.64100
N 3.53600 -3.02600 -0.90800
C 1.91800 -4.25900 -1.71100
C 2.33700 -2.46700 -0.58100
N 1.34700 -3.18900 -1.05500
C -0.25500 -5.01300 1.19600
S -1.50400 -4.81700 -0.15600
S -3.14600 0.56400 0.84700
C -1.60000 0.77100 0.51900
Zn -0.79800 -2.71300 -0.94400
O -0.74100 -1.47100 0.62500
H -0.68800 1.55000 3.43400
H -1.47700 0.25800 4.35400
H 1.56900 -1.00500 6.58800
H -0.16000 -1.23100 6.23500
H 1.06600 -1.68100 5.01500
H -2.54900 7.05700 -2.85300
H -1.73800 6.86700 -4.41100
H -1.20600 3.48200 -2.70600
H 1.09100 3.51700 -1.52800
H 1.97700 5.88300 -1.34500
H -9.11300 2.22000 0.83300
H -6.86200 1.14500 0.41500
H -6.75600 0.70500 2.12300
H -6.73600 2.40700 1.63900
H -10.48300 0.24900 1.56000
H -9.02000 -0.59900 2.03700
H -9.84300 -1.35600 -0.21600
H -9.66700 0.29100 -0.84200
H -8.23900 -0.63500 -0.36100
H -0.67500 -3.49800 -4.48600
H 0.70100 -2.44100 -4.09500
H 0.97800 0.05500 -1.49200
H 2.30000 -0.49900 -2.54200
H 6.44500 2.52900 2.58600
H 6.83900 0.82500 2.20400
H 4.66100 1.27400 0.52000
H 1.90800 1.49700 2.53100
H 2.51200 1.97900 0.94000
H 3.34400 1.47700 4.31300
H 4.98100 1.05600 4.04300
H 3.90000 -5.91000 -2.70500
H 5.03300 -4.56100 -2.85700
H 4.42900 -2.57900 -0.75600
H 1.30200 -5.02100 -2.16700
H 2.24100 -1.55200 -0.02000
H -0.56400 -5.83700 1.85100
H 0.73700 -5.24800 0.79500
O 4.73500 4.00600 -0.58300
H -1.38000 -1.89200 1.21500
O -0.55800 1.25400 0.29400
H 4.82200 4.40700 -1.45600
H 3.95800 3.40300 -0.65900
H -0.17300 -4.10200 1.79600
H -0.54600 -1.89300 -5.23500
H 4.96900 -5.50300 -1.35600
H 6.85100 2.06600 0.92700
H 1.46100 1.04600 -2.84800
H -2.78300 5.57000 -3.79500
H -8.97700 1.85600 2.54600
H -0.26400 -0.14900 3.10200

Int1

C 0.55153 1.78877 5.45359
S 1.91710 0.63655 5.85002
C 1.05421 -0.95473 5.57097
C -2.66419 6.44971 -3.34441
C -1.39707 5.83915 -2.81244
N -0.29880 6.62070 -2.50379
C -1.11033 4.51330 -2.55088
C 0.62313 5.77753 -2.06685
N 0.18533 4.48913 -2.07442
C -8.82237 1.75732 1.92622
C -9.69921 0.49330 1.89967
C -7.31417 1.48073 1.97835
C -9.54084 -0.35863 0.63293
C -1.05925 -2.32737 -4.41098
S -1.97914 -1.59828 -2.98060
C 1.35489 0.48538 -2.35215
C 2.44133 1.24870 -1.61914
O 3.53627 0.67261 -1.35096
O 2.17224 2.44634 -1.23571
C 6.19642 1.80964 1.14145
N 4.77055 1.63668 0.90844
C 3.83325 1.54199 1.85493
N 4.17628 1.29645 3.14882
N 2.55329 1.74338 1.54105
C 3.85144 -4.96125 -2.73830
C 2.86119 -4.03629 -2.12085
N 3.21393 -2.85853 -1.47527
C 1.49559 -4.10078 -2.00941
C 2.07775 -2.27034 -1.01912
N 1.02364 -2.99517 -1.32878
C -0.42889 -4.94644 1.03396
S -1.72214 -4.67562 -0.25613
S -1.05400 -0.60601 0.48670
C 0.09341 -0.81219 1.74583
Zn -1.10787 -2.60656 -1.07201
O 0.64031 -2.06564 1.86276
H 0.95891 2.80192 5.49383
H -0.25003 1.69860 6.19200
H 1.80030 -1.75035 5.64653
H 0.29027 -1.11413 6.33716
H 0.60394 -0.94556 4.57762
H -3.07863 7.17565 -2.63445
H -2.47939 6.98271 -4.28500
H -1.69690 3.61204 -2.65256
H 0.73632 3.67147 -1.78766
H 1.61540 6.04502 -1.72944
H -9.05173 2.37005 1.04233
H -6.96205 0.94744 1.08982
H -7.05471 0.87150 2.85287
H -6.74354 2.41348 2.04355
H -10.75219 0.79169 1.99989
H -9.47014 -0.12026 2.78297
H -10.22323 -1.21627 0.64596
H -9.76277 0.22934 -0.26642
H -8.52462 -0.75157 0.52755
H -1.21839 -3.40790 -4.48808
H 0.01656 -2.13380 -4.34557
H 0.57755 0.17139 -1.64934
H 1.75123 -0.39000 -2.86604
H 6.36791 2.48555 1.98526
H 6.71415 0.85814 1.33302
H 4.42416 1.46834 -0.06327
H 1.79852 1.27356 2.05208
H 2.33679 2.09587 0.59944
H 3.43933 1.16994 3.84280
H 5.03715 0.80246 3.32799
H 3.32908 -5.81194 -3.18383
H 4.43246 -4.47605 -3.53365
H 4.12370 -2.42024 -1.45012
H 0.81700 -4.87129 -2.34684
H 2.08281 -1.32919 -0.49832
H -0.46294 -5.99875 1.33819
H 0.57192 -4.72488 0.65486
O 4.38838 4.15430 -1.24759
H 1.27422 -2.01288 2.59650
O 0.49057 0.05799 2.54880
H 4.44654 4.42796 -2.17030
H 3.61106 3.54803 -1.21579
H -0.60967 -4.31993 1.90980
H -1.43761 -1.85791 -5.32435
H 4.56267 -5.35722 -2.00139
H 6.62467 2.26593 0.24707
H 0.85945 1.13819 -3.07731
H -3.42184 5.68142 -3.52951
H -9.10548 2.36314 2.79825
H 0.16987 1.58159 4.45124

Ts2

C -0.02600 1.67900 5.05900
S 1.58100 0.89700 5.45700
C 1.06300 -0.85100 5.62300
C -2.56800 6.46100 -3.40400
C -1.32300 5.85500 -2.81500
N -0.24200 6.64100 -2.46000
C -1.04300 4.53100 -2.53900
C 0.66400 5.80300 -1.98200
N 0.23200 4.51200 -2.00500
C -8.70900 1.86200 1.52600
C -9.70200 0.68600 1.56200
C -7.26300 1.48100 1.86800
C -9.44800 -0.38800 0.49600
C -0.86100 -2.30700 -4.38600
S -1.96300 -1.72200 -3.02300
C 1.49000 0.54200 -2.30300
C 2.55300 1.30300 -1.53800
O 3.66900 0.75700 -1.29900
O 2.23800 2.46900 -1.08900
C 6.25600 1.93400 1.26700
N 4.85100 1.65400 1.00500
C 3.91200 1.51700 1.93600
N 4.23800 1.32600 3.24100
N 2.62100 1.62800 1.60200
C 4.04400 -4.88300 -2.61000
C 2.98600 -3.99900 -2.04400
N 3.25700 -2.82900 -1.34600
C 1.61800 -4.10500 -2.02900
C 2.07600 -2.28500 -0.95200
N 1.06900 -3.03100 -1.35400
C -0.30200 -4.88500 1.08600
S -1.61200 -4.73700 -0.20400
S -1.09100 -0.67000 0.42500
C -0.37300 -1.07900 1.94800
Zn -1.06900 -2.69000 -1.11900
O 1.01200 -1.30500 1.88700
H 0.18000 2.71200 4.76800
H -0.67500 1.67700 5.93800
H 1.96500 -1.46800 5.57300
H 0.58300 -1.01500 6.59200
H 0.37100 -1.11500 4.81900
H -3.01700 7.18400 -2.71200
H -2.34200 6.99600 -4.33400
H -1.62100 3.62700 -2.66400
H 0.77400 3.69800 -1.69700
H 1.63900 6.07600 -1.60100
H -8.73900 2.32500 0.53000
H -6.83900 0.77600 1.14600
H -7.20100 1.01500 2.86000
H -6.61400 2.36300 1.87600
H -10.72000 1.07900 1.43800
H -9.67300 0.22400 2.56000
H -10.21200 -1.17400 0.53800
H -9.47200 0.04500 -0.51200
H -8.47400 -0.87100 0.62300
H -0.90600 -3.39400 -4.50800
H 0.18200 -2.01500 -4.22300
H 0.71200 0.20000 -1.61100
H 1.90900 -0.31600 -2.82800
H 6.35800 2.60500 2.12600
H 6.83900 1.02100 1.45000
H 4.52200 1.49300 0.02200
H 1.92200 1.10500 2.11500
H 2.37600 1.97800 0.66400
H 3.47900 1.23200 3.91600
H 5.12000 0.89500 3.46800
H 3.58000 -5.73700 -3.10900
H 4.66500 -4.36300 -3.35200
H 4.15100 -2.37000 -1.24500
H 0.99000 -4.88600 -2.43200
H 1.99900 -1.38100 -0.37400
H -0.35000 -5.89200 1.51600
H 0.69800 -4.73600 0.66900
O 4.42600 4.21000 -0.95700
H 1.25000 -1.65500 2.76200
O -0.89800 -1.13800 3.05700
H 4.50000 4.55600 -1.85300
H 3.65300 3.59700 -0.98700
H -0.45900 -4.16300 1.89300
H -1.20400 -1.84400 -5.31700
H 4.71300 -5.27500 -1.83200
H 6.66100 2.44100 0.38900
H 0.99500 1.20400 -3.02000
H -3.31200 5.69000 -3.62400
H -9.05300 2.63200 2.23100
H -0.51200 1.14900 4.23700

Int2

C 1.64527 3.62380 3.73675
S 3.20173 2.87277 4.34169
C 2.58288 1.28650 5.01226
C -2.88741 5.44257 -5.03460
C -2.07810 5.33157 -3.78032
N -1.41264 6.41652 -3.24266
C -1.86147 4.21752 -2.99643
C -0.81104 5.95193 -2.15774
N -1.05086 4.62875 -1.96201
C -6.24558 1.27800 2.30393
C -7.40465 0.62051 1.53587
C -6.35454 2.80416 2.42616
C -8.77168 0.73818 2.22521
C -1.39139 -3.25264 -3.32990
S -2.34788 -4.22151 -2.08308
C 1.33300 -0.04114 -2.83730
C 2.40259 0.52844 -1.91086
O 3.60528 0.13563 -1.99788
O 2.02033 1.40480 -1.06007
C 6.73656 2.12692 -0.96648
N 5.42335 1.69059 -0.51454
C 4.92040 2.00806 0.68329
N 5.71504 2.61619 1.61628
N 3.64778 1.76853 0.95473
C 3.78676 -5.42249 -1.80396
C 2.69776 -4.50120 -1.35970
N 2.88113 -3.15093 -1.08164
C 1.37005 -4.73414 -1.10038
C 1.68480 -2.63857 -0.68011
N 0.75746 -3.57159 -0.68405
C 0.31232 -3.91594 2.46335
S -1.37219 -4.44972 1.92205
S -1.67610 -0.96002 -0.18799
C -1.00827 -0.24230 1.24020
Zn -1.35705 -3.45157 -0.14238
O -1.60995 0.99360 1.53496
H 0.93683 3.73520 4.56163
H 1.19658 3.03743 2.93077
H 3.45265 0.72368 5.35868
H 1.91717 1.46840 5.85984
H 2.05965 0.70574 4.24774
H -3.69399 6.17780 -4.92335
H -2.26774 5.76783 -5.87952
H -2.20544 3.19704 -3.07910
H -0.68513 4.02409 -1.20803
H -0.19375 6.53375 -1.48651
H -6.18431 0.83403 3.30820
H -7.24318 3.11106 2.99049
H -6.40928 3.27436 1.43636
H -5.47952 3.22012 2.93845
H -7.16836 -0.44063 1.38843
H -7.46336 1.06027 0.53008
H -9.54693 0.20985 1.65794
H -8.74238 0.30184 3.23199
H -9.09240 1.78148 2.32551
H -0.32224 -3.48154 -3.28333
H -1.52391 -2.17564 -3.18895
H 0.43943 -0.30844 -2.26653
H 1.70713 -0.90529 -3.38860
H 6.87043 3.20528 -0.82328
H 7.55434 1.59393 -0.46012
H 4.79114 1.13406 -1.14234
H 3.31587 1.96303 1.89150
H 2.96027 1.59040 0.14973
H 5.30563 2.76559 2.53124
H 6.70950 2.45147 1.59166
H 3.37176 -6.42062 -1.96629
H 4.24890 -5.09687 -2.74512
H 3.70057 -2.58638 -1.26471
H 0.80988 -5.65423 -1.18674
H 1.53645 -1.61192 -0.38377
H 0.43619 -4.19954 3.51536
H 1.09383 -4.41629 1.88340
O -0.04285 2.87226 -0.01687
H -1.21741 1.24994 2.38641
O -0.10863 -0.61288 1.98825
H 0.60714 2.29014 -0.47400
H -0.71755 2.27275 0.35061
H 0.42824 -2.83460 2.37815
H -1.75698 -3.51486 -4.32865
H 4.58403 -5.50995 -1.05422
H 6.81269 1.90959 -2.03314
H 1.04326 0.73843 -3.55196
H -3.33874 4.47967 -5.29477
H -5.30385 1.02566 1.80256
H 1.89741 4.61348 3.34973

Ts3

C 0.40100 2.67500 3.55600
S 2.03000 2.98900 4.33300
C 2.26100 1.40900 5.22200
C -3.42000 5.19700 -4.85400
C -2.58600 5.22800 -3.61200
N -1.98900 6.39000 -3.16100
C -2.27600 4.18600 -2.76300
C -1.33400 6.04000 -2.06400
N -1.47800 4.72000 -1.77700
C -6.39400 0.32200 2.07400
C -7.47800 -0.41400 1.26900
C -6.71700 1.79300 2.37100
C -8.81600 -0.57500 2.00400
C -0.78600 -3.27500 -3.42500
S -1.56100 -4.50500 -2.28400
C 1.45000 0.25700 -2.71200
C 2.40400 0.97200 -1.76000
O 3.64000 0.68300 -1.75100
O 1.89400 1.85600 -0.99100
C 6.44000 3.06500 -0.56500
N 5.17400 2.46000 -0.17900
C 4.55300 2.73400 0.97400
N 5.21400 3.44100 1.94600
N 3.30100 2.36200 1.16200
C 4.59400 -4.77600 -1.79400
C 3.37500 -4.03300 -1.35100
N 3.36700 -2.69500 -0.97000
C 2.08100 -4.45400 -1.17300
C 2.10000 -2.37200 -0.59000
N 1.30400 -3.41500 -0.70400
C 0.81400 -3.94300 2.39800
S -0.80800 -4.60500 1.80800
S -1.65100 -1.27800 -0.32300
C -0.92000 -0.35300 0.93200
Zn -0.81600 -3.63500 -0.27800
O -1.66800 0.84600 1.23200
H -0.38400 2.69300 4.31600
H 0.39900 1.71500 3.03200
H 3.23800 1.44300 5.71000
H 1.48900 1.28500 5.98700
H 2.22700 0.56400 4.53000
H -4.27700 5.87800 -4.77600
H -2.83800 5.50800 -5.73000
H -2.54300 3.14000 -2.77600
H -1.03900 4.18100 -1.01200
H -0.74500 6.70600 -1.44800
H -6.22500 -0.21400 3.02000
H -7.61500 1.90100 2.98900
H -6.88200 2.35400 1.44300
H -5.89200 2.27800 2.90500
H -7.09800 -1.40700 0.99800
H -7.64300 0.11700 0.32100
H -9.53100 -1.14700 1.40100
H -8.68200 -1.11000 2.95300
H -9.27900 0.39200 2.23100
H 0.30700 -3.30700 -3.36600
H -1.11800 -2.25800 -3.19900
H 0.59600 -0.14400 -2.16000
H 1.95600 -0.54400 -3.25500
H 6.42500 4.15100 -0.41800
H 7.29400 2.64200 -0.01600
H 4.65400 1.82700 -0.83600
H 2.87700 2.50400 2.07500
H 2.70300 2.11500 0.31000
H 4.72500 3.57000 2.82100
H 6.21700 3.35700 2.00900
H 4.31700 -5.80000 -2.05500
H 5.06100 -4.32100 -2.67700
H 4.10500 -2.01300 -1.08700
H 1.64700 -5.42700 -1.35800
H 1.80900 -1.40200 -0.21600
H 0.92500 -4.20700 3.45600
H 1.64900 -4.38400 1.84700
O -0.27900 3.08800 0.14400
H -2.21800 0.67100 2.00700
O 0.11900 -0.52000 1.55200
H 0.42000 2.60300 -0.35000
H -0.84800 2.39500 0.52800
H 0.85600 -2.85600 2.30400
H -1.08400 -3.51600 -4.45100
H 5.35400 -4.82600 -1.00300
H 6.59900 2.86300 -1.62600
H 1.05700 0.98700 -3.42900
H -3.80300 4.18900 -5.04100
H -5.45000 0.26700 1.51900
H 0.22300 3.46500 2.82300

