Electronic Structure

CrossMark

RECEIVED
7 September 2022

REVISED
16 December 2022

ACCEPTED FOR PUBLICATION
4 January 2023

PUBLISHED
26 January 2023

PAPER

QM-cluster model study of CO_2 hydration mechanisms in metal-substituted human carbonic anhydrase II

Thomas J Summers[1] **and Nathan J DeYonker**[*]

Department of Chemistry, The University of Memphis, 213 Smith Chemistry Building, Memphis, TN 38152-3550, United States of America

[1] Current address: Department of Chemical and Materials Engineering, University of Nevada, Reno, Reno, NV 89557, United States of America.

[*] Author to whom any correspondence should be addressed.

E-mail: ndyonker@memphis.edu

Keywords: carbonic anhydrase, reaction mechanism, metal substitution, computational enzymology

Supplementary material for this article is available online

Abstract

Human carbonic anhydrase (CA) metalloenzymes utilize a Zn^{2+}-containing active site to catalyze the interconversion of carbon dioxide to bicarbonate. The Zn^{2+} ion may be replaced with other divalent transition metals, though the catalytic efficiency of the enzyme will be reduced. In this work, quantum mechanical cluster models of the active site are used to map the reaction profile for the hydration mechanism of carbon dioxide. The Lipscomb proton transfer and Lindskog rotation mechanisms were examined for the native Zn^{2+}-enzyme along with variants where the metal was substituted with Cd^{2+}, Ni^{2+}, Fe^{2+}, and Fe^{3+}. The findings highlight the impact the metal coordination geometry has on the reaction profile. The results also suggest Fe^{2+}, which is the functional metal for a prototypical CA of an anaerobic bacterium, might also be functional for human CA if cultured within an anaerobic environment.

1. Introduction

Metalloenzymes have crucial roles in living organisms, ranging from facilitating cellular signal transduction pathways to functionalizing substrates. Carbonic anhydrases [1] (CAs) are a family of metalloenzymes that rely upon a zinc-bound active site to catalyze the reversible hydration of carbon dioxide through the reaction:

$$CO_2 + H_2O \rightleftharpoons HCO_3^- + H^+.$$

Because of this functionality, CAs are found widely throughout organisms in all three biological domains and play important roles in cellular respiration, pH and fluid homeostasis, and carbon dioxide fixation [1–3].

Although there are several different isoforms of CA, one of the most well studied is human carbonic anhydrase II (HCAII), which is a 32 kDA monomeric α-class protein that supports a four-coordinate Zn^{2+} center (figure 1(A)) [4–6]. This protein is cambialistic in which the native Zn^{2+} may be substituted with other divalent metal ions and the enzyme still retains some catalytic activity. In general, the metal binding affinity of CAs follow the Irving–Williams series (defined as $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$) with the exceptions of the native CA Zn^{2+} having a greater affinity than Cu^{2+}, and Fe^{2+} having a smaller affinity than Mn^{2+} [7, 8]. While HCAII and other α-class CAs are capable of binding to different divalent transition metals, the activity of metal-substituted variants is significantly reduced. For example, α-class CAs bound to Co^{2+}, Mn^{2+}, Fe^{2+}, Ni^{2+}, and Cd^{2+}-bound show approximately 50%, 8%, 4%, 2%, and 2% activity, respectively compared to Zn^{2+} [7, 9–11].

It has been reported that the enzyme carbonic anhydrase of *Methanosarcina thermophila* (CAM), a prototypical γ-class CA from an anaerobic archaeon, contains Fe^{2+} when synthesized in an anaerobic environment *in vivo*; furthermore, the Fe^{2+}-CAM enzyme has a catalytic efficiency greater than Zn^{2+}-CAM [12, 13]. Exposing the Fe^{2+}-CAM enzyme to hydrogen peroxide or air inactivated the enzyme by oxidizing

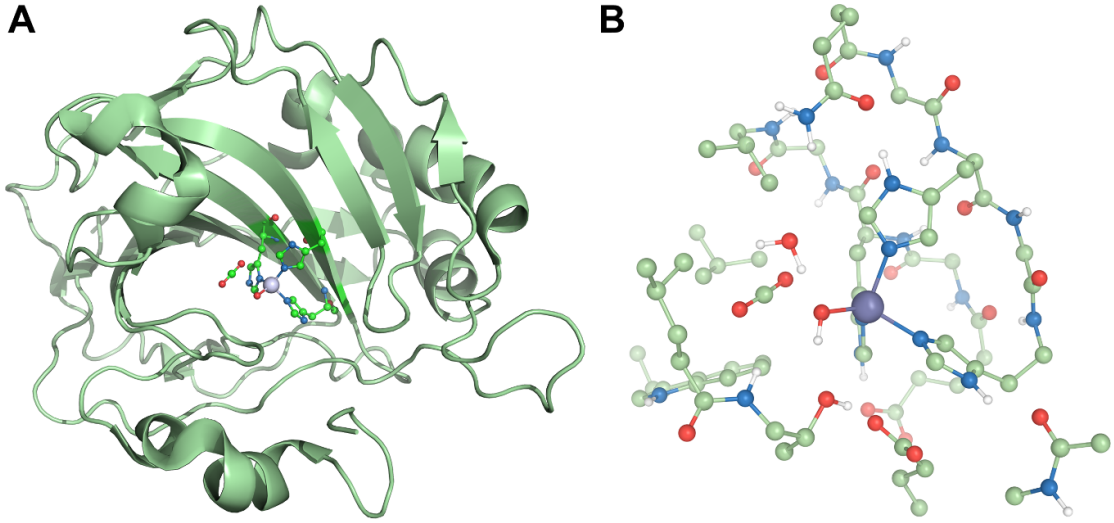

Figure 1. (A) Cartoon representation of HCAII (PDB: 3D92) highlighting the location of CO_2 within the active site nearby the metal ion coordinated to three histidines and a metal-bound hydroxide. (B) Ball-and-stick model representation of the QM-cluster model. Non-polar hydrogens are omitted for clarity. Carbons, oxygens, hydrogens, and the metal are colored green, blue, white, and grey, respectively.

Fe^{2+} to Fe^{3+}, which then dissociates from the active site [12, 13]. These results suggest that the previous reports of poor activity for Fe-substituted CAs may originate from sample purification in an aerobic environment causing Fe^{2+} to oxidize and be replaced with trace Zn^{2+} contaminants present in the buffer [9, 13]. Since both α- and γ-class CAs possess similar metal-coordination sites (three histidines and a hydroxide, figure 1(A)) [14], and since there is evidence that including Fe^{2+} in the reaction medium with Zn^{2+} can improve α-class CA catalytic activity [15], it is warranted to re-investigate the structure, activity, and catalytic mechanism of Fe^{2+}-substituted α-class CAs within an anaerobic environment.

The catalytic cycle (figure 2) begins with the enzyme sequestering and orienting the CO_2 into a hydrophobic region of the active site. There are two commonly proposed CO_2 hydration mechanisms for native HCAII. Both begin with a Zn^{2+}-bound hydroxide performing a nucleophilic attack on carbon dioxide to form bicarbonate with two oxygens bound to the metal [16, 17]. From this intermediate, the Lipscomb mechanism proposes that dual proton transfers occur between the bicarbonate and the nearby, evolutionarily conserved threonine and glutamic acid residues before the bicarbonate detaches from the metal [18, 19]. Alternatively, the Lindskog pathway proposes the bicarbonate breaks away without additional proton transfer steps [6]. A water molecule then replaces bicarbonate, and the catalytic hydroxide is regenerated via a proton dissociation shuttle transferring a water proton to the protein surface and subsequent solvent [20–23]. For decades, there have been numerous theoretical studies and debates over which CO_2 hydration pathway is preferred [16, 17, 19, 24–35]. Historical differences in opinion arise in part from the advancement of computational resources over time. Initial studies from the 1990s were limited to minimal 18–33 atom quantum mechanical-cluster (QM-cluster) models of only the metal-coordinated atoms and CO_2 [26, 35, 36], but modeling capabilities have since grown to include simulating the surrounding protein and water environment through both larger QM-cluster models [19, 24, 37, 38] and hybrid quantum mechanical/molecular mechanics (QM/MM) models [27, 34, 38, 39]. Although many QM-only and recent QM/MM models propose the Lindskog path to be the energetically favored mechanism, thermodynamic differences between the two paths are frequently noted to be miniscule and that both paths may be competitive with each other [26, 27, 35–37, 39].

Although there have been experimental and theoretical studies on the CO_2 hydration mechanisms for HCAII substituted with different transition metals (such as Co^{2+}, Ni^{2+}, Cu^{2+}, and Cd^{2+}) [27, 28, 40, 41], there has not been further investigation on the mechanistic feasibility of Fe^{2+}-HCAII or other α-class CAs within an anaerobic environment. In this work, we seek to investigate the CO_2 hydration mechanism of Fe-substituted HCAII by quantum mechanically simulating both the proposed Lipscomb and Lindskog reaction mechanisms. For improved context and comparison, reaction pathways for models with Zn^{2+}, Ni^{2+}, and Cd^{2+} as the active site ions will also be computed. The system will be modeled with QM-cluster models designed using information from protein residue interaction networks (RINs) [42, 43]. Through this work, we seek to effectively obtain atomic-level insight into the HCAII active site and the impact different metals ions have on the CO_2 hydration mechanism.