Int3

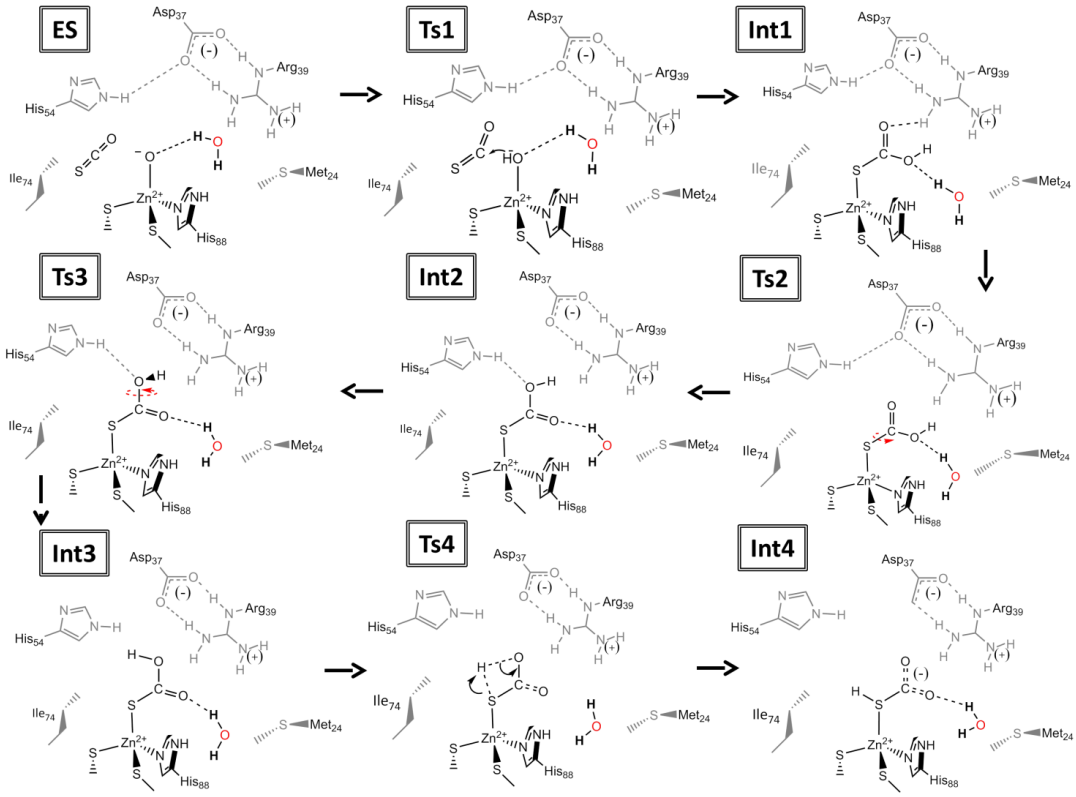
C 1.02670 4.14621 3.89839
S 2.68775 3.49526 4.30427
C 2.24513 1.78965 4.80864
C -4.76009 3.94471 -4.90831
C -4.14817 4.01434 -3.54766
N -4.34103 5.09266 -2.70552
C -3.33605 3.08475 -2.93713
C -3.65127 4.80634 -1.61147
N -3.02823 3.60179 -1.70032
C -6.34823 -0.86968 2.39136
C -7.40026 -1.79445 1.75614
C -6.86650 0.53123 2.74391
C -8.56497 -2.16377 2.68579
C -0.64815 -3.87860 -3.28395
S -1.65441 -3.13166 -1.92437
C 1.01763 0.00660 -2.95373
C 1.82638 1.13992 -2.34612
O 3.09145 1.12449 -2.42515
O 1.17684 2.06821 -1.75098
C 5.62346 3.71837 -1.40197
N 4.39674 3.06037 -0.98851
C 3.86516 3.14356 0.23705
N 4.57246 3.67595 1.27369
N 2.61136 2.75541 0.42039
C 5.01724 -4.38085 -2.05875
C 3.82720 -3.63772 -1.54939
N 3.82316 -2.27373 -1.28648
C 2.56627 -4.06025 -1.21161
C 2.59273 -1.93525 -0.81742
N 1.81134 -2.99311 -0.76739
C 1.49218 -3.89857 2.40137
S 0.05697 -4.72201 1.57473
S -0.68761 -0.80744 0.81507
C 0.54641 0.04418 1.73748
Zn -0.25095 -3.17246 -0.08043
O 0.38613 1.41738 1.80414
H 0.42803 4.24364 4.80836
H 0.52413 3.48276 3.19177
H 3.16294 1.32037 5.17199
H 1.52118 1.82603 5.62737
H 1.85514 1.20961 3.97012
H -5.85517 3.99061 -4.85712
H -4.43152 4.78245 -5.53623
H -2.96273 2.12480 -3.26107
H -2.39346 3.13874 -1.03787
H -3.57598 5.43646 -0.73559
H -5.94617 -1.34961 3.29482
H -7.66250 0.49893 3.49718
H -7.26743 1.03748 1.85675
H -6.06076 1.15689 3.14370
H -6.90245 -2.71455 1.42494
H -7.79575 -1.31831 0.84740
H -9.25588 -2.86029 2.19635
H -8.19863 -2.64764 3.60020
H -9.14465 -1.28345 2.98618
H -0.42320 -4.93267 -3.09056
H 0.29271 -3.34055 -3.43627
H 0.38787 -0.47312 -2.19829
H 1.66496 -0.73725 -3.42080
H 5.63403 4.76840 -1.08827
H 6.52647 3.22499 -1.01181
H 3.89660 2.40600 -1.64643
H 2.24628 2.50314 1.33236
H 2.02820 2.52456 -0.41971
H 4.12878 3.64777 2.19201
H 5.57802 3.59593 1.25906
H 4.76113 -5.43635 -2.18185
H 5.35563 -4.00477 -3.03307
H 4.55517 -1.60900 -1.49176
H 2.15741 -5.06008 -1.23229
H 2.32427 -0.93597 -0.51789
H 1.68091 -4.41231 3.35055
H 2.39638 -3.96998 1.78837
O -1.13354 2.02842 -0.39637
H -0.28413 1.71978 1.14290
O 1.51452 -0.42132 2.32324
H -0.34679 2.03025 -1.00824
H -1.36420 1.07636 -0.30340
H 1.29217 -2.84511 2.60408
H -1.23362 -3.81719 -4.20668
H 5.86608 -4.32310 -1.36506
H 5.67235 3.68405 -2.49205
H 0.33960 0.41795 -3.71022
H -4.48339 3.01322 -5.41263
H -5.50222 -0.77351 1.70054
H 1.15906 5.13396 3.45026

Ts4

C -1.05700 -4.11900 3.78100
S -2.71500 -3.53600 4.29200
C -2.29300 -1.82700 4.80500
C 4.68600 -3.92300 -4.94300
C 4.27900 -3.69800 -3.52100
N 5.00000 -4.21900 -2.46300
C 3.20300 -2.98200 -3.04400
C 4.36500 -3.81900 -1.37300
N 3.27000 -3.06700 -1.66900
C 6.28500 0.82500 2.57700
C 7.26500 1.79700 1.89900
C 6.82300 -0.60300 2.74200
C 8.53800 2.08200 2.70900
C 0.62500 3.90800 -3.23200
S 1.59400 3.03400 -1.92200
C -1.06500 0.02900 -2.94000
C -1.69900 -1.29500 -2.55400
O -2.92500 -1.49300 -2.76000
O -0.93400 -2.16600 -1.99300
C -5.69100 -3.67000 -1.42000
N -4.33400 -3.22600 -1.15400
C -3.74400 -3.25400 0.05100
N -4.45200 -3.57500 1.16700
N -2.43800 -3.03500 0.13900
C -5.03600 4.43200 -1.99400
C -3.90000 3.66600 -1.40100
N -4.05800 2.54300 -0.59800
C -2.54500 3.85600 -1.46300
C -2.82300 2.11100 -0.21200
N -1.89100 2.88700 -0.73000
C -1.50800 3.88100 2.45700
S -0.02700 4.57300 1.59400
S 0.60300 0.70100 0.91900
C -0.73200 -0.35300 1.62100
Zn 0.18500 3.03200 -0.09200
O -0.49700 -1.63000 1.56400
H -0.40700 -4.21300 4.65500
H -0.62500 -3.42000 3.06000
H -3.19500 -1.39700 5.24900
H -1.50700 -1.85500 5.56500
H -1.98900 -1.22300 3.94700
H 5.68800 -3.52200 -5.13700
H 4.71400 -4.99300 -5.18300
H 2.41700 -2.43400 -3.54100
H 2.60000 -2.64000 -1.02800
H 4.65700 -4.04400 -0.35700
H 6.00400 1.22600 3.56100
H 7.70700 -0.63800 3.38900
H 7.10400 -1.03100 1.77100
H 6.06500 -1.25900 3.18500
H 6.74400 2.74300 1.70800
H 7.54200 1.39900 0.91200
H 9.17400 2.81700 2.20100
H 8.29100 2.48500 3.70000
H 9.13900 1.17800 2.86000
H 0.38700 4.93800 -2.94500
H -0.30300 3.37900 -3.47100
H -0.67700 0.54100 -2.05400
H -1.78600 0.67600 -3.44100
H -5.86900 -4.66700 -1.00100
H -6.45400 -2.98000 -1.02700
H -3.80900 -2.68500 -1.88400
H -1.99800 -2.64300 0.98000
H -1.90500 -2.85300 -0.73300
H -3.97000 -3.51900 2.06700
H -5.44200 -3.38400 1.18000
H -4.64500 5.27700 -2.56500
H -5.63900 3.81900 -2.67700
H -4.93200 2.11000 -0.33800
H -2.00300 4.63800 -1.97400
H -2.65100 1.28700 0.47100
H -1.61000 4.38800 3.42200
H -2.42400 4.06000 1.88200
O 1.01100 -1.72900 -0.40600
H 0.45200 -1.91700 0.46600
O -1.75800 0.15100 2.10000
H 0.30900 -1.83800 -1.15600
H 1.04000 -0.58900 -0.09200
H -1.40700 2.80800 2.63700
H 1.24000 3.94100 -4.13600
H -5.70700 4.83400 -1.22300
H -5.82100 -3.72500 -2.50300
H -0.20700 -0.14300 -3.59800
H 3.98700 -3.43800 -5.63100
H 5.35900 0.79200 1.98900
H -1.18100 -5.10200 3.32000

Int4

C 3.79262 2.14318 3.99549
S 4.36382 0.48131 4.51033
C 2.75862 -0.28698 4.95403
C -0.06805 6.42072 -4.80061
C 0.81170 5.88830 -3.71359
N 1.89895 6.59679 -3.24130
C 0.71191 4.68804 -3.04123
C 2.43426 5.82672 -2.30477
N 1.75404 4.66249 -2.14238
C -4.92361 4.24891 2.52971
C -6.27316 4.30873 1.79476
C -4.24586 5.61290 2.71860
C -7.37803 5.05175 2.55811
C -3.12311 -1.90290 -3.33958
S -4.46551 -1.81031 -2.06034
C 0.85718 -0.51770 -2.87359
C 2.19514 -0.96163 -2.29497
O 2.49549 -2.18661 -2.21381
O 2.95952 -0.02714 -1.87542
C 6.65251 -1.43135 -1.14783
N 5.30250 -1.52481 -0.60591
C 4.92921 -0.96608 0.56839
N 5.81771 -0.83913 1.58568
N 3.68950 -0.53977 0.68790
C 0.28555 -6.43667 -2.04496
C -0.31438 -5.14390 -1.60559
N 0.36487 -3.94159 -1.48084
C -1.59732 -4.87644 -1.19711
C -0.50624 -3.01957 -1.01046
N -1.70293 -3.55107 -0.83298
C -1.87283 -3.54449 2.37410
S -3.63253 -3.09362 1.97246
S -0.86392 0.64700 0.80336
C 0.68787 0.27379 1.78671
Zn -3.46717 -2.61824 -0.20961
O 1.64367 1.09454 1.67097
H 3.40097 2.68758 4.85958
H 3.02586 2.05099 3.22210
H 2.98130 -1.29344 5.31815
H 2.29193 0.28413 5.76176
H 2.08921 -0.35383 4.09214
H -0.52782 7.37350 -4.51009
H 0.50253 6.60468 -5.71964
H 0.00793 3.87296 -3.12052
H 1.98187 3.87631 -1.50916
H 3.31206 6.06899 -1.72121
H -5.06883 3.77322 3.51003
H -4.83933 6.28444 3.35001
H -4.09209 6.11228 1.75370
H -3.26384 5.49981 3.19084
H -6.60748 3.28334 1.59167
H -6.12690 4.78090 0.81266
H -8.32643 5.02259 2.00874
H -7.55068 4.59602 3.54147
H -7.12632 6.10578 2.72139
H -2.68809 -2.90309 -3.39622
H -2.33358 -1.17633 -3.13737
H 0.24720 -0.07814 -2.07701
H 0.31307 -1.34320 -3.33624
H 6.99575 -0.39231 -1.22996
H 7.37388 -1.99595 -0.54558
H 4.54027 -1.55946 -1.28928
H 3.28082 -0.13327 1.52824
H 3.20576 -0.35219 -0.20539
H 5.45670 -0.49393 2.47806
H 6.60256 -1.47123 1.62405
H -0.48937 -7.20727 -2.06778
H 0.72433 -6.37005 -3.04847
H 1.30361 -3.65775 -1.79035
H -2.43889 -5.55138 -1.13271
H -0.24100 -1.99629 -0.79512
H -1.87017 -3.89597 3.40992
H -1.51202 -4.34900 1.73062
O 2.35874 2.41770 -0.65757
H 2.05075 2.06079 0.20515
O 0.61910 -0.76591 2.45054
H 2.54508 1.63120 -1.21378
H -0.40738 1.74874 0.17454
H -1.21024 -2.67996 2.29995
H -3.58049 -1.66236 -4.30327
H 1.07566 -6.77279 -1.36183
H 6.64011 -1.87629 -2.14387
H 1.02802 0.26955 -3.61458
H -0.87108 5.71411 -5.03496
H -4.24817 3.58928 1.97196
H 4.65961 2.67857 3.60024



Scheme S1. Proposed reaction mechanism for the conversion of COS in H_2S and CO_2 by considering water molecule in not X-ray position.

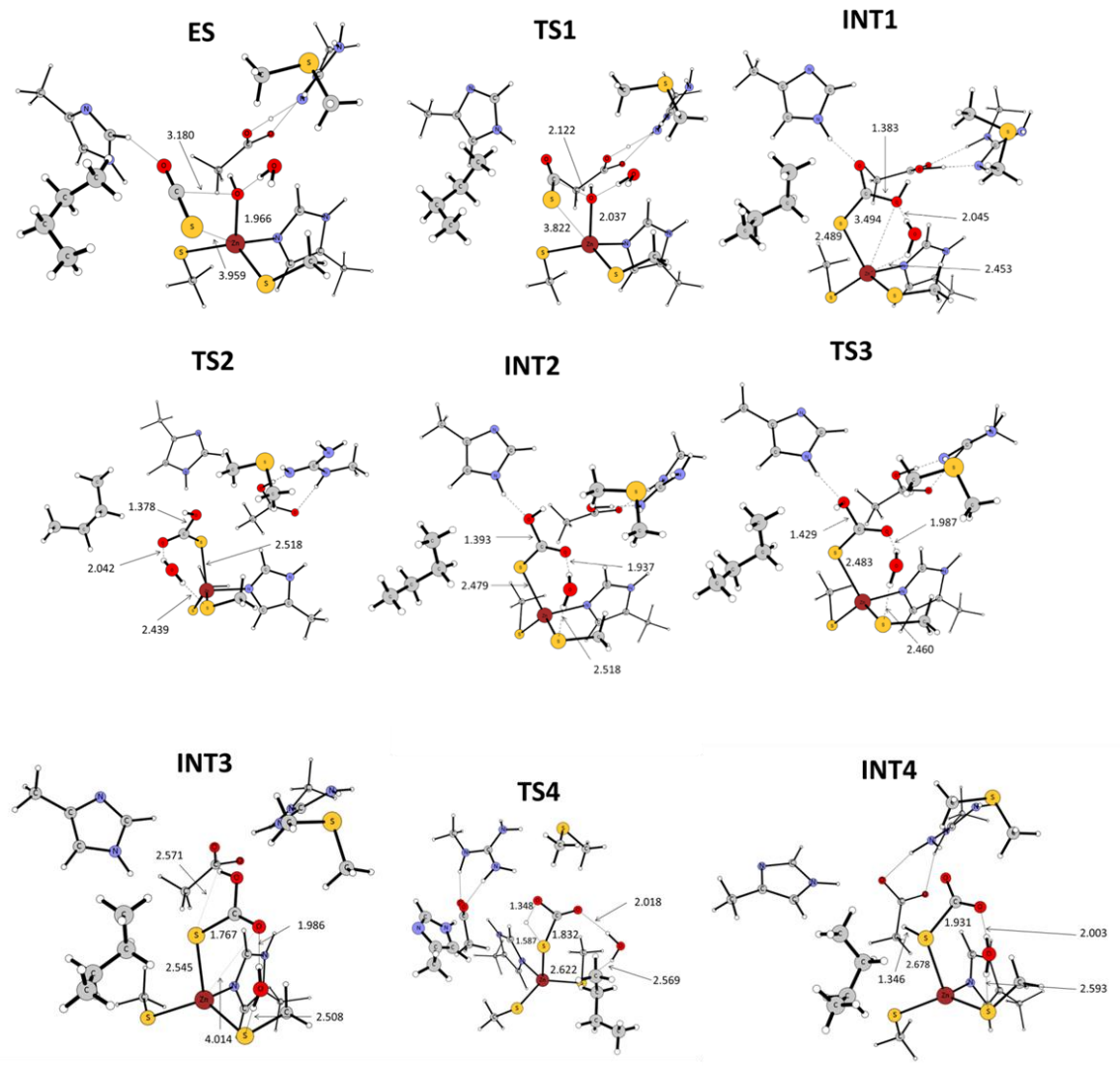


Figure S1. Optimized stationary points present along the PES by considering water molecule in not X-ray position. Distances are in Å.