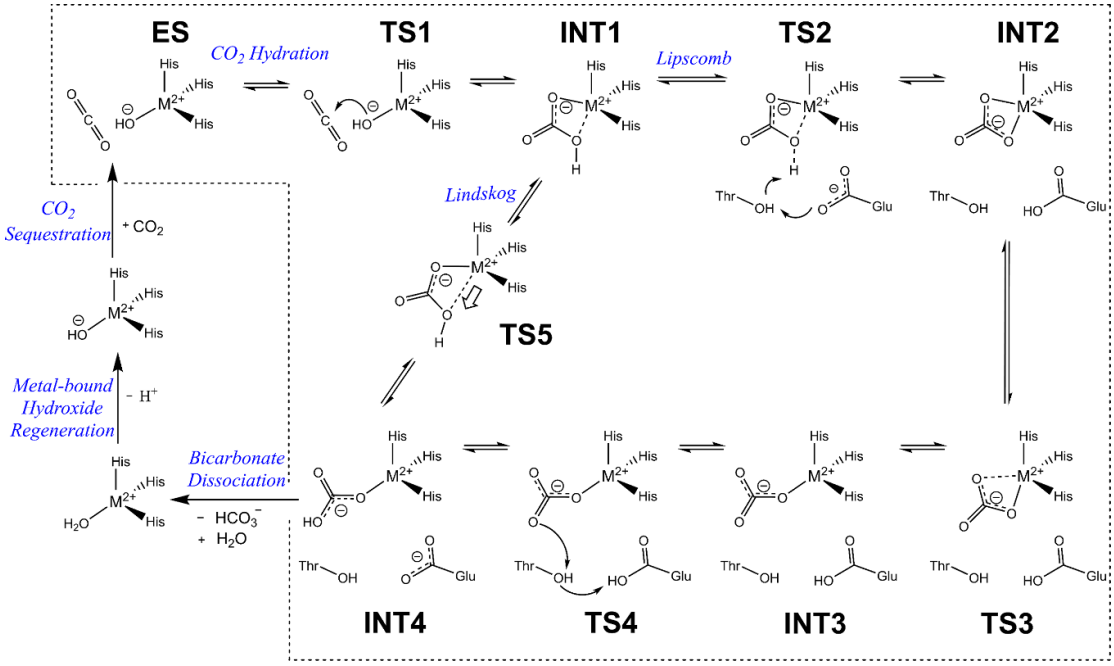

Figure 2. Scheme of the catalytic cycle of HCAII. The proposed Lipscomb and Lindskog reaction steps for carbon dioxide hydration mechanism examined in this work (encircled) are detailed with proposed intermediate and transition state structures (M = Zn^{2+}, Cd^{2+}, high-spin Fe^{2+}, low-spin Fe^{3+}, and high-spin Fe^{3+}).

2. Methods

The QM-cluster models of the CA active site were constructed using the x-ray crystallographic coordinates of a cobalt-substituted derivative of HCAII [Protein Data Bank code (PDB): 3KOI] [44]. Hydrogen atoms were added to the structure using the program *reduce* [45], and correct residue protonation states were verified against a neutron diffraction-resolved HCAII structure [46]. To identify the residues that craft the active site microenvironment, the inter-residue topologies for three HCAII x-ray crystal structures (PDBs 3KOI, 3D92, and 1CAH) [44, 47, 48] were mapped by using the *probe* software [49] to identify inter-residue contact interactions. The computed inter-residue contact interactions were used to construct a RIN, a graph that translates the three-dimensional protein structure into a two-dimensional network of residues (called 'nodes') interconnected by their contact interactions (called 'edges') [42, 43]. Collectively, this workflow for model design follows a prototype of the Residue Interaction Network ResidUe Selector toolkit [50].

The three x-ray crystal structures examined each have unique characteristics which ensures crucial interactions are captured in modeling the various metal-substituted systems. The cobalt-substituted structure PDB:3KOI has three waters coordinated to the metal to give an octahedral geometry [44]. The PDB:3D92 structure is obtained from CO_2-pressurized, cryo-cooled crystals, capturing the location of CO_2 within the active site [47]. The cobalt-substituted structure PDB:1CAH has bicarbonate product complexed to the metal [48]. The contact maps of these three proteins were analyzed to identify residues (table S1) interacting with either the metal-coordinated waters, the three metal-coordinated histidine side chains (His94, His96, His119), the unbound CO_2 molecules (within 3D92) and the metal-bound bicarbonate molecule (within 1CAH). A total of 15 residues were identified as having contact interactions with the aforementioned species. Three interacting residues were excluded from the final model (Asn67 and Thr200 had very few contact interactions and Phe66 only had minor non-hydrogen bonding backbone contacts with His94), allowing the final list of residues contained in the QM-cluster models to be Gln92, Phe93, Phe95, Glu106, Glu117, Leu118, Val121, Val143, Leu198, Thr199, Trp209, and Asn244. Though the contact maps do not predict the function of the identified residues, the residues have been previously identified as having important roles in crafting the active site. Val121, Val143, Leu198 and Trp209 form the hydrophobic cavity responsible for entrapping and orienting CO_2 before the hydration reaction begins. Gln92, Glu117, and Asn244 hydrogen bond with the three metal-coordinated histidine side chains to stabilize and orient them. The main chains of Phe93, Phe95, and Leu118 are involved in the hydrogen bonding networks formed among the metal-coordinated histidine main chains. Lastly, Glu106 and Thr199 form the hydrogen bonding network involved in the Lipscomb hydrogen transfer mechanism.

The rest of the protein was trimmed, cutting at the N–C_α or C_α–CO bonds adjacent to a residue side chain and capping the C_α where the bond was broken with a C–H bond to satisfy valency. Residues that only had main chain interactions (and no side chains interactions) with the active site similarly had their side chain trimmed and capped with hydrogen (see table S2 in supporting information). A water molecule was also included in the model positioned at the crystallographic coordinates of Wat592 of PDB entry 3D92. The final model for the Zn-, Fe-, and Cd-composed active sites is composed of 224 atoms (figure 1(B)). An additional metal-coordinated water is added to the Ni-substituted model to yield a 227-atom model. To mimic the generally rigid nature of the protein backbone, a total of 17 C_α and 5 C_β atoms were frozen to their crystallographic coordinates (table S2 and figure S1). Ni-HCAII was modeled as a high-spin state complex, as is observed in biomimetic complexes and other Ni^{2+}-enzymes [51–53]. Fe-HCAII models are examined in both low- and high-spin states, though the high-spin state is what had been observed experimentally [12].