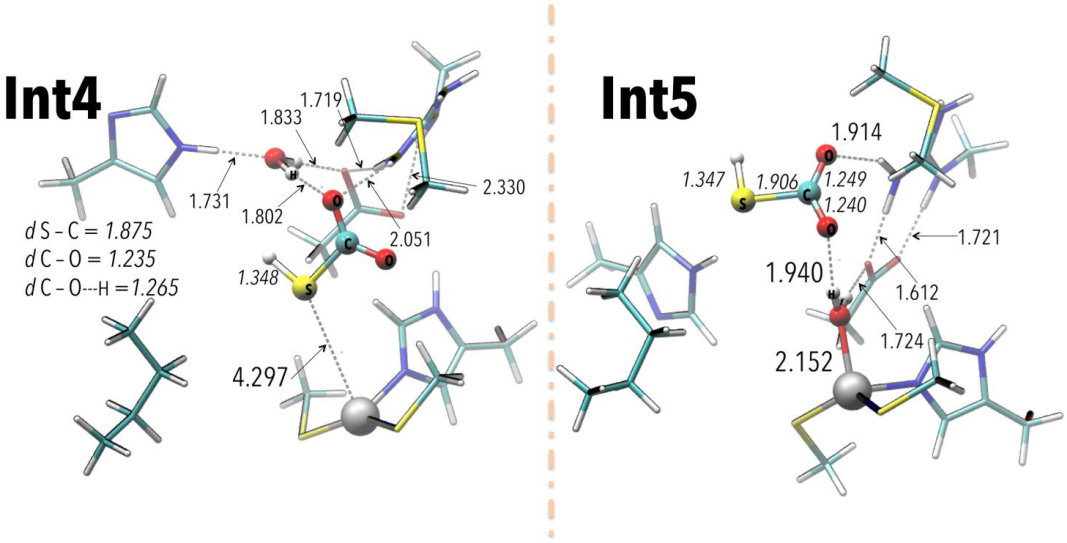


Figure S2. Optimized structures of the intermediates Int4 and Int5 present along the PES for the mechanism with water in X-ray position.

Paper IV

Direct Hydrogenation of Carbon Dioxide by an Artificial Reductase Obtained Substituting Zinc with Rhodium in the Carbonic Anhydrase Catalytic Center. A mechanistic study.

Paolo Piazzetta[1], Tiziana *Marino*[1], Nino Russo[1], D. R. Salahub[2]

[1]Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Rende, Italy

[2]IQST – Inst. for Quantum Science and Technology, CMS – Centre for Molecular Simulation, ISEEE – Inst. for Sustainable Energy, Environment and Economy BI 556
University of Calgary, 2500 University Drive NW
Calgary, Alberta, Canada T2N 1N4

Manuscript in preparation

Direct Hydrogenation of Carbon Dioxide by an Artificial Reductase Obtained Substituting Zinc with Rhodium in the Carbonic Anhydrase Catalytic Center. A mechanistic study

P. Piazzetta[1], *T. Marino*[1]*, *N. Russo*[1], *D. R. Salahub*[2]

[1]Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Rende, Italy

[2]IQST – Inst. for Quantum Science and Technology, CMS – Centre for Molecular Simulation, ISEEE – Inst. for Sustainable Energy,
Environment and Economy BI 556
University of Calgary, 2500 University Drive NW
Calgary, Alberta, Canada T2N 1N4

Abstract

Recently, a new artificial carbonic anhydrase enzyme in which the native zinc cation has been replaced with a Rh(I) has been proposed as a new reductase able to efficiently catalyze the hydrogenation of olefins. In this paper we propose the possible use of this modified enzyme in the direct hydrogenation of carbon dioxide.
In our theoretical investigation we have considered different reaction mechanisms such as reductive elimination and σ-bond metathesis. In addition, the release of the formic acid and the restoring of the catalytic cycle has been also studied. Results show that the σ-bond metathesis potential energy surfaces lies below the reactant species. The rate determining step is the release for the product with an energy barrier of 12.8 kcal/mol. On the basis of our results we can conclude that this artificial enzyme can efficiently catalyze the direct hydrogenation of CO_2.
*Corresponding author: tmarino@unical.it

Keywords: DFT, Rh-carbonic anhydrase, carbon dioxide hydrogenation, artificial enzyme

Introduction

The increasing world energy demand imposes to search alternative source to the largely used fossil fuels. The new chemical feedstock must possess some characteristics that can be summarized in the following points: low cost and abundance, environmentally friendly, easily produced, stored and transportable. Although the use of hydrogen suffers of a series of unsolved problems such as security and storage, there are many evidences that indicate this source as a better candidate for the new "energy era" already called Hydrogen Economy.[1-5] On the other hand, the enormous consumption of fossil fuel introduce into the atmosphere a huge amount of carbon dioxide that strongly contribute to the global climate change[6] that in the last years have caused social and economic undesirable consequence. In this scenario, an ideal combination should be the concomitant use of H_2 and CO_2 for the production of energy and, simultaneously, the reduction of environmental pollution. The nontoxic economic and abundant carbon dioxide is thermodynamically stable and its conversion into useful products is difficult and represents a challenge for the modern chemical research.[7-15] In particular, its transformation in formic acid[16-19] can be considered as a valid hydrogen storage way [20-22] since its possible in site consumption. The reduction of CO_2 in HCOOH can be done by an hydrogenation reaction by using efficient and specific catalysts. There are in the recent literature many examples of the research devoted to this field[23-28] but the proposed catalysts do not give catalytic turnover such to justify their employment at industrial scale. Nature can inspire new research since uses extraordinary chemical machines, the enzymes, that efficiently catalyze a huge number of reactions.

Very recently, a bacterial hydrogen-dependent carbon dioxide reductase from *Acetobacterium woodii* catalyzing the hydrogenation of CO_2, the hydrogen-dependent carbon dioxide reductase (HDCR), has been discovered and characterized.[29] In particular, this kind of enzyme makes physiological use of H_2 + CO_2 as the sole growth substrate and produces formate as intermediate. This discovery opened a new frontier in biocatalysis although the structure is not known and the reaction mechanism seems to involve different subunits, cofactors and electron sources. Nature offers another well studied class of enzymes, Carbonic Anhydrases, widely present in different organisms, that are able to transform, at physiological conditions, the CO_2 in carbonate[30-44] but is not

able to catalyze its direct hydrogenation. In a recent paper Jing, Okrasa and Kazlauskas[45], have demonstrate that the substitution of the natural zinc cation, in the active site of human CA, with a Rh(I) activates the reductase activity of this modified enzyme. In fact, the authors observe the stereoselective direct hydrogenation of olefins with high catalytic efficiency.[45,46] Inspired by these recent and important discoveries, considering that some rhodium organometallic complexes have been previously proposed as promising catalysts for the CO_2 hydrogenation[27,28,47,48] and continuing our previous works on the promiscuous ability of hCA,[49,50] we have undertaken a detailed theoretical work with the aim to verify if the "artificial" Rh(I)-CA is able to efficiently catalyze the direct hydrogenation of carbon dioxide. Due to the novelty of this research, different reaction mechanisms have been explored by using density functional theory and modeling the enzyme active site with a model cluster in which other than the primary ligand of the Rh(I) also other amino acid residues that come into play in the catalytic process have been included.

Computational Methods

All the calculations have been performed at DFT level of theory employing the B3LYP exchange-correlation functional[51,52], as implemented in the Gaussian 09 program.[53] The geometries optimization has been carried out using the 6-31+G(d,p) basis set on all elements except for Rh, which has been described by the relativistic compact Stuttgart/Dresden effective core potential (SDD)[54] pseudopotential associated with its corresponding basis sets.

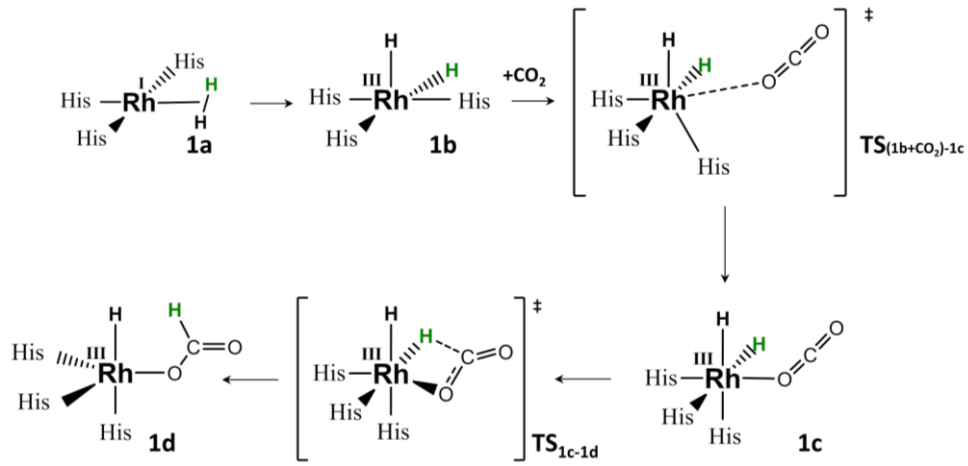
The final and the solvation energies have been calculated as single point corrections on the optimized structures using the PCM[55,56] continuum solvation model employing the larger 6-311+G(2d,2p) basis set and taking into account the dispersion contribution as suggested by Grimme et al.[57,58] The dielectric constant was chosen to be ε=4, which is the most used value used for the protein environmental representation.[59-64]

In order to identify all the stationary points as minima (zero imaginary frequency) or transition states (one imaginary frequency) and to obtain Gibbs free energy corrections at 298.15 K analytic frequency calculations have been performed on all the optimized structures at the same level of theory used for the optimization step. The exact connection of that

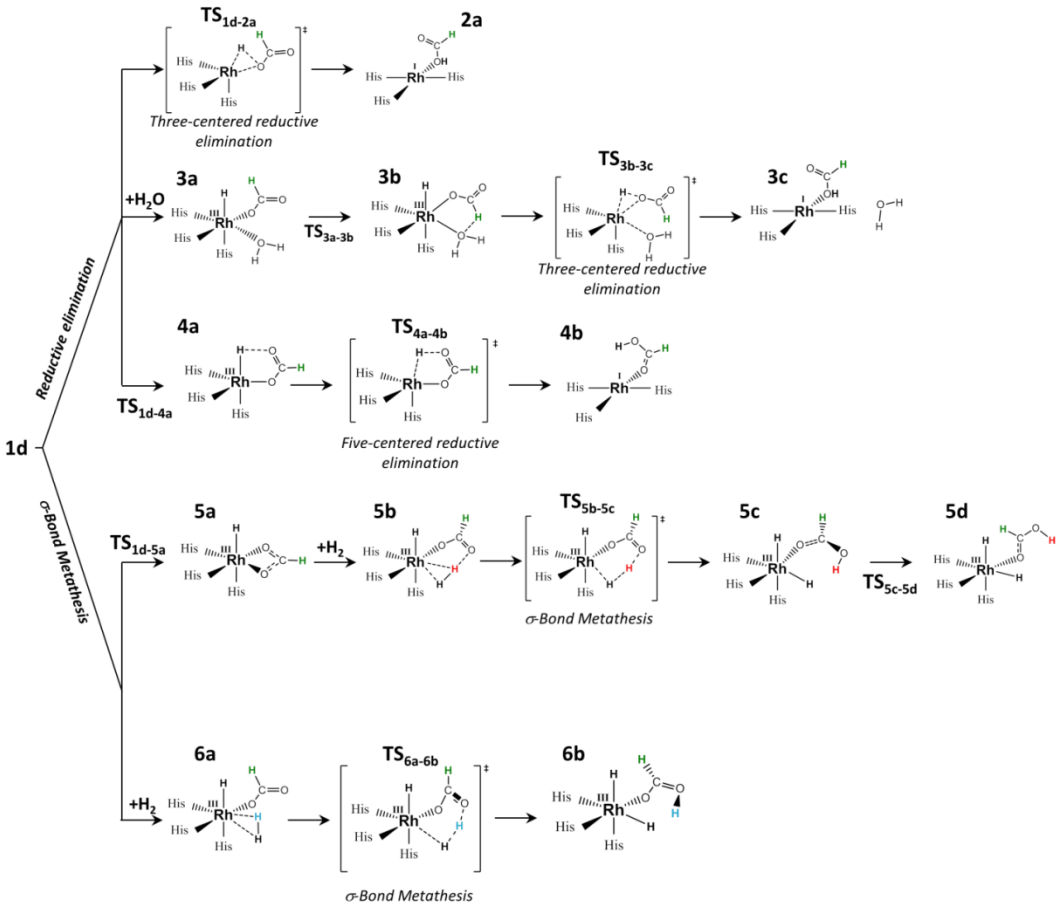
transition state with the two related minima has been checked by using the intrinsic reaction coordinate (IRC) computational procedure.[65,66] NBO charge analysis[67] was carried out on the structures of some intercepted stationary points.

Results and discussion

On the basis of previous studies on the H_2 and CO_2 activation on a series of complexes and considering the catalytic process of the native substrate of the hCAII we have considered different reaction paths as depicted in Schemes 1 and 2. These paths include the hydrogen activation and the subsequent CO_2 insertion (**Path 1**), the formation of formic acid via reductive elimination (**Paths 2, 3 and 4**) and σ -bond metathesis (**Path 5** and **6**) mechanisms. Finally, we have also considered the restoring of the catalytic cycle by analyzing the releasing process of the formed product (Path **4, 5, 6**).



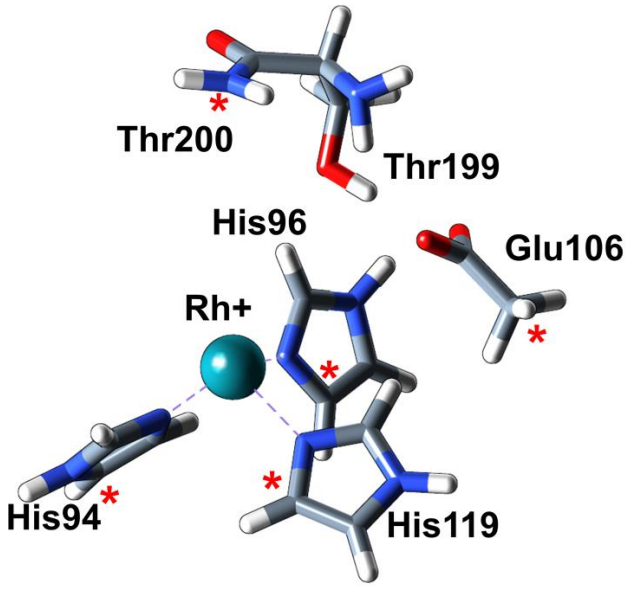
Scheme 1. Proposed pathway for hydrogen activation and CO_2 insertion (path **1**).



Scheme 2. Overview of the investigated formic acid formation pathways starting from **1d** complex. Paths **2**, **3** and **4** are related to reductive elimination and paths **5** and **6** to σ -bond metathesis.

Used model. The model used to describe the reductase behavior of the Rh(I)-containing hCAII is depicted in Scheme 3. It has been obtained by crystallographic structure (PDB=2CBA) where the native zinc ion has been replaced by Rh(I). The obtained cluster mimicking the active site of the enzyme contains a Rh(I) coordinated to the imidazoyl nitrogen atoms modelling the three histidine residues (His94, His96, His119) as primary ligands. The tetracoordination of Rh(I) is then reached adding a water molecule occupying the fourth position around the metal ion. Furthermore, three indirect ligands of the metal center represented by truncated models of the Glu106, Thr199 and Thr200 complete the model (50 atoms, charge =0). The presence of these amino acid residues is believed crucial in the phase of CO_2 approaching and HCOOH releasing due to the formation of an extended H-bond network. The atoms where the truncation has been done, have been kept frozen to their X-ray crystal positions during optimization procedure, to preserve the spatial arrangement of the residues. These centers are indicated by stars in the Scheme 3.

This cluster model revealed to be adequate in the elucidation of the catalytic mechanism followed by hCAII towards native CO_2[68-70] and other isoelectronic promiscuous substrates studied in previous works.[49,50]



Scheme 3. Selected model of the hCAII active site. Stars indicate the atoms that are fixed to their X-ray positions (PDB =2CBA).

Oxidative addition and CO_2 insertion.

The first step of the [Rh(hCAII)] activation concerns the substitution of the coordinated water molecule to metal center Rh(I) with a H_2 molecule and occurs with a destabilization energy of 4.5 kcal/mol (**1a**). The formed species can undergo oxidative addition with the formation of the **1b** product without energy barrier (see Scheme 1). As consequence, the geometry changes from a square planar to a square pyramidal geometry with an apical hydride due to the presence of σ donor ligands (His94, His96, His119) in a d^6 transition metal complex.[71] The barrierless oxidative addition is exothermic and the **1b** product has an energy that lies at 2.1 kcal/mol below **1a**. These findings confirm that Rh is an effective metal in H_2 oxidative addition.[72]

CO_2 insertion into Rh(III) – H bond of [*cis*-Rh**H**$_2$(hCAII)] starts with the formation of an octahedral precursor complex (**1b+CO$_2$**) where the CO_2 is placed at a distance of 4.608 Å from the metallic center. This behavior is different from that found in the case of cis-$[Rh(III)H_2(PH_3)_3]^+$ and cis-$[Rh(III)H_2(PH_3)_2\text{-}(H_2O)]^+$ catalyst models that did present a CO_2 directly coordinated to the rhodium center (2.464 and 2.320 Å, respectively).[47] In the enzyme the presence of the Glu106-Thr199-Thr200 triad, retained in our model, does allow the formation of a stable complex (**1b+CO$_2$**). The formation of the coordinate intermediate (**1c)** requires an activation barrier of 6.9 kcal/mol. In the transition state (**1b+CO$_2$-1c**) CO_2 establishes by O atom a weak interaction with the NH_2 moiety of Thr199 and is located at 3.073 Å from the rhodium center. In order to accommodate the incoming CO_2, the His119 ligand is involved in a gradual migration from the equatorial position to the axial one with an increment in Rh – His119 distance of 0.233 Å due to trans effect of H_α.[73]
The structure of this transition state evidences as the CO_2 acts as a Lewis base coordinating through one oxygen atom, as normally occurs in electron – deficient metal centers.[74] The **1c** formed complex results to be more stable than **1b+CO$_2$** by 2.0 kcal/mol. From **1c** CO_2 is inserted in Rh(III) – H bond by a four centered TS (**TS$_{1c\text{ - }1d}$**) giving rice to the **1d** species in which the OCOH fragment is coordinated to the rhodium center in η_1 fashion. The imaginary frequency in **TS$_{1c\text{ - }1d}$** (60.1i cm^{-1}) mainly involves Rh-H_β bond breaking and C - H_β bond formation. The geometrical features of this transition state indicate the η_1-formate anion is almost formed. The **1d** complex is characterized by a five coordinated

pseudo square pyramidal geometry (see **Figure 2**) and lies at only 0.1 kcal/mol below the reactant energies.

The NBO population of C atom increases as the insertion proceeds, while that of the Rh one decreases. This behavior (see **Figure 1**) clearly indicates that a charge transfer from Rh to CO_2 occurs during the η_1-formate formation. The net charges at the TS are almost the same as that of the product, confirming that the formate anion is almost formed. H_α and H_β population remains substantially unchanged during the reaction, allowing to hypothesize that both the charge transfers from H_β to C and from Rh to H_β participate in the CO_2 activation.

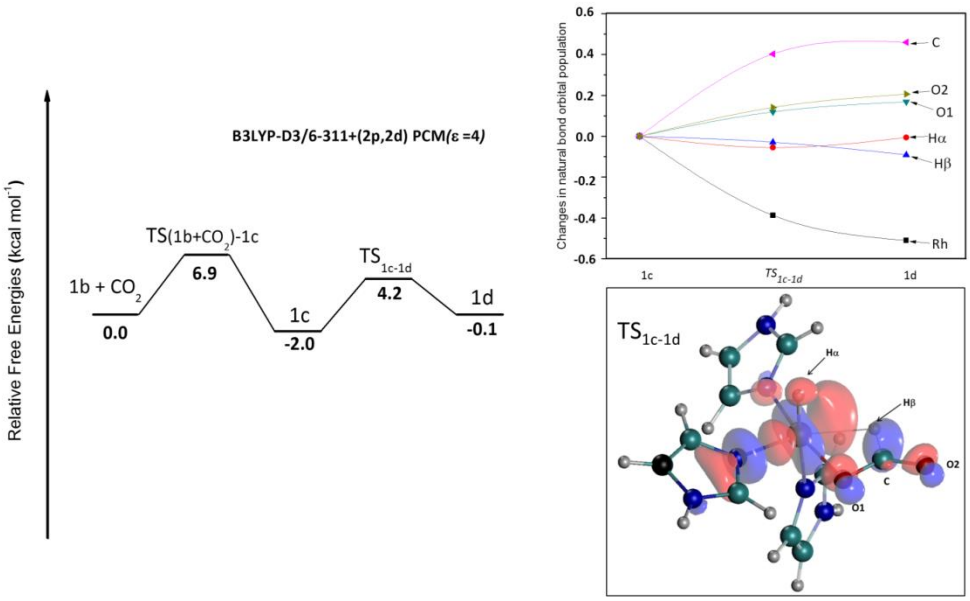


Figure 1. Calculated B3LYP-D3 free-energy profile for the hydrogen activation and CO_2 insertion (path **1**). Energies ($kcalmol^{-1}$) are relative to the asymptote of the reactants. Population changes occurring in path **1**. A positive value represents an increase in population relative to **1c** species. HOMO contour plot in the ***TS*$_{1c\text{-}1d}$** species.

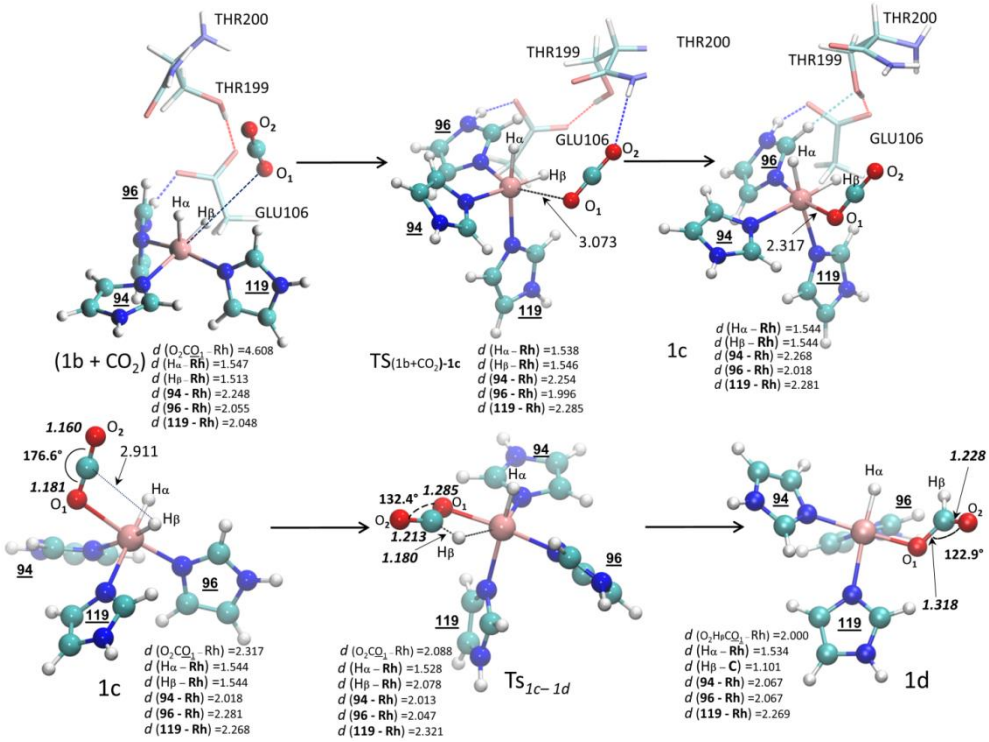


Figure 2. DFT optimized geometrical structures for all the species intercepted through the hydrogen activation and CO_2 insertion (path **1**)with selected structural parameters (distances in Å). In **1c**, $\mathbf{TS_{1c\ -\ 1d}}$ and **1d** species the Glu106-Thr200-Thr199 triad is omitted for clarity.

Reductive elimination (Path 2,3,4).

As depicted in Scheme 2, the reductive elimination of the η_1-formic acid can occur following two different paths characterized by the formation of three centered (paths **2**, **3**) and five centered (path **4**) transition states.
The computed potential energy surfaces for the paths **2** and **4** are reported in **Figure 3**.

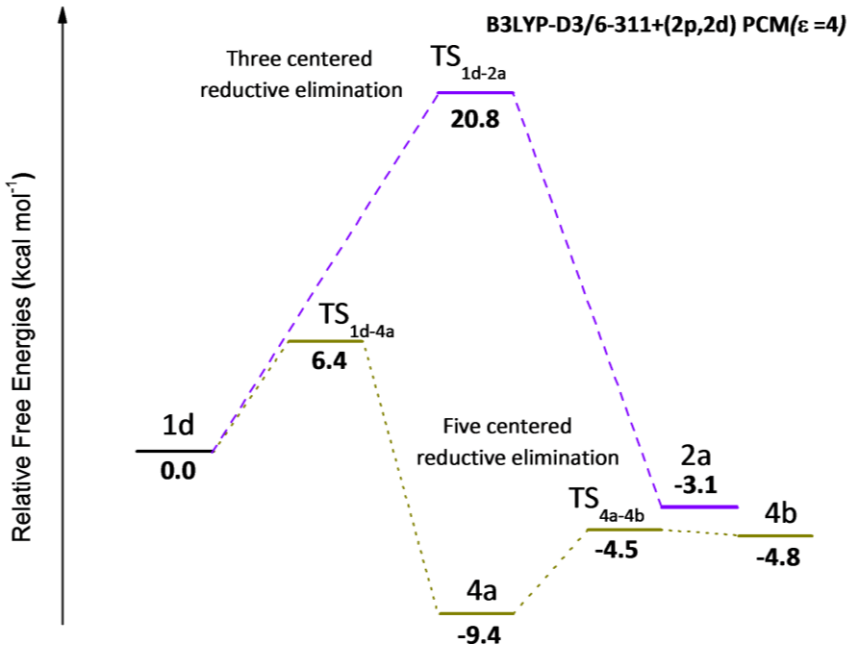


Figure 3. Calculated B3LYP-D3 free-energy profiles for the three centered (path **2**) and five centered (path **4**) reductive elimination. Energies (kcalmol-1) are relative to the asymptote of the reactant (**1d**).

Starting from **1d**, in the path 2, the proton transfer of apical H_α atom moves toward the O_1 atom (without a filling of the geometrical vacancy of the pseudo square pyramidal structure) giving rise to the three centered **TS**$_{1d\text{-}2a}$ characterized by an imaginary frequency (1030.8i cm^{-1}) whose eigenvector involves the H_α shifting to O_1. This TS lies at 20.8 kcal/mol, and the final product (**2a**) in which the formic acid is coordinated to the Rh cation throughout the OH group, is exothermic by 3.1 kcal/mol with respect to the **1d** starting species. The **2a** product takes a four coordinated planar structure (**Figure 4**) due to the d^8 electron configuration of Rh(I) and the Rh-O_1 distance (2.255 Å) is longer than the corresponding in **1d** formate adduct. Following this path, the formation of formic acid requires an activation energy higher enough (20.8 kcal/mol) to hypothesize an efficient catalytic process. Similar behavior (with an activation energy of 19.6 kcal/mol) has been previously found in the case of CO_2 reductive elimination catalyzed by a Wilkinson-like complex.[47]

Due to the presence of this higher energy barrier, we have considered the possibility that in this process a water molecule, that could fill the vacancy in the **1d** structure, can assist the formic acid formation (path **3** in **Figure 5**). In this situation, the first step of the reaction is the formation without energy barrier of a mono hydrate octahedral complex **3a (**see Scheme 2 and **Figure 4)** with a stabilization energy of 19.5 kcal/mol with respect to **1d**. In this complex, the water is coordinated to the Rh with a distance of 2.046 Å and rhodium coordination geometry is now octahedral. The subsequent isomerization (**3a–->3b**) occurs with a rotation along O_1-C bond and the relative transition state (**TS**$_{3a-3b}$) lies at only 0.7 kcal/mol above the reactants. The resulted intermediate (**3b)** is very stable (-30.8 kcal/mol) and is characterized by an hydrogen bond between the formate oxygen and one hydrogen of the water molecule (1.510 Å). From this point an energetically prevented three centered TS (**TS**$_{3b-3c}$, Ea = 42.2 kcal/mol) involves the H_α shifting toward O_1, with the concomitant expulsion of the water molecule from the rhodium coordination sphere (**3c)**. From these results it is evident that, also in the presence of water molecule the three centered transition state reductive elimination requires high energies.

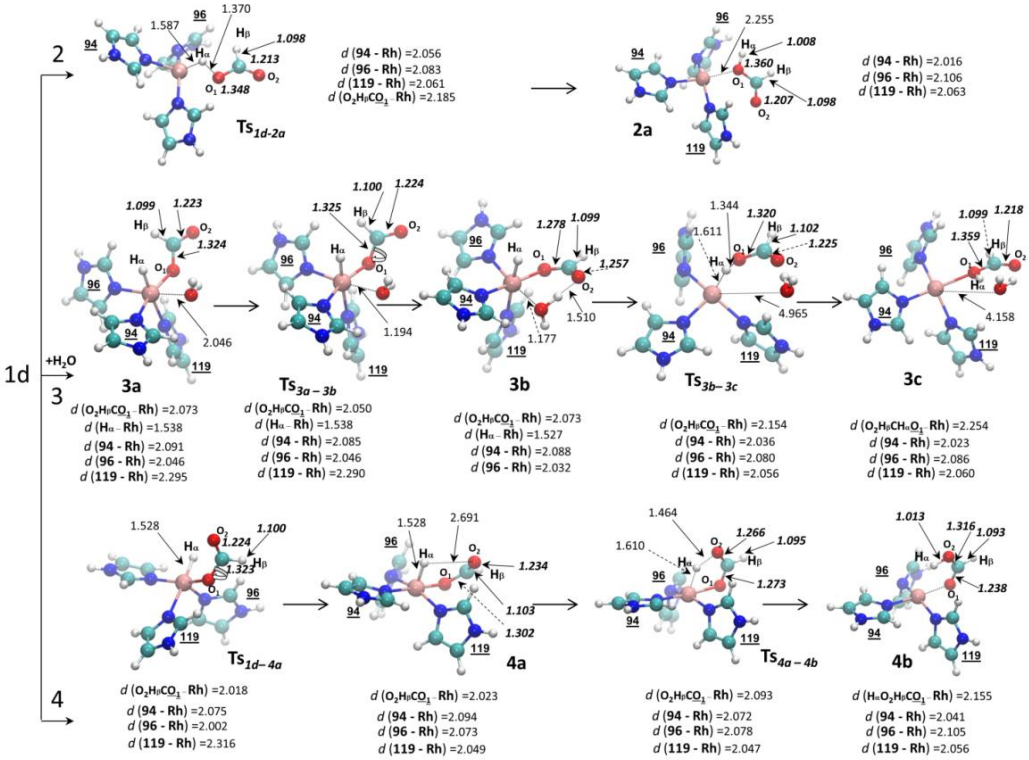


Figure 4. DFT optimized geometrical structures for all the species intercepted through paths **2**, **3** and **4** with selected structural parameters (distances in Å). The Glu106-Thr200-Thr199 triad is omitted for clarity.

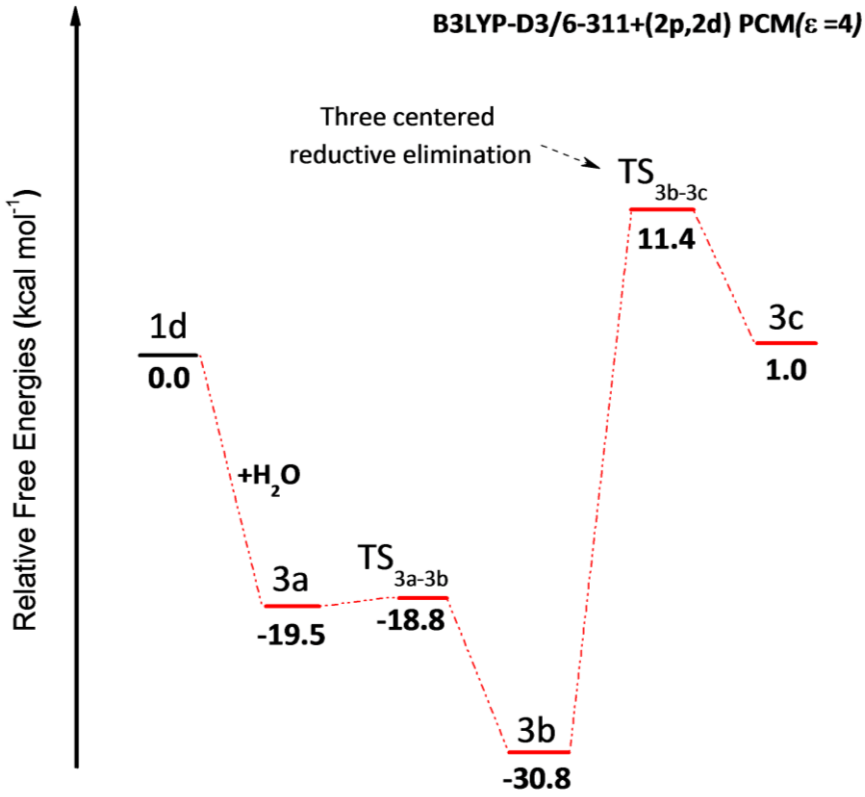


Figure 5. Calculated B3LYP-D3 free-energy profile for the three centered (path **3**) reductive elimination water assisted. Energies (kcalmol-1) are relative to the asymptote of the reactant (**1d**).