QM-cluster model computations were conducted using the Gaussian16 software and employed density functional theory [54]. The hybrid B3LYP exchange-correlation functional [55, 56] was used with 6–31G(d′) basis set for N and O atoms [57], 6–31G basis set for C and H atoms [58], and LANL2DZ basis and effective core potentials for the metals as modified by Couty and Hall [59, 60]. The models were simulated with the inclusion of the GD3BJ dispersion correction and a conductor-like polarizable continuum model (CPCM) environment using universal force field (UFF)-based sets of atomic radii, a non-default electrostatic scaling factor of 1.2, and a dielectric constant of $\varepsilon = 4$ [61–63]. Unscaled harmonic vibrational frequency calculations were used to confirm all stationary points as minima (no imaginary vibrational frequencies) or transition states (one imaginary vibrational frequency).

3. Results and discussion

3.1. Proposed mechanism for Zn^{2+}-HCAII

The relative Gibbs free energies for stationary points along both Lipscomb and Lindskog reaction pathways (figure 2) were computed using the designed HCAII QM-cluster models with a Zn^{2+} centered first coordination sphere. The resulting free energy profile for the proposed mechanisms are shown in figure 3. The results indicate that at the point where the mechanisms deviate (INT1), the barrier to immediately break the Zn–O bond and form the product following the Lindskog mechanism requires 9.3 kcal mol^{-1} (TS5) whereas a dual hydrogen transfer following the Lipscomb mechanism is barrierless (TS2) and forms an intermediate (INT2) more stable than INT1 by 4.5 kcal mol^{-1}. However, continuing along the Lipscomb path then requires overcoming a 15.1 kcal mol^{-1} barrier (TS3) to break the metal-bicarbonate Zn–O bond, a more energetically costly feat compared to the Lindskog path. The dual proton transfer forming INT2 is readily reversible, though, given the low activation free energy to return to INT1. This reversibility may permit the reaction to still follow the less energetically costly Lindskog mechanism even if the proton transfer occurred. Nonetheless, since the Lipscomb and Lindskog rate determining steps are feasible channels, and since both pathways are energetically more stable than the starting reactants, both paths are expected to occur at physiological conditions.

In comparing these QM-cluster model results to other free energy profiles in the literature, the thermodynamics of our computed mechanisms have characteristics attributable to the absence of extensive water networks within the active model (e.g. within QM/MM or QM-cluster models with additional water molecule networks) [24, 27]. For example, the TS1 activation energy ($\Delta\Delta G = 1.6$ kcal mol^{-1}) is lower than the ~7 kcal mol^{-1} activation energy reported for water-packed models, as the hydroxide and CO_2 substrate do not have to break through a 'wall' of waters to bind. The calculated value is instead comparable to the ~1 kcal mol^{-1} activation energy reported in QM-cluster models with fewer waters [19, 39]. The issue of how to properly ensure water networks are identified and modeled appropriately for solvent-exposed active sites remains a complicated topic, compounded by the mobility of waters and their typically poor resolution in x-ray crystal structures. Efforts to design QM-cluster models of solvent-accessible active sites that better account for the presence of waters through solvated MM or molecular dynamics (MD) simulations are underway by our lab. Apart from water network differences, the results reported here are comparable to those from reported 'larger' QM-cluster models, providing good evidence to support the rigor of this enzyme model building approach. Results with Zn^{2+} in the active site will serve as a reference of comparison against the other metal-substituted HCAII models.

3.2. Proposed mechanism for Ni^{2+}-HCAII

The Ni^{2+}-substituted metallovariant of HCAII was reported as having poor enzymatic activity compared to the activity of Zn^{2+}-coordinated HCAII. Unlike the native structure, an additional water molecule is observed to coordinate to Ni^{2+}, forming a square pyramidal starting geometry. This difference in metal and coordination alters the predicted Lipscomb- and Lindskog-like mechanisms (figure 4) and subsequent

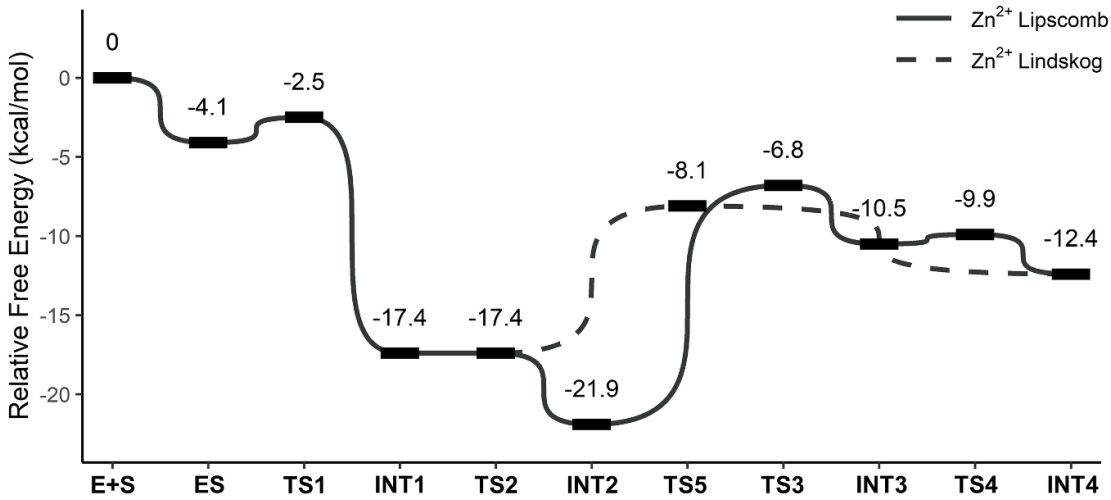

Figure 3. Free energy profile for the Lipscomb and Lindskog reaction mechanisms for the native Zn^{2+}-bound HCAII.

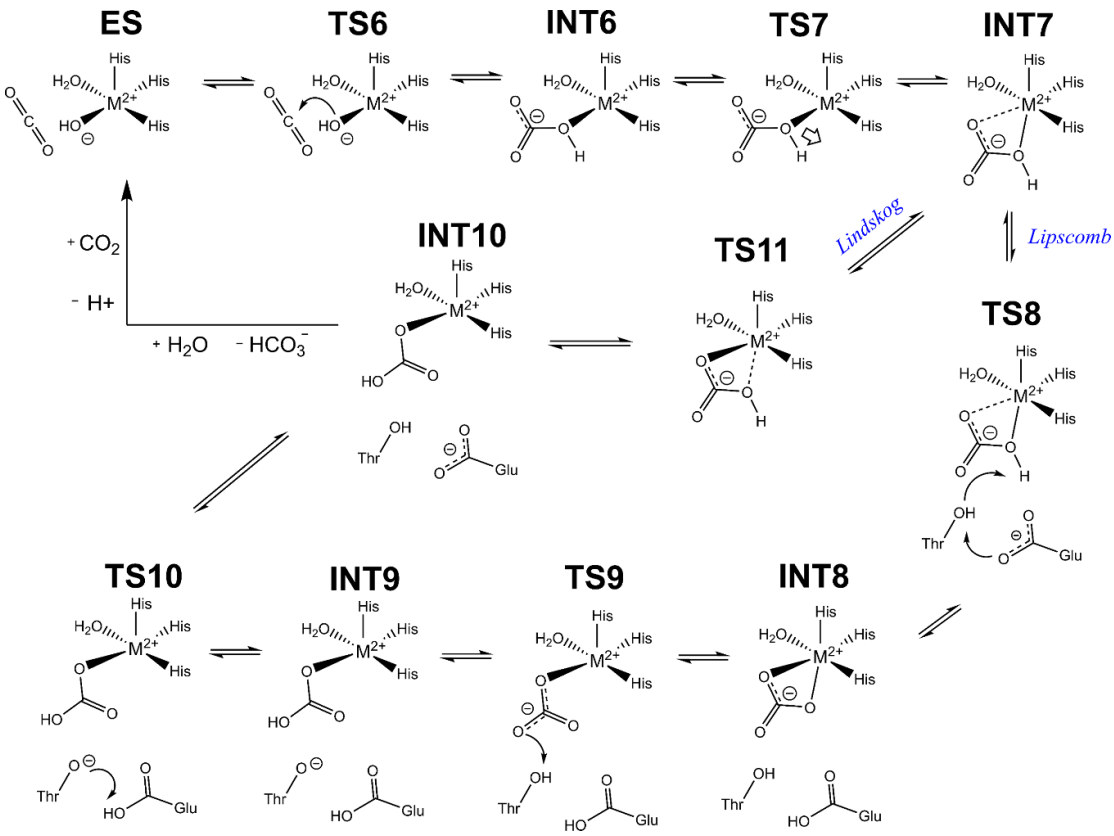

Figure 4. Proposed mechanism for carbon dioxide hydration by HCAII (M = high-spin Ni^{2+}, and low-spin Fe^{2+}).

reaction energy profiles (figure 5). There is a remarkable change in activation energy for the initial nucleophilic attack of the hydroxide to the carbon dioxide, increasing from 1.6 kcal mol^{-1} for Zn^{2+}-HCAII to 11.2 kcal mol^{-1} (TS6). This substantial increase in activation energy suggests the square pyramidal coordination geometry is not as conducive toward reaction initiation as the tetrahedral geometry. After reaction initiation, the Lipscomb-like path is computed to be preferred ($\Delta\Delta G = 7.3$ kcal mol^{-1}) though there is only a 0.8 kcal mol^{-1} difference between the two mechanisms. Lastly, it is noted that the transition state and intermediate structures are not as stabilized relative to the reactants, and the hydration of CO_2 mechanism is less exergonic within the Ni^{2+}-active site compared to the Zn^{2+} equivalent.

Previous experiments have highlighted the impact that changing metal-coordination geometry has on CO_2-hydration efficiency. X-ray crystal structures reveal hexacoordinate geometries for Ni^{2+}-HCAII, which