As shown in Scheme 2, an alternative way is the five centered reductive elimination. The potential energy profile for this process is reported in **Figure 3** while the geometries of the species characterizing this path are depicted in **Figure 4**. The initial step is the isomerization of **1d** to **4a** that occurs with a rotation around the O_2 – C bond. This topological change is necessary to prearrange the formate in a fashion suitable to form a five centered interaction with the Rh-H moiety. The located $\mathbf{TS}_{1d-4a}$ (176.1i cm^{-1}) has a rearrangement of geometry, that remains a square planar pyramidal, involving the His119 that moves on the plane formed by His94, His96 and η_1-formate. The related energy is calculated to be 6.4 kcal/mol above **1d**. The five centered reductive elimination occurs through $\mathbf{TS}_{4a-4b}$ (526.9i cm^{-1}) with an activation energy of only 4.9 kcal/mol. The structure of this transition state indicates that the Rh – H_α bond is practically broken (1.610 Å) while the value of the O_2 – H_α distance (1.464 Å) suggests the initial bond formation (see **Figure 4**).
The analysis of the NBO behaviors for the considered reductive eliminations paths are reported in **Figure 6**. As shown in this Figure, H_α atomic population decreases in both reductive eliminations because H_α mutates into proton. The differences in term of activation energy between the two mechanisms can be related to the reason that in the three centered reductive elimination paths (**2** and **3**), the p orbital of O_1 atom must change its direction toward H_α suppressing the charge transfer to the Rh and lightly increasing the O_1 atomic population. On the other hand, the O_2 p orbital in the five centered reductive elimination, is well oriented and expands toward the H_α while O_1 population decreases with the elongation of the Rh - O_1 bond (see **Figure 6**).
We can conclude that the reductive elimination path of **1d** that occurs through the formation of a five centered transition state ($\mathbf{TS}_{4a-4b}$), is the favored one.

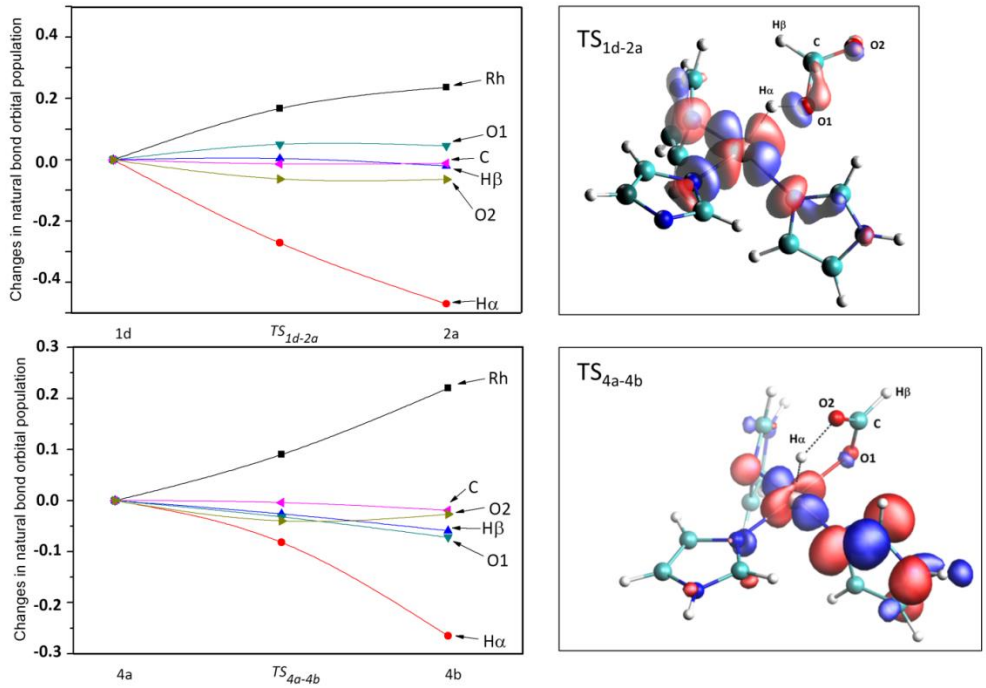


Figure 6. Population changes occurring in paths **2** and **4.** A positive value represents an increase in population relative to **1d** (path **2**) and **4a** (path **4**) species**.** HOMO contour plots are related to the ***TS*$_{1d-2a}$** and ***TS*$_{4a-4b}$** species.

σ-bond metathesis (**Paths 5, 6)**

Other than the reductive elimination we have considered also the possibility that the enzymatic production of formic acid can follow other paths in which a σ-bond metathesis should be the main mechanism. In this hypothesis, two distinct paths have been considered (see Scheme 2). In the first path (path **5**) the formate that is mono-coordinated to the Rh cation (**1d**) undergoes a rotation forming a η_2-formate product coordinated to Rh in a bidentate way (**5a**). This step can be understood since the Rh(III) ion tends to form a six coordinated complex due to its d^6 electron configuration. The transition state connecting these two minima (**TS**$_{1d\ -\ 5a}$) requires 3.7 kcal/mol respect to the asymptotic limit (see **Figure 7**) and occurs through a rotation of –OCOH moiety. The value of the imaginary frequency (159.2i cm^{-1}) well accounts for this movement. The O – Rh distances are now 2.128 and 2.136 Å evidencing the bidentate nature of the Rh(III) formate interaction (see **Figure 8**). The product **5a** is -12.3 kcal/mol more stable than **1d** and the system assumes an octahedron-like geometry in which the O_1 – Rh – NHis96 angle is 102.2° and the His96 is in a staggered plane with respect to the plane formed by the other equatorial ligands. This arrangement allows the insertion of a new H_2 molecule in the same side of the O_1 atom of [RhH(η_2 – O_2CHO_1 (hCAII)]. The **5b** formed species is 3.8 kcal/mol more stable than **5a** and the insertion of the molecular hydrogen breaks the formate bi-coordination (see **Figure 8**). From this intermediate the σ-bond metathesis takes place by means a six centered TS (**TS**$_{5b\ -\ 5c}$) with a negligible activation energy (0.9 kcal/mol). The analysis of the related imaginary frequency (473.6i cm^{-1}) clearly shows that the H – Rh bond is in formation and the hydrogen transfer to the oxygen is in progress (H1 – O_1 distance is 1.453 Å). We underline that the PES region describing σ-bond metathesis process appears flat enough such that the reaction proceeds without barrier for the formation of formic acid containing complex (**5c)** that is stabilized by 20.0 kcal/mol. No significant geometrical changes occur during the process apart from the elongation of Rh – His94 distance due to trans effect of H_1. As shown in **Figure 9**, H_1 atomic population decreases in the σ-bond metathesis because H_1 takes a favorable position to form interaction with the O_1 p orbital of HOMO. The lone pair on the oxygen atom is oriented toward the incoming hydrogen and acts as trigger in the H_1-H_2 heterolytic cleavage. Moreover, η_1-formic acid product (**5c**) can isomerize through a rotation of the Rh – O_2 – C – H_β

dihedral leading to the formation of a less stable isomer of 9.4 kcal/mol (**5d**). This isomerization occurs through a transition state (**TS**$_{5c-5d}$) with an activation energy of 12.4 kcal/mol.

The other explored path includes a pseudo σ-bond metathesis (path **6** of **Figure 7**). In this alternative way the molecular hydrogen is directly linked into to the **1d** intermediate filling the vacancy in this species with the formation of an octahedral complex (**6a**) lying at 12.2 kcal/mol below **1d**. The reaction directly proceeds with the formation of a complex, in which the formic acid is linked to the Rh(III) (**6b**), with an activation energy of only 5 kcal/mol. The process requires the rotation of the η_1–formate with the concomitant formation of a concerted σ-bond metathesis **TS**$_{6a-6b}$. (see ***Figure 8***). The imaginary frequency of 171.2i cm^{-1} is related to rotation modes and both a Rh – O_1 and C – O_2 stretches. The topology of this TS is different from the corresponding transition state in the path **5** and results to be a pseudo six – membered ring. The low energy barrier associated with **TS**$_{6a-6b}$ can be elucidated by analyzing the MO representation and the NBO population (see **Figure 9**). During the formate rotation a lone pair of O_2 atom orients itself to establish a positive overlap with one H' hydrogen atom. Compared to the path **5** both Rh and O_1 population slightly decreases since the formate moiety rotation. We can conclude that during this dualistic process a new hybridization occurs and allows a best overlap between H' and O_2 with a formation of a covalent interaction.

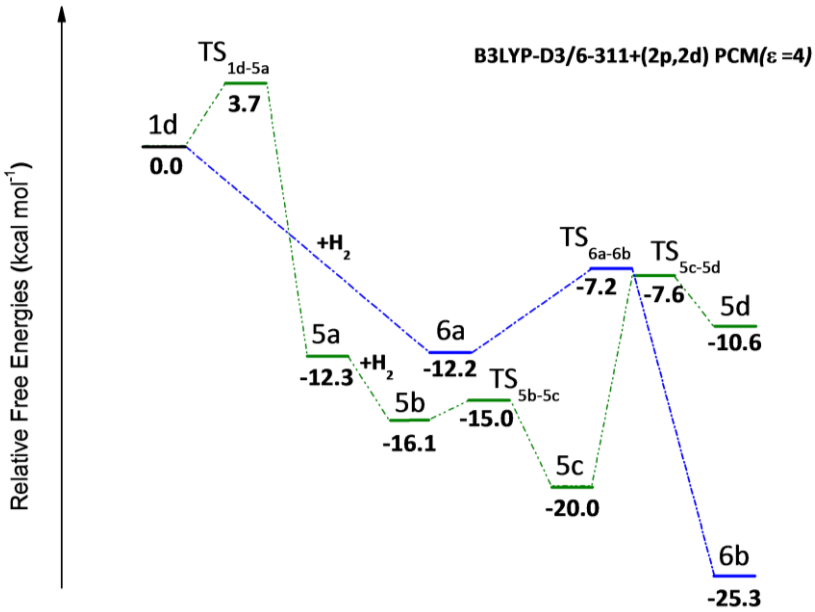


Figure 7. Calculated B3LYP-D3 free-energy profiles for the σ-bond metathesis explored pathways (**5** and **6**). Energies (kcalmol-1) are relative to the asymptote of the reactant (**1d**).

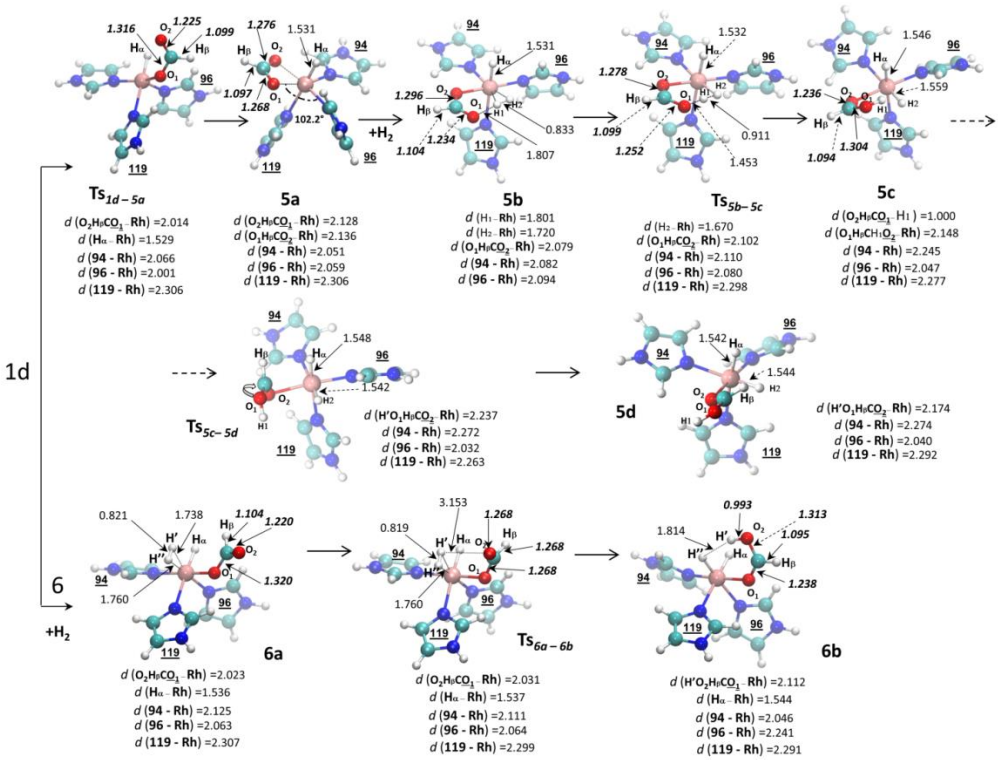


Figure 8. DFT optimized geometrical structures for all the species intercepted through paths **5** and **6** with selected structural parameters (distances in Å). The Glu106-Thr200-Thr199 triad is omitted for clarity.

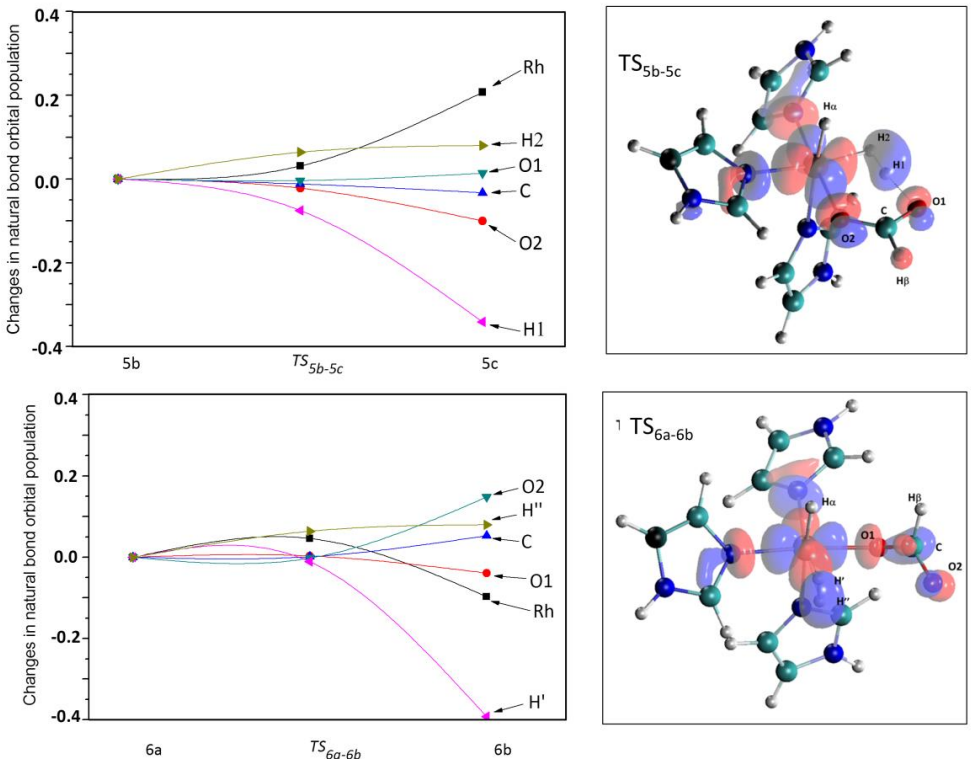


Figure 9. Population changes occurring in path **5** and **6**. A positive value represents an increase in population relative to **5b** (path **5**) and **6a** (path **6**) species. HOMO contour plots are related to the ***TS*$_{5b-5c}$** and ***TS*$_{6a-6b}$** species**.**

Release of the formic acid and restoring of the catalytic cycle.

In both enzymatic and non-enzymatic reactions, the restoring of the catalytic cycle can be the rate determining step.[75-77] In our case, the release of the formic acid product should be a crucial point in restoring the catalyst. Kinetic and theoretical investigations demonstrate how the presence of a base can facilitate the liberation of formic acid.[79] This aspect has been confirmed in the case of Ru complexes in which a Lewis base is necessary to release formic acid.[78]

For this reason, we have explored also this process considering the addition of a hydrogen or a carbon dioxide molecule necessary to restore the catalytic cycle for removing the formic acid produced at the end of paths **4, 5** and **6** (see **Figure 10**). The final product of reductive elimination (path **4**) is the species **4b** so, an entering H_2 (H3 – H4) can give rise to the formation of **1a** and the detachment from the Rh center of the formic acid. As shown in **Figure 10** the potential energy surface for this process includes the formation of a new intermediate **4c** in which H_2 weakly interacts with Rh cation (see **Figure 11**) and its energy lies at only 2.1 kcal/mol above **4b**. The subsequent insertion of molecular hydrogen in the Rh coordination sphere requires the bypassing of a barrier of 10 kcal/mol and occurs through the transition state (**TS**$_{4c\,-\,4d}$). The imaginary frequency (307i cm^{-1}) clearly indicates the H_2 insertion step. In the final product **4d**, that is stabilized by only 1.1 kcal/mol with respect to **1d**, the H-Rh distances are 1.662 and 1.670 Å, the H3-H4 distance is 0.881 Å and the formic acid is detached from the Rh center (4.492 Å). As shown in **Figure 11** the Thr200 residue seems to help the release of the formic acid.

The restoring of the catalyst starting from the products obtained from σ-bond metathesis mechanisms (paths **5** and **6**) require the interaction with a new carbon dioxide molecule. According to what happens in the CO_2 insertion, where CO_2 acts as a Lewis base coordinating through one oxygen atom the metal center, we investigated the release of the formic acid from both **5d** and **6b** complexes. In the first adduct, for both the paths, the CO_2 is linked to the complex with weak interaction with consequent slight energy stabilization. Both the starting complexes have octahedral coordination and the insertion is more difficult. In fact the located transition states, ***TS***$_{5e\,-\,5f}$ and ***TS***$_{6c\,-\,6d}$, impose barriers of 20.5 and 12.8 kcal/mol respectively (see **Figure 10**). The lowest energy barrier for the ***TS***$_{6c\,-\,6d}$ is also due to the presence of an H-bond between OH

moiety of formic acid and the corresponding group of Thr199 (1.564 Å, see **Figure 11**). This hydrogen bond is retained in the **6d** species (1.586 Å) that results again more stable than **5f**. In both final complexes the carbon dioxide results to be coordinated to the Rh center (Rh-O distances of 2.285 Å and 2.225 Å for **5f** and **6d**, respectively) while the formic acid is clearly released (Rh-O distances of 4.049 and 5.130 for **5f** and **6d** respectively).