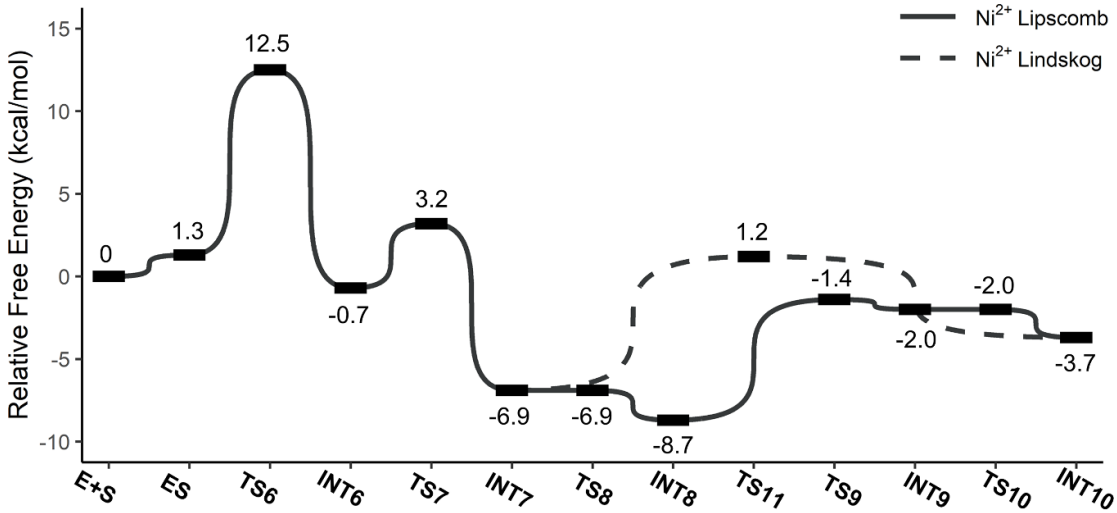

Figure 5. Free energy profiles for the Lipscomb and Lindskog reaction mechanisms for high-spin Ni^{2+}-bound HCAII.

will have to have one of their waters displaced for the reaction to commence [40, 41]. The non-tetrahedral geometry leads to steric hindrance between the CO_2 and the metal-bound waters as CO_2 enters the active site pocket, distorting the final orientation of CO_2 within the Ni^{2+}-HCAII cavity [40]. Although our water-sparse QM-cluster model is not designed to accurately capture the magnitude of steric effects, the QM-cluster model can validate a steric explanation by both the heightened energy barrier for the initial nucleophilic attack on CO_2 (TS6) and the enzyme-substrate complex being computed to be less stable within the Ni^{2+}-model than the separate enzyme and substrate species (E + S), a feature not observed with the Zn^{2+} model. Analysis of x-ray crystal structures and biomimetic catalysts also point toward reduced dissociation of bicarbonate from the metal as a factor for the poor activity of Ni^{2+}-HCAII [40, 64]. That step of the catalytic cycle was not modeled in this work but should be investigated in future studies.

3.3. Proposed mechanism for Cd^{2+}-HCAII

While Cd^{2+} and Zn^{2+} are in the same periodic group, Cd^{2+}-HCAII has a catalytic activity of only ~2% to that of Zn^{2+}-HCAII. The energy profile for the CO_2 hydration reaction (figure 6) was computed using the tetrahedral metal starting geometry, and intermediates and transition state structures similar to those for Zn^{2+}-HCAII were identified. The energy profile indicates the reaction is slightly more favorable thermodynamically and kinetically when catalyzed by Cd^{2+} than Zn^{2+}-HCAII. The Lindskog pathway is the energetically favored pathway with a relative activation free energy of 7.5 kcal mol^{-1} (TS5). Notably, the energy to break the Cd–O bond of the Lindskog path is 1.8 kcal mol^{-1} lower than the equivalent Zn–O bond breakage for the native enzyme. The monodentate carbonate intermediate present for the Zn^{2+}-HCAII Lipscomb path (INT3) is also absent from the Cd^{2+}-HCAII mechanism, as the proton transfer (TS4 with $\Delta\Delta G = 0.6$ kcal mol^{-1} in Zn^{2+}-HCAII) immediately proceeds to form the bicarbonate product.

Comparing the energy profile computed here to the profile reported by Toscano *et al* [28] using a 58-atom model, both profiles identify the reaction catalyzed by Cd^{2+}-HCAII as being more thermodynamically favorable than the Zn^{2+}-catalyzed reaction. However, while our model also suggests Cd^{2+}-HCAII to be slightly more kinetically favorable, the Toscano and coworkers model describes a kinetically unfavored reaction pathway characterized by highly stabilized intermediates separated by energy barriers of 35–70 kcal mol^{-1}. This discrepancy between the theoretical models is likely due to differences in QM-cluster model design, demonstrating that a more comprehensive treatment of the active site environment will qualitatively change proposed mechanisms of Cd^{2+}-HCAII [28] and other enzymes [65–68].

Given the similarity between Zn^{2+} and Cd^{2+} energy profiles, the findings suggest the experimental observation of poor catalytic activity for Cd^{2+}-HCAII is due to hindrance of other steps of the catalytic mechanism; specifically the regeneration of the metal-bound hydroxide step that occurs before CO_2 hydration may be impacted. In HCAII, after the bicarbonate product dissociates from the metal and is replaced with water (figure 2), the metal-bound water is deprotonated to a hydroxide by a histidine-directed water shuttle that drives the proton from within the protein pocket into the bulk solvent [69–73]. Given that Cd^{2+} is a soft Lewis acid compared to the borderline hard-soft Zn^{2+}, it may be that the Cd^{2+}-bound water is less reactive toward deprotonation to the harder hydroxide ligand than Zn^{2+}-bound water [74]. As a result,