From the data reported in **Figure 10** we underline that both paths **4** and **5** show transition state energies that lies above the asymptotic limit (**1d** energy) while the whole species intercepted along the path **6** lie below **1d** energy. This means that we can propose path **6** as the preferred one for the restoring of the catalytic cycle.

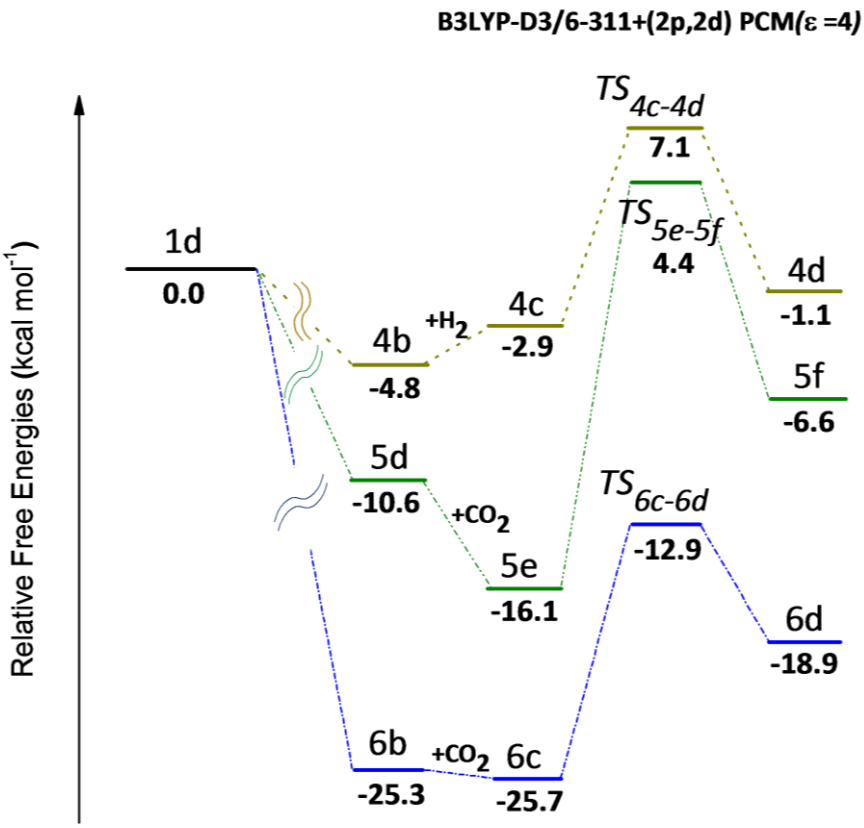


Figure 10. Calculated B3LYP-D3 free-energy profiles for formic acid release explored pathways. Energies (kcalmol-1) are relative to the asymptote of the reactant (**1d**).

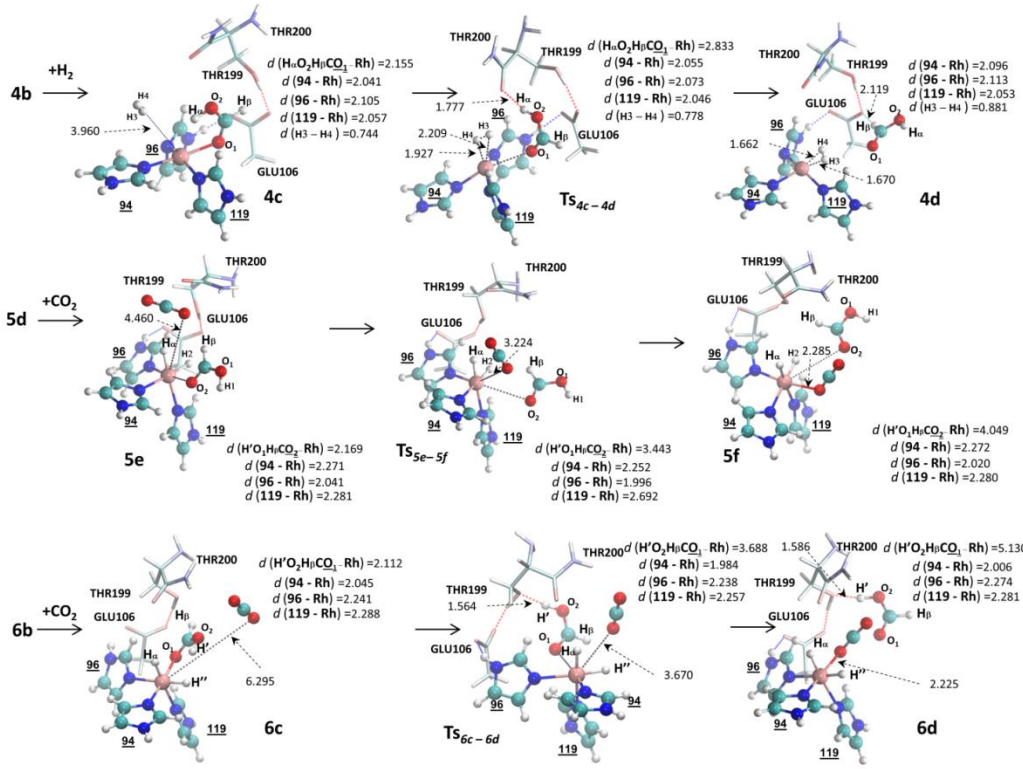


Figure 11 DFT optimized geometrical structures for all the species intercepted through the explored paths for formic acid release starting from **4d**, **5b** and **6b** species with selected structural parameters (distances in Å). The Glu106-Thr200-Thr199 triad is retained.

Conclusions

Starting from the recent experimental evidence that a modified carbonic anhydrase in which the native zinc cation is substituted with a Rh(I) center, is able to act as a reductase enzyme able to efficiently catalyze the hydrogenation of olefins, we have explored at DFT level of theory, the possibility that the carbon dioxide native substrate of carbonic anhydrase can be also reduced by a direct hydrogenation process. We have considered the following possible reaction mechanisms: (i) reductive elimination (two different paths in which a three- or five-centered transition state is possible); (ii) σ-bond metathesis and pseudo σ-bond metathesis; (iii) release of the formic acid product and the restoring of the catalytic cycle. In addition, for the reductive elimination the possible mechanism that involves the assistance of a deep water molecule has been also explored. Results show that:

-In the first step of the reaction, the insertion and the activation of H_2 occur without energy barrier;

-The subsequent insertion of carbon dioxide is possible with a relative low energy request (6.9 kcal/mol);

-The favored reductive elimination path is that in which a five-centered transition state is formed. The required activation energy for this last process results to be 6.4 kcal/mol. The water assistance can be excluded due to the found high energetic barrier (11.4 kcal/mol);

- The σ-bond metathesis reaction mechanism is the favored path and all the minima and maxima intercepted on the relative potential energy surface lie below the considered asymptotic limit;

- The rate determined step is the restoring catalytic cycle in which the formic acid product is released and the catalyst regenerated.

Our result indicate that this new modified Rh-carbonic anhydrase can act as reductase also for the direct hydrogenation of carbon dioxide. We hope that our work can stimulate experimental studies that confirm our prevision.

Acknowledgements

This work was financially supported through Department of Chemistry and Chemical Technologies of the Università della Calabria (Italy). P. Piazzetta gratefully acknowledges Commissione Europea, Fondo Sociale Europeo, and Regione Calabria for the financial support.

References

1. Suh, M. P.; Park, H. J.; Prasad, T. K.; Lim, D.-W. *Chem. Rev.* **2012**, *112*, 782.
2. Coontz, R.; Hanson, B. *Science* **2004**, *305*, 957.
3. Crabtree, G. W.; Dresselhaus, M. S.; Buchanan, M. V. *Phys. Today* **2004**, *57*, 39.
4. U. S. DOE. Hydrogen, Fuel Cells & Infrastructure Technology Program (http://www.eere.energy.gov/hydrogenandfuelcells/storage).
5. The American Physical Society. The Hydrogen Initiative (http://www.aps.org/public_affairs/index.cfm).
6. Appel, A.; M.; Bercaw, J. E.; Bocarsly, A. B.; Dobbek, H.; DuBois, D. L.; Dupuis, M.; Ferry, J. G.; Fujita, E.; Hille, R.; Kenis, P. J. A.; Kerfeld, C. A.; Morris, R. H.; Peden, C. H. F.; Portis, A. R.; Ragsdale, S. W.; Rauchfuss, T. B.; Reek, J. N. H.; Seefeldt, L. C.; Thauer, R. K.; Waldrop, G. L. *Chem. Rev.* **2013**, *113*, 6621.
7. Reda, T.; Plugge, C. M.; Abram, N. J.; Hirst J. *PNAS* **2008**, 105, 10654.
8. Aresta, M.; Dibenedetto, A. *Dalton Trans.* **2007**, 2975.
9. Benson, E. E.; Kubiak, C. P.; Sathrum, A. J.; Smieja, J. M. *Chem. Soc. Rev.* **2009**, *38*, 89.
10. Finn, C.; Schnittger, S.; Yellowlees, L. J.; Love, J. B. *Chem. Commun.* **2012**, *48*, 1392.
11. Peters, M.; Köhler, B.; Kuckshinrichs, W.; Leitner, W.; Markewitz, P.; Müller, T. E. *ChemSusChem* **2011**, *4*, 1216.
12. Lewis, N. S.; Nocera, D. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15729.
13. Activation of Small Molecules: Organometallic and Bioinorganic Perspectives; Tolman, W. B., Ed.; *Wiley-VCH*: Weinheim, Germany, **2006**.
14. Riduan, S. N.; Zhang, Y. *Dalton Trans.* **2010**, *39*, 3347.
15. Sakakura, T.; Choi, J.; Yasuda, H. *Chem. Rev.* **2007**, *107*, 2365.
16. Jessop, P. G.; Ikariya, T.; Noyori, R. *Chem. Rev.***1995**, *95*, 259.
17. Jessop, P. G.; Joo, F.; Tai, C. C. *Coord. Chem. Rev.* **2004**, *248*, 2425.
18. Leitner, W. *Angew. Chem., Int. Ed. Engl.***1995**, *34*, 2207.
19. Wang, W.; Wang, S.; Ma, X.; Gong, *J. Chem. Soc. Rev.* **2011**, *40*, 3703.
20. Laurenczy, G. *Chimia* **2011**, *65*, 663.
21. Boddien, A.; Mellmann, D.; Gaertner, F.; Jackstell, R.; Junge, H.; Dyson, P. J.; Laurenczy, G.; Ludwig, R.; Beller, M. *Science* **2011**, *333*,1733.
22. Johnson, T. C.; Morris, D. J.; Wills, M. *Chem. Soc. Rev.* **2010**, 39, *81*.
23. Federsel, C.; Boddien, A.; Jackstell, R.; Jennerjahn, R.; Dyson, P. J.; Scopelliti, R.; Laurenczy, G.; Beller, M. *Angew. Chem. Int. Ed.* **2010**, *49*, 9777.

24. Langer, R.; Diskin-Posner, Y.; Leitus, G.; Shimon, L. J. W.; BenDavid, Y.; Milstein, D. *Angew. Chem., Int. Ed.* **2011**, *50*, 9948.
25. Jeletic, M. S.; Mock, M. T.; Appel, A. M.; Linehan, J. C. *J. Am. Chem. Soc.* **2013**, *135*, 11533.
26. Tai, C. C.; Chang, T.; Roller, B.; Jessop, P. G. *Inorg. Chem.* **2003**, *42*, 7340.
27. Annibale V. T.; Song D. *Organometallics* **2014**, *33*(11), 2776.
28. Hou, C.; Jiang, J.; Zhang, S.; Wang, G.; Zhang, Z.; Ke, Z.; Zhao, C. *ACS Catalysis* **2014**, *4* (9), 2990.
29. Schuchmann K.; Müller, V. *Science* **2013**, *342*, 1382.
30. Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides G. M. *Chem. Rev.* **2008**, *108* (3), 946.
31. Khalifah, R. G. *J. Biol. Chem.* **1971**, *246*, 2561.
32. Jewell, D. A.; Tu, C.; Paranawithana, S. R.; Tanhauser, S. M.; LoGrasso, P. V.; Laipis, P. J.; Silverman, D. N. *Biochemistry* **1991**, *30*, 1484.
33. Ren, X.; Jonsson, B. H.; Millqvist, E.; Lindskog, S. *Biochim. Biophys. Acta* **1988**, *953*, 79.
34. Baird, T. T., Jr.; Waheed, A.; Okuyama, T.; Sly, W. S.; Fierke, C. A. *Biochemistry* **1997**, *36*, 2669.
35. Hurt, J. D.; Tu, C.; Laipis, P. J.; Silverman, D. N. *J. Biol. Chem.* **1997**, *272*, 13512.
36. Heck, R. W.; Tanhauser, S. M.; Manda, R.; Tu, C.; Laipis, P. J.; Silverman, D. N. *J. Biol. Chem.* **1994**, *269*, 24742.
37. Feldstein, J. B.; Silverman, D. N. *J. Biol. Chem.* **1984**, *259*, 5447.
38. Earnhardt, J. N.; Qian, M.; Tu, C.; Lakkis, M. M.; Bergenhem, N. C. H.; Laipis, P. J.; Tashian, R. E.; Silverman, D. N. *Biochemistry* **1998**, *37*, 10837.
39. Lindskog, S.; Silverman, D. N. In The Carbonic Anhydrases: New Horizons; Chegwidden, W. R., Carter, N. D., Edwards, Y. H., Eds.; Birkha¨user Verlag: Basel, Switzerland, **2000**; Vol. 90.
40. Chegwidden, W. R.; Carter, N. D. In The Carbonic Anhydrases: New Horizons; Chegwidden, W. R., Carter, N. D., Edwards, Y. H., Eds.; Birkha¨user Verlag: Basel, Switzerland, **2000**; Vol. 90.
41. Lehtonen, J.; Shen, B.; Vihinen, M.; Casini, A.; Scozzafava, A.; Supuran, C. T.; Parkkila, A.-K.; Saarnio, J.; Kivela¨, A. J.; Waheed, A.; Sly, W. S.; Parkkila, S. *J. Biol. Chem.* **2004**, *279*, 2719.
42. Nishimori, I.; Vullo, D.; Innocenti, A.; Scozzafava, A.; Mastrolorenzo, A.; Supuran, C. T. *J. Med. Chem.* **2005**, *48*, 7860.
43. Nishimori, I.; Minakuchi, T.; Onishi, S.; Vullo, D.; Scozzafava, A.; Supuran, C. T. *J. Med. Chem.* **2007**, *50*, 381.
44. Nishimori, I.; Vullo, D.; Innocenti, A.; Scozzafava, A.; Mastrolorenzo, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3828.

45. Jing, Q.; Okrasa, K.; Kazlauskas, R. J. *Chem. Eur. J.* **2009**, *15*, 1370.
46. Jing, Q.; Kazlauskas, R. J. *ChemCatChem* **2010**, *2*, 953.
47. Musashi, Y.; Sakaki, S. *J. Am. Chem. Soc.* **2002**, *124*, 7588.
48. Aresta, M.; Dibenedetto, A.; Angelini A. *Chem. Rev.* **2014**, *114* (3), 1709.
49. P. Piazzetta, T. Marino, N. Russo *Inorg. Chem.* **2014**, *53*, 3488.
50. P. Piazzetta, T. Marino, N. Russo *Phys. Chem. Chem. Phys.* **2014**, *16*, 16671.
51. Becke, A. D. J. *J. Chem. Phys.* **1993**, *98*, 5648.
52. Lee, C. T. Yang, W. T. Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
53. Gaussian 09, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, M. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2009**.
54. Andrae, D.; Haussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123.
55. Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995.
56. Cossi, M.; Rega, N.; Scalmani, G. *J. Comput. Chem.* **2003**, *24*, 669.
57. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg H. *J. Chem. Phys.* **2010**, *132*, 154104.
58. Grimme, S.; Ehrlich, S.; Goerigk L. *J. Comput. Chem.* **2011**, *32*, 1456.
59. Siegbahn; P. E. M; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421.
60. Warshel, A. Computer Modeling of Chemical Reactions in Enzymes and Solutions; *Wiley*: New York, **1991.**
61. Ramos, M. J.; Fernandes, P. A. *Acc. Chem. Res.* **2008**, *41*(6), 689.
62. Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn P. E. M. *Chem. Rev.* **2014**, *114*, 3601.
63. Himo, F.; Siegbahn P. E. M. *J. Biol. Inorg. Chem.* **2009**, *14*, 643.
64. Siegbahn, P. E. M.; Himo, F. Wiley Interdiscip. Rev.: *Comput. Mol. Sci.* **2011**, *1*, 323.
65. Fukui, K. *J. Phys. Chem.* **1970**, *74*, 4161.
66. Gonzalez, C.; Schlegel H. B. *J. Chem. Phys.* **1989**, *90*, 2154.

67. Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. NBO, version 3.1.
68. Marino, T.; Russo, N.; Toscano, M. *J. Am. Chem. Soc.* **2005**, *127*, 4242.
69. Bottoni, A.; Lanza, C. Z.; Miscione, G. P.; Spinelli, D. *J. Am. Chem. Soc.* **2004**, *126*, 1542.
70. Miscione, G. P. Stenta, M. Spinelli, D. Anders, E. Bottoni, A. *Theor. Chem.* Acc. **2007**, *118*, 193.
71. Rossi, A.; Hoffmann, R. *Inorg. Chem.* **1975**, *14*, 365.
72. Siegbahn; P. E. M; Blomberg, M. R. A. In Theoretical Aspects of Homogeneous Catalysis; van Leeuwen, P. W. N. M.; Morokuma, K.; van Lenthe J. H. Eds. Kluwer Academic Publishers: Dordrecht, **1995**
73. Musashi, Y.; Shigeyoshi, S. *J. Chem. Soc.,* Dalton Trans., **1998**, 577.
74. Braunstein, P.; Matt, D.; Nobel, D. *Chem. Rev.* **1988**, *88*, 747.
75. Mellot-Draznieks, C.; Valayannopoulos, V.; Chrétien, D.; Munnich, A.; de Lonlay, P.; Toulhoat H. *ACS Catal.* **2012**, *2* (12), 2673.
76. Jackson, C. J.; Foob, J.-L.; Tokurikic, N.; Afriatc, L.; Carr, P. D.; Kimb, H.-K.; Schenke, G.; Tawfik, D. S.; Ollis, D. L. *PNAS*, **2009**, 106, 21631.
77. Hutschka, , F; Dedieu, , A; Eichberger, M; Fornika, R.; Leitner, W. *J. Am. Chem. Soc.* **1997**, *119*, 4432.
78. Jessop, P. G.; Ikariya, T.; Noyori, R. *J. Am. Chem. Soc.* **1996**, *118*, 344.

Paper V

Theoretical investigation on the restoring step of the carbonic anhydrase catalytic cycle

Paolo Piazzetta, Tiziana Marino and Nino Russo

Manuscript in preparation

Introduction

An enzyme is defined as a species that increases the rate of a chemical reaction by providing an alternative pathway of lower activation energy[1,2].

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} EP \xrightarrow{k_4} E + P \qquad \textbf{Eq. 1}$$

In order to efficiently catalyze a chemical reaction, enzymes are required to maintain fast rates for formation of the Michaelis complex (ES), the chemical reaction (formation of EP from ES) and product release (E+P). The turnover number (k_{cat}) is a resultant of a series of microscopic rate constants related to the formation (k_1) and dissociation (k_2) of the Michaelis complex (ES), chemical event (k_3), and product release (k_4), as shown in Eq 1.

In some cases can occur that the chemical step in the reaction (k_3) does not represent the rate determining step but the occurring conformational changes required in the product release make the last step governed by (k_4) that having the lowest rate.

Having this in mind we have taken into account the case of human carbonic anhydrase in order to inspect the factors preventing product release as in the particular case of the suicide inhibitor ureate in comparison with the release of natural product (HCO_3^-). [3,4]

The generally accepted catalytic mechanism of carbonic anhydrase is constituted by two fundamental steps[5], where the Zn-OH moiety catalyzes the interconversion of CO_2 to HCO_3^- (**a**), that can be replaced by a water molecule to re-establish the next enzymatic cycle (**b**):

$$E\text{-}Zn\text{-}OH^- + CO_2 + H_2O \rightleftharpoons E\text{-}Zn\text{-}HCO_3^- \qquad \textbf{a)}$$

$$E\text{-}Zn\text{-}HCO_3^- + H_2O \rightleftharpoons E\text{-}Zn\text{-}H_2O + HCO_3^- \qquad \textbf{b)}$$

While the step (**a**) related to the chemical event described by the activated water to the polar carbonyl group of the native substrate,

represented the main focus of different theoretical studies [6-8](**Paper I**), the dissociation of the product did not receive the same attention.

The examined inhibitor of hCA is the carbodiimide molecule that normally binds to the active site of the enzyme as proposed by experimental inspection [9-11] and normally processed as recently confirmed by theoretical investigation (**Paper I**), but it prevents the release of the final product (ureate) which results to be strongly bonded to the enzyme. The enzyme consequently commits suicide by activating the inhibitor (carbodiimide). Since suicide inhibitors take advantage of the enzyme mechanism to activate an initially innocuous compound that then irreversibly inactivate the enzyme, they can be extremely used as effective drugs. Adding detailed insights into the native reaction catalyzed by the enzyme can just represent the first step for clarifying its inhibiting activity but a focused research is required for finding the reason for this behaviour. While the QM and QM/MM study of the chemical behavior of hCA towards carbodiimide (**Paper I**) in comparison with the native CO_2 substrate has been performed to gain insights on the main steps characterizing the catalysed reaction, the physical process by which final product is released in order to restore the catalytic cycle remains still unclear.

As far as we are aware, the steered molecular simulations (SMD) of products detachment proposed in the present study, represents the first investigations intended to unravel at atomic level the factors involved in the releasing process of the product and deducible from the molecular dynamics simulations. In fact, the SMD simulation can shed light on the optimal binding – unbinding pathway of the ligand

Starting from our previous mechanistic investigations on the hCAII activity towards carbon dioxide and carbodiimide (**Paper I**), we validated a strategy based on nonequilibrium SMD simulations method to determine the potential energy related to the release's process of the two reaction products bicarbonate and ureate, for CO_2 and carbodiimide respectively. In order to restore the catalytic cycle a water molecule must substitute the reaction product. In the case of native substrate, our QM and QM/MM computations showed that this substitution can easily occur, but on the contrary, in the case of ureate the water coordination on Zn^{2+} and the release of the product seem to be a difficult process and strictly related to the interaction between the reaction product and the metal core.

Conventional MD

MD simulations were carried out with NAMD 2.9 [12] using CHARMM[13] force field. The native binding poses for both HCO_3^-- hCAII and H_2NCOHN^-- hCAII complexes obtained by X-ray investigations (PDB Code 2VVB [14] and 1BV3 [9] respectively) were used as starting structures. Both ligands required custom parameters for the force field. They were obtained from parameters by similar models through analogies (bicarbonate) and parameterized in order to create the topology and input files required by the CHARMM calculations starting from the Gaussian output files (ureate). In order to reproduce the natural flexibility of the enzymatic metal core and to preserve the zinc coordination, a model was built up combining bonded and non–bonded terms: the zinc center was described introducing explicit bonds but devoid of a bond to the ligand. Constraints on the zinc binding site were introduced using the Colvars module of NAMD with values of bonds, angles and dihedral arising from QM/MM calculations (**Table1** and **Table 2**) (**Paper I**), which were restrained with harmonic constraints[15]. The charges for the ligands atoms, for the metal core and for the coordinated histidines were set using the NBO charges from DFT calculations [**Supporting Information Paper I**].

The constraints of the angles defined by substrate – Zn – histidine ligands were only used during the MD simulation and not for the subsequent analysis. A water layer of 16 Å was built around each studied enzyme-product complex by creating a water box of 80 x 78 x 89 $Å^3$ for hCAII- HCO_3^- and 81 x 78 x 88 $Å^3$ for H_2NCOHN^-- hCAII. The resulting number of atoms was to 52101 for hCAII- HCO_3^- and 51054 for H_2NCOHN^-- hCAII. Due to the neutral charge of the studied systems no further neutralization was needed. SHAKE[16] algorithm was practiced in order to constrain all bonds involving hydrogen atoms. Periodic boundary conditions were used. The long-range electrostatic interactions were evaluated by Particle Mesh Ewald (PME) method[17]. The supramolecular systems obtained including protein, waters and the ligand (HCO_3^- and ureate) were minimized using the conjugate gradient algorithm: in first instance, minimizing the water molecules keeping the protein and ligand fixed, and then relaxing the entire system. A time step of Δt = 2 fs was used for both the investigated systems. A cutoff of 12 Å for non-bonded interactions was applied and a switching scheme was used. Scaled 1–4 parameters were enabled for 1–4 interactions. A first

equilibration (60 ps) was performed on the minimized models by increasing temperature until to 300 K in increments of 20 K. Then a simulation time equal to 10 ns was performed in the NPT ensemble with a Langevin dampening coefficient of 5 ps^{-1} controlling pressure with the Langevin piston pressure control. The global stability of two models was assessed by the *root mean square deviation* (RMSD) values (**Figure 1**). In order to identify the local stability of the active site, coordination bonds lengths and dihedral angle values were monitored during the NPT equilibration process. The results indicate that all the coordination bonds to the metal were stable in the simulation. Carbonyl carbon - Zn^{2+} length in bicarbonate undergoes an elongation at 1 ns due to the formation of the pentacoordinated complex with the coordination of a free water molecule. Instead during the ureate equilibration the carbonyl carbon - Zn^{2+} distance has a different behavior maintaining an average length of 2.5 Å (**Figure 2** and **Figure 3**). All the checks on the interaction details suggest that the active sites of two models are locally stable. In general, the global and local conformations of the two complexes both validate the stability at the free MD stages describing accurately the different coordination, i.e, in term of water activated complex for the bicarbonate complex, which is really necessary for the sequential SMD simulations.

Table 1: Parameters for the zinc coordination in the bicarbonate–hCAII complex.

Atoms	Force Constant	Equilibrium
Bonds	k	r_e
Zn N_{119}	30 kcal/mol/Å^2	2.07 Å
Zn N_{94}	30 kcal/mol/Å^2	2.01 Å
Zn N_{96}	30 kcal/mol/Å^2	2.01 Å
Angles	K	θ_e
N_{119} Zn N_{96}	23 kcal/mol/rad^2	96.9°
N_{96} Zn N_{94}	23 kcal/mol/rad^2	107.2°
N_{94} Zn N_{119}	23 kcal/mol/rad^2	118.2°
$O_{bicarbonate}$ Zn N_{119}	0.05 kcal/mol/rad^2	104°
$O_{bicarbonate}$ Zn N_{94}	0.05 kcal/mol/rad^2	108°
$O_{bicarbonate}$ Zn N_{96}	0.05 kcal/mol/rad^2	120°
Improper Dihedrals	K	ϕ_0
Zn N_{119} N_{96} N_{94}	25 kcal/mol/rad^2	0

Table 2: Parameters for the zinc coordination in the ureate–hCAII complex.

Atoms			Force Constant	Equilibrium
Bonds			k	r_e
Zn	N_{119}		30 kcal/mol/Å^2	2.10 Å
Zn	N_{94}		30 kcal/mol/Å^2	2.04 Å
Zn	N_{96}		30 kcal/mol/Å^2	2.05 Å
Angles			K	θ_e
N_{119}	Zn	N_{96}	23 kcal/mol/rad^2	114.0°
N_{96}	Zn	N_{94}	23 kcal/mol/rad^2	100.0°
N_{94}	Zn	N_{119}	23 kcal/mol/rad^2	95.0°
O_{ureate}	Zn	N_{119}	0.05 kcal/mol/rad^2	111.5°
O_{ureate}	Zn	N_{94}	0.05 kcal/mol/rad^2	121.0°
O_{ureate}	Zn	N_{96}	0.05 kcal/mol/rad^2	110.0°
Improper Dihedrals			K	ϕ_0
Zn	N_{119}	N_{96} N_{94}	25 kcal/mol/rad^2	0

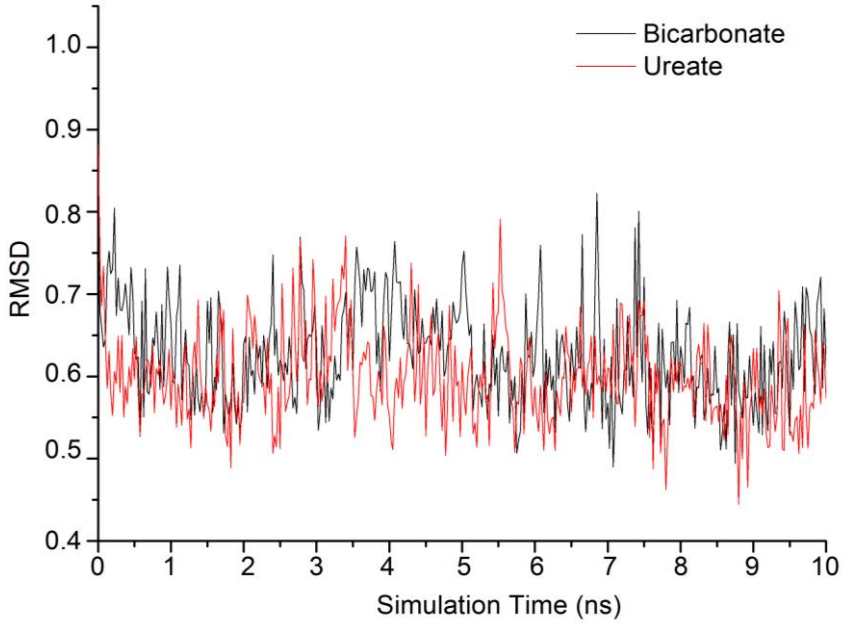


Figure 1 RMSD analysis referred to the MD simulation.

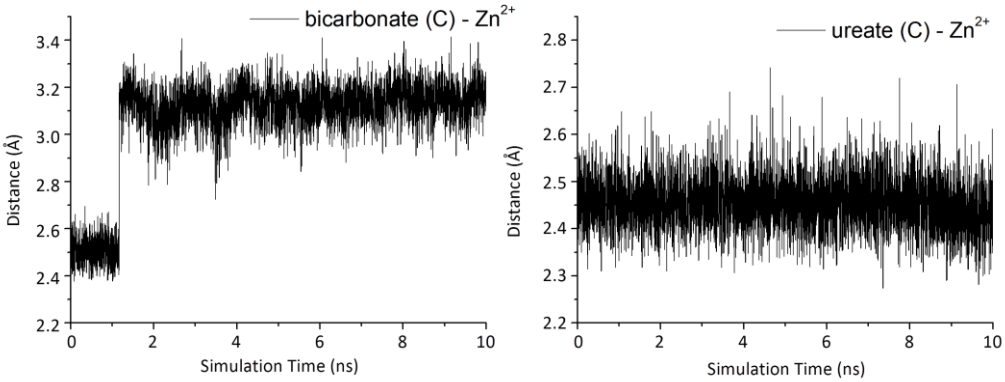


Figure 2. Time dependence of the Zn-coordinated bond lengths in the active sites of hCAII – bicarbonate in the 10 ns NPT MD simulations.

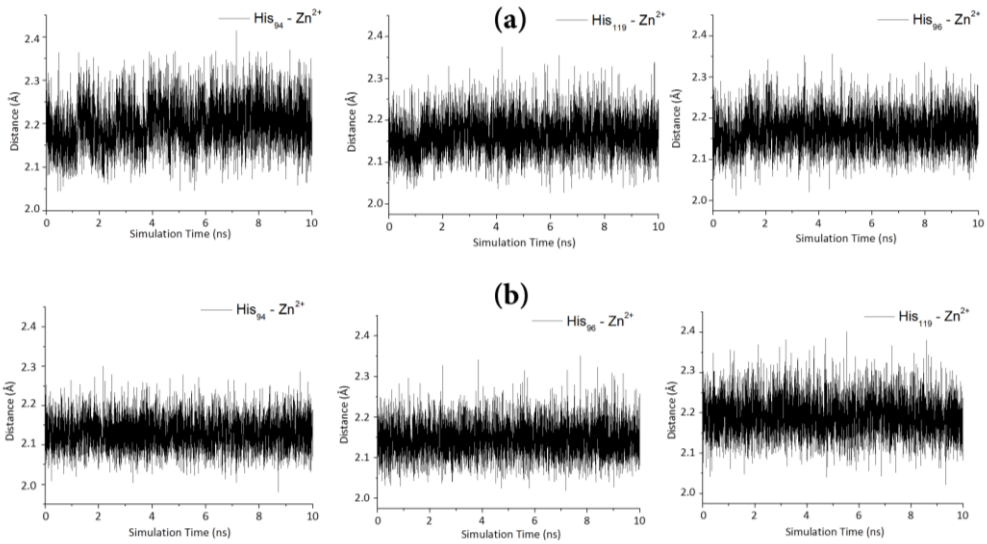


Figure 3. Time dependence of the Zn-coordinated bond lengths in the active sites of hCAII – bicarbonate (**a**) and hCAII – ureate (**b**) in the 10 ns NPT MD simulations.

SMD simulation

The final equilibration state was used as initial point for further SMD. For both ligands ten 1 ns SMD simulations were carried out. The external force was applied to the carbonyl carbon of each substrate which coincides approximately with the center of mass. In the adopted constant velocity SMD, the value of the spring force varies significantly depending on what pulling velocity and spring constant were applied to the system. Higher pulling velocities result in disequilibrium, that introduce a significant error into the simulation results. We carried out 5 different pulling velocity simulations (ranging from 0.000025 Å /ts to 0.005 Å /ts) to identify a suitable and compatible velocity for monitoring the detachment step of two examined ligands. The "rupture" force was defined as the highest peak value of the pulling force. A pulling velocity of 0.00005 Å /ts was chosen because it better reproduces the different behavior of both ligand. In order to compare binding properties of the tow different complexes, the SMD simulations must be carried out by using the same v as previously suggested [18,19].

It is noteworthy that the spring constant value must not be so large because high level of background noise could be induced with consequent difficulty in the measurement. Due to this problem different spring constant values

The other parameter to be set is the spring constant which must be large enough to enable the local dissociation and to ensure small deviation of the reaction coordinate from the constraint position. It is noteworthy that the spring constant value must not be so large because high level of background noise[20,21] could be induced with consequent difficulty in the measurement. Due to this problem different spring constant values (5, 7, 10, 15 kcal/mol/ $Å^2$) have been tested in the case of bicarbonate ligand. **Figure 4** shows as under the same pulling velocity, the spring constants of 10 and 15 kcal/mol/ $Å^2$ caused large fluctuations in rupture force, while spring constants of 5 and 7 kcal/mol/ $Å^2$ showed a reduced effect. A spring constant value of 5 kcal/mol/ $Å^2$, produced a measurable and consistent level of rupture forces and was therefore chosen for subsequent SMD simulations for both ligands.