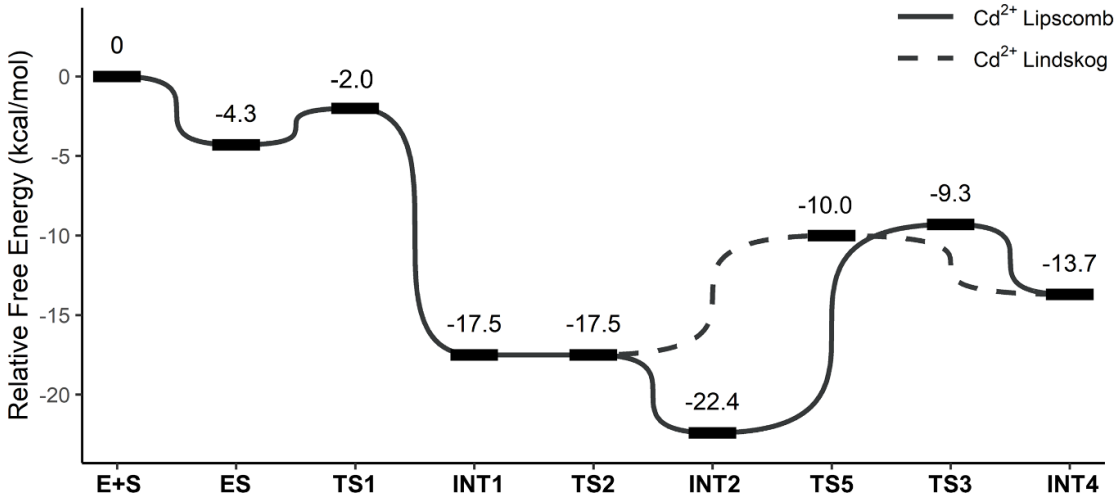

Figure 6. Free energy profiles for the Lipscomb (solid) and Lindskog (dashed) reaction mechanisms for Cd^{2+}-bound HCAII.

the predominant form of the enzyme in solution would be the water-bound Cd^{2+}-HCAII rather than the catalytically active hydroxide-bound Cd^{2+}-HCAII. There is experimental evidence that Cd^{2+}-bound human CA B isozyme becomes active at pHs higher than physiological pH, and the activity profile corresponds to the deprotonation of a Cd^{2+}-bound water molecule [75]. Furthermore, Cd^{2+} successfully serves as the catalytic metal for the ζ-class CA of diatoms [76]. In the ζ-class CA, Cd^{2+} is coordinated to only one histidine (an intermediate hard-soft base) and two (soft base) cysteines instead of the three histidines present in α-class HCAII. Theoretical models have noted that swapping histidine with cysteine affects the charges and polarizability of both metal and ligand [77]. Additional modeling to more closely investigate how the metal and its coordinating residues affect the water deprotonation shuttle mechanism is beyond the abilities of this QM-cluster model and outside the scope of this current work.

3.4. Proposed mechanism for Fe^{2+} and Fe^{3+}-HCAII

Low enzymatic activity was reported for Fe^{2+}-HCAII, though it is believed that this poor activity may be attributed to the aerobic environment oxidizing Fe^{2+} to Fe^{3+}, and the Fe^{3+} ion subsequently dissociating from the active site [12]. To investigate the potential CO_2 hydration mechanism for Fe^{2+}-HCAII within an anaerobic environment, the Fe^{2+}-HCAII QM-cluster model was simulated at both low and high-spin configurations. A study on the γ-class CAM enzyme in an anaerobic environment reported a high-spin state for the Fe^{2+} coordination center [12]. Given both α-class and γ-class CA active sites utilize a three histidine metal coordination, the Fe^{2+} is expected to likewise be in a high spin state within HCAII. Structures for the low-spin Fe^{2+}-HCAII were computed using the model with the additional metal-bound water as the metal ion adopted a square pyramidal starting geometry. The resulting Lipscomb- and Lindksog-like mechanisms (figures 4 and S2) and free energy profile (figure S3) are similar to the Ni^{2+}-HCAII pathways with several additional stationary points identified corresponding to intermediary ligand translation (figure S2(A)) and unbound water reorienting hydrogen bonds within in the active site (figure S2(B)). The Lindskog-like mechanism is the energetically favored pathway, and the reaction is slightly more thermodynamically favorable with low-spin Fe^{2+} compared to Ni^{2+}. However, as the reaction begins from the square pyramidal geometry, the low-spin Fe^{2+}-HCAII reaction faces a substantial activation energy barrier required for reaction initiation ($\Delta\Delta G = 13.7$ kcal mol^{-1}, figure S3) compared to reactant metal ions beginning from a tetrahedral geometry.

Unlike in the low-spin model, the high-spin Fe^{2+}-HCAII metal adopted a tetrahedral starting geometry and is more stable than the low-spin model by 4.2 kcal mol^{-1} (table S3), in agreement with experimental observation [12]. The proposed mechanisms (figure 2) and free energy profile (figure 7) follow a similar reaction path to the native Zn^{2+}-HCAII pathway. The intermediates are shown to be more thermodynamically stable, and the reaction is slightly more thermodynamically favorable than the Zn^{2+}-catalyzed pathway by 1.7 kcal mol^{-1}. High spin Fe^{2+}-HCAII also has a reduced activation energy required for the Lipscomb pathway (TS3; $\Delta\Delta G_{Zn} = 15.1$ kcal mol^{-1}, $\Delta\Delta G_{Fe} = 11.7$ kcal mol^{-1}), and a slightly increased activation energy required for the Lindskog pathway (TS5; $\Delta\Delta G_{Zn} = 9.3$ kcal mol^{-1}, $\Delta\Delta G_{Fe} = 10.6$ kcal mol^{-1}). Both metal active sites kinetically favor the Lindskog pathway over the Lipscomb pathway, although the difference is reduced for high spin Fe^{2+}-HCAII, making the two pathways more