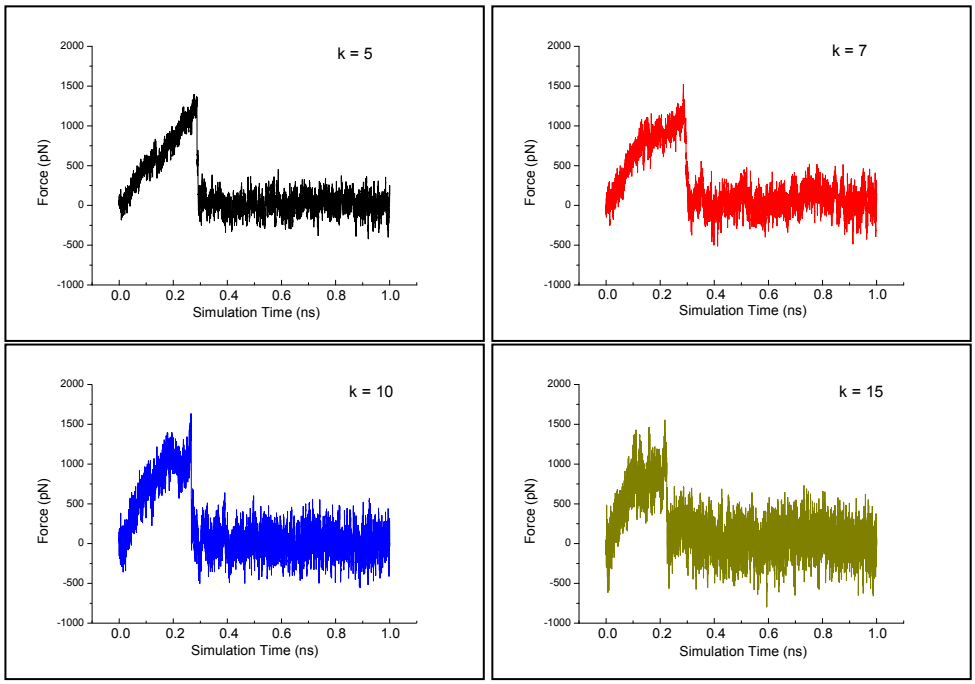


Figure 4. Applied force in bicarbonate using different spring constant (kcal/ mol / Å^2) at the same pulling velocity equal to 0.00005 Å /ts.

Choice of pulling path

Since the rupture force is sensitive to pulling direction,[22-24] it is fundamental to determine the opportune pulling path to be chosen. hCAII, as above described, is characterized by a conical cavity at the active site which is 15 Å deep and about Å in diameter at its mouth. To probe pulling pathways in term of normalized vector we used Caver 2.0, a plugin of Pymol[25]. **Figure 5** shows the optimal path that is directed outside the binding pocket. This is the univocal pathway (lowest rupture force) to pull carbonate or ureate from the active site of hCAII. For all the SMD simulations a normalized vector was deduced from the calculated channel (**Figure 6**). When performing the pulling simulations, the Cα of four residues (Asp247, Ser48, Pro195, Val223) away from the active site were kept fix to maintain the same system orientation during all the SMD simulations.

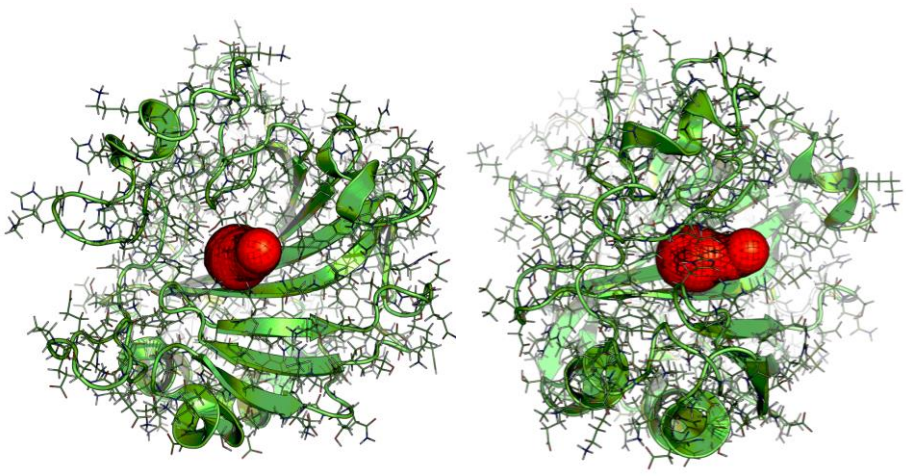


Figure 5. Mechanical unbinding path obtained obtained by Caver 2.0

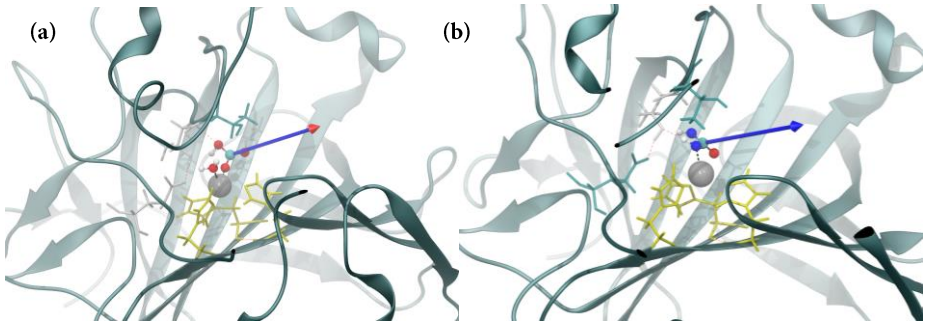


Figure 6. Initial structure for pulling related to bicarbonate (**a**) starting from the pentaccordinated complex and ureate (**b**) starting from tetrahedral complex. The evidenced residues are His94, His96, His119, Glu106, Thr199, Thr200. The solid arrow represents the normalized pulling vector.

Ligand's dissociation

The rupture force $\mathbf{F_{max}}$ which defines the stability of the ligand, directly related to unbinding process suggests that the larger is the force needed to unbind a ligand the higher is the related binding affinity. For this reason, the rupture force is used as score function to measure binding affinity.

Figure 7 shows the profile of the applied force along the chosen expulsion direction as a function of simulation time for both HCO_3^- and H_2NCOHN^- products in a single pulling event. The ureate relese starts taking place after 0.33 ns and the required force is equal to 1573 pN while a shorter time (0.28 ns) with an applied force of 1280 pN is required for releasing the bicarbonate product.

The starting complexes are characterized by different substrate – Zn^{2+} distances as above mentioned and depicted in figure **Figure 2** (3.1 – 3.3 Å $Zn - C\,HCO_3^-$, 2.4 – 2.6 Å $Zn - C\,H_2NCOHN^-$)

PMF curves for both the substrates were constructed from ten repeated SMD simulations along the predefined direction path.

The calculated PMF profiles as function of the simulation time using the second cumulant expansion (*see Chapter 2*) for bicarbonate and ureate are 10.8 kcal/mol and 25.2 kcal/mol, respectively. As reported in **Figure 8,** it is possible to deduce that the standard deviation assumes a significant value when the ligands start the departure from the metallic center.

After the unbinding events the ligands show a different behaviour in term of linear diffusion in the solvent. Bicarbonate has a relative standard deviation of 2 kcal/mol (1.2 k_BT) less than ureate which seems to diffuse differently with a deviation from the standard PMF value of 7 kcal/mol (4.2 k_BT). In general the minor deviations of bicarbonate respect that of ureate can be attributed to fluctuations due to short sampling time and could be improved by using a larger number of samples.

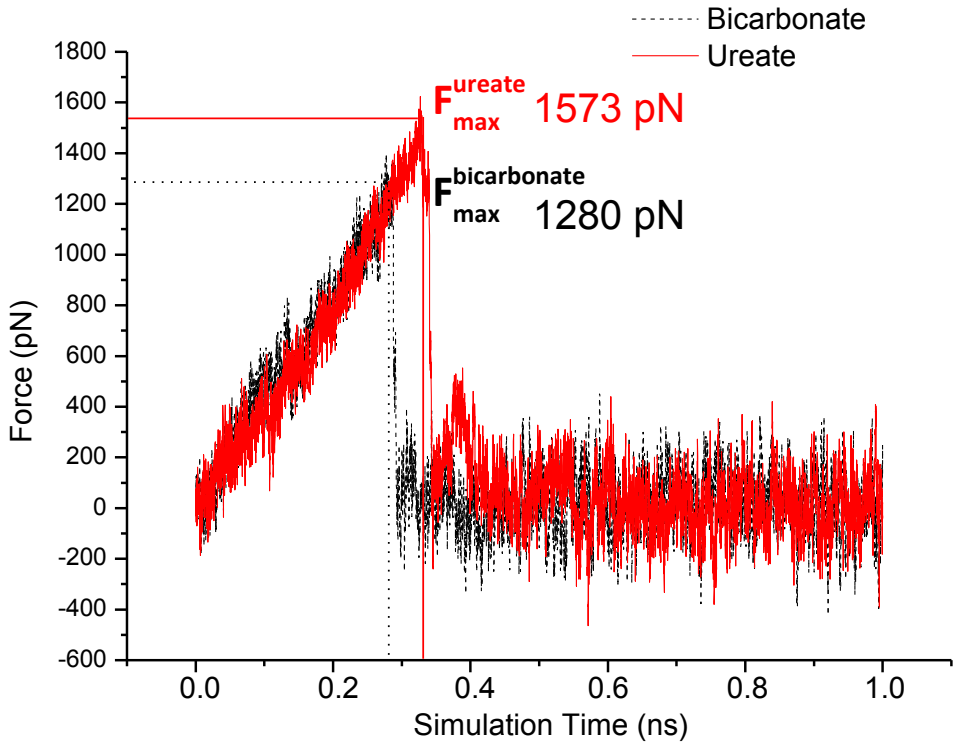


Figure 7. Single pulling event in bicarbonate and ureate complex using k=5 kcal/mol/Å^2 and v=0.00005 Å/ts.

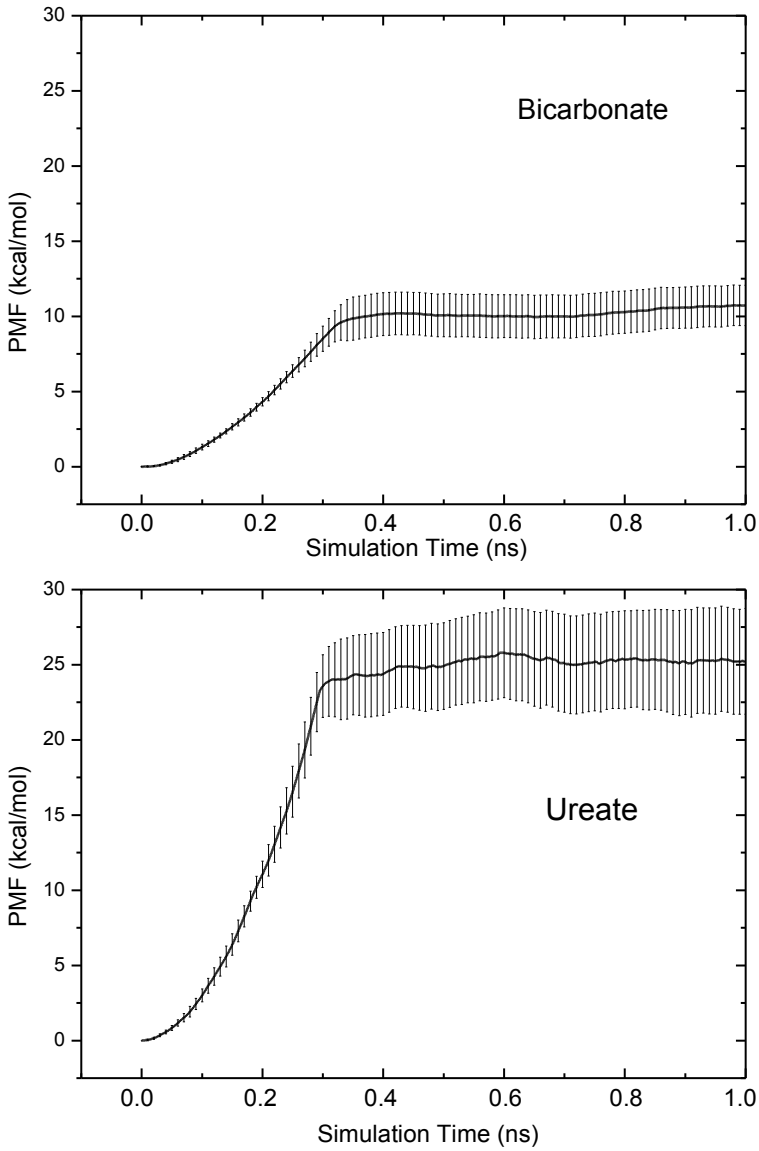


Figure 8. PMF profile as function of the simulation time of bicarbonate and ureate (k=5 kcal/mol/Å^2 and v=0.00005 Å/ts) through the block of 10 trajectories. The error bars indicate the standard deviation over the blocks.

Conclusion

In order to analyze the unbinding process of hCA – bicarbonate vs the unbinding process of hCA – ureate useful in the elucidation of the inhibitory behaviour of the product of the hydrolysis of carbodiimide, constant velocity SMD simulations have been performed. From the pulling simulations can be evinced that ureate binds to the catalytic metal centre of hCA with a higher affinity than that of the natural reaction product bicarbonate. The observed different behaviour of the two products strengthens what observed in the previous QM and QM/MM investigations (**Paper I**): the bicarbonate interacts with the enzyme allowing the formation of a pentacoordinate complex following the water entrance that promotes the expulsion with the consequent restoring of the catalytic cycle. The lacking of the pentacoordinated complex due to the prevention of the water molecule coordination in the case of ureate, is also evidenced by MD and SMD simulations. The computed free energy for the hCA – ureate complex can account for its difficult dissociation and points out the suicide inhibitory nature of the formed products.

The resulting computed free energy for the ureate – hCAII is about twice of that relating to the bicarbonate.

[1] Radzicka, A.; Wolfenden, R. *Science* 1995, *267*(5194), 90.
[2] Hong, S. B.; Raushel, F. M. *Biochemistry* 1999, *38*(4),1159.
[3] Mellot-Draznieks, C.; Valayannopoulos, V.; Chrétien, D.; Munnich, A.; de Lonlay, P.; Toulhoat, H. *ACS Catal.* 2012, *2* (12), 2673.
[4] Jackson, C. J.; Foob, J.-L.; Tokurikic, N.; Afriatc, L.; Carr, P. D.; Kimb, H.-K.; Schenke, G.; Tawfik, D. S.; Ollis, D. L. *PNAS*, 2009, *106*, 21631.
[5] Silverman, D. N. ; Lindskog, S. *Acc. Chem. Res.* 1988, *21*, 30.
[6] Bottoni, A.; Lanza, C. Z.; Miscione, G. P.; Spinelli, D. J. *Am.Chem. Soc.* 2004, *126*, 1542.
[7] Mauksch, M.; Braüer, M.; Weston, J.; Anders, E. ChemBioChem 2001, *2*, 190.
[8] Cui, Q.; Karplus, M. *J. Phys. Chem. B* 2003, *107*, 1071.
[9] Briganti, F.; Mangani, S.; Scozzafava, A.; Vernaglione, G.; Supuran, C.T. J. *Biol. Inorg. Chem*. 1999, *4*, 528.
[10] Scolnick, L. R.; Christianson, D. W. Biochemistry 1996, *35*,16429.
[11] Guerri, A.; Briganti, F.; Scozzafava, A.; Supuran, C. T.; Magani, S. *Biochemistry* 2000, *39*, 12391.
[12] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comp. Chem.* 2005, *26*, 1781.
[13] Vanommeslaeghe, K. et al *J. Comp. Chem.* 2010, *31*, 671.
[14] Sjoblom, B.; Polentarutti, M.; Djinovic-Carugo, K. *Proc.Natl.Acad.Sci.USA* 2009 *106*, 10609.
[15] Schmid, M.; Nogueira, E. S.; Monnard, F.W.; Ward, T. R.; Meuwly, M. *Chem Sci*. 2012, *3*, 690.
[16] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys*. 1977, *23*, 327.
[17] Darden, T.; York, D.; Pedersen. L. *J. Chem. Phys.* 1993, *98*, 10089.
[18] Merkel, R.; Nassoy, P.; Leung, A.; Ritchie, K.; Evans, E. *Nature* 1999; *397*, 50.
[19] Heymann, B.; Grubmuller, H. *Chem Phys Lett* 1999, *303*, 1.
[20] Zhang, J.; Zheng, Q.; Zhang, H. *Chem PhysLett* 2010, *484*, 338.
[21] Park, S., Khalili-Araghi, F.; Tajkhorshid, E.; Schulten; K. *J. Chem. Phys*. 2003, *119*, 3559.
[22] Kumar, S.; Li, M. S.: Phys. Rep. 2010, *486*, 1.
[23] Kouza, M.; Hu, C. K.; Li, M. S. *J. Chem. Phys.* 2008, *128*, 045103.
[24] Carrion-Vazquez, M.; Li, H. B.; Lu, H.; Marszalek, P. E.; Oberhauser, A. F.; Fernandez, J. M. *Nat. Struct. Biol.* 2003, *10*, 738.
[25] Petrek, M.; Otyepka, M.; Banas, P.; Kosinova, P.; Koca1, J.; Damborsky, J. CAVER: *BMC Bioinf*. 2006, *7*, 316.