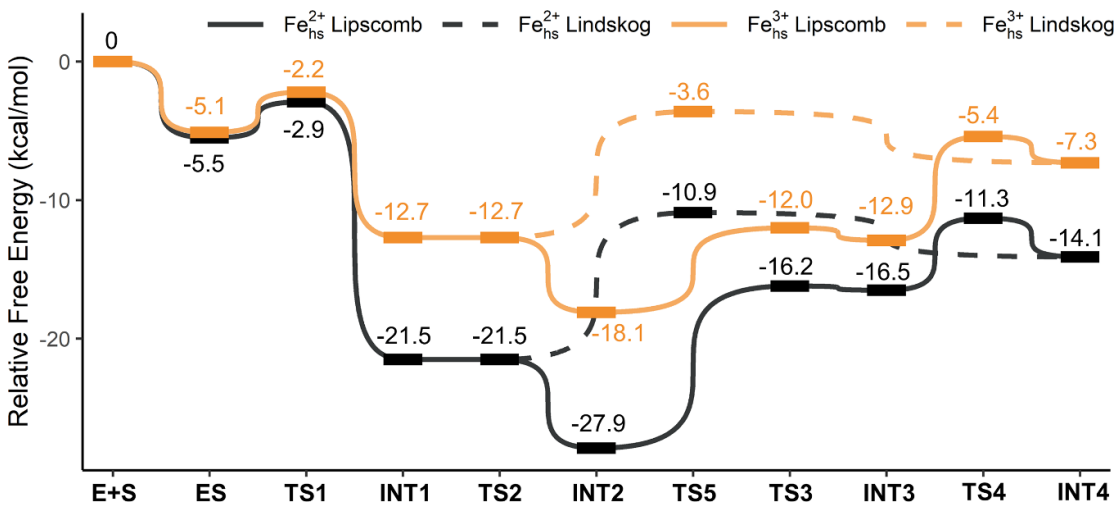

Figure 7. Free energy profiles for the Lipscomb and Lindskog reaction mechanisms for high-spin Fe^{2+} and Fe^{3+}-bound HCAII.

competitive. Based upon these results, the hydration of CO_2 is predicted to be thermodynamically and kinetically feasible for high spin Fe^{2+}-HCAII.

The models were also used to simulate the CO_2 hydration mechanism for both low and high-spin Fe^{3+}-HCAII. The metal in both spin configurations is arranged in a tetrahedral geometry, and, as seen in experiment [12], the high-spin state is more stable (table S3). The reaction mechanism (figure 2) and free energy profiles (figures S4) for the low-spin Fe^{3+}-HCAII mirrored the high-spin Fe^{2+}-HCAII with the exception of the very thermostable Lipscomb intermediate INT2. This increases the effective activation free energy for the Lipscomb reaction pathway to 21.1 kcal mol^{-1}; the Lindskog pathway thus remains kinetically favored with a $\Delta\Delta G = 11.5$ kcal mol^{-1}. For the high-spin Fe^{3+}-HCAII model, the reaction path thermodynamics are comparable to both the native Zn^{2+} and high-spin Fe^{2+}-HCAII catalyzed pathways, though the intermediates are less thermodynamically stable and the reaction is slightly less favorable than both. The activation energy for the Lipscomb pathway is notably smaller than that for the Lindskog pathway by 1.6 kcal mol^{-1}, making it both the thermodynamically and kinetically favored path. These results support the feasibility of Fe^{3+}-HCAII to catalyze the CO_2 hydration reaction. However, this feasibility assumes the starting metal-bound hydroxide structure is readily formed, and experiments report poor metal binding affinity for Fe^{3+} to HCAII [12, 78].

4. Conclusions

In this work, the feasibility for multiple transition metal-substituted HCAII enzymes to catalyze CO_2 hydration were investigated using QM-cluster models designed from RINs. Two reaction mechanisms were considered, the Lipscomb path involving a dual proton transfer between the bicarbonate and nearby Thr and Glu residues and the Lindskog path involving a bicarbonate metal-oxygen bond breakage and rotation. The free energy profile for the native Zn^{2+}-HCAII was first computed, and the results agreed with several previous theoretical studies finding that while both Lipscomb and Lindskog reaction paths are feasible, the Lindskog path is more energetically favored. When the Zn^{2+} ion was substituted with Ni^{2+}, an additional water binds to the metal that leads to a square pyramidal starting geometry rather than a tetrahedral geometry. This change in the coordination chemistry substantially increases the activation energy required for the hydroxide to bind to CO_2 and initiate the reaction. The proposed mechanisms for Ni^{2+} substitutions are also less thermodynamically favored compared to the Zn^{2+}-catalyzed reactions. The CO_2 hydration mechanism catalyzed by Cd^{2+}-HCAII is shown to be comparable to Zn^{2+} even though experiments report generally poor catalytic activity. This discrepancy may arise from Cd^{2+}-HCAII not readily deprotonating the metal-bound water to form the starting metal-bound hydroxide, which is the step in the catalytic cycle previous to those modeled in this work.

Lastly, the mechanisms for both Fe^{2+}- and Fe^{3+}-HCAII in low and high spin states were examined. The high spin state Fe^{2+} is expected to be the predominant catalytic form, and the reaction pathways computed are thermodynamically more favored than Zn^{2+}-HCAII along with being kinetically comparable. These results suggest that, in an anaerobic environment where the Fe^{2+} is not able to be oxidized to Fe^{3+}, the hydration of CO_2 by Fe^{2+}-HCAII is theoretically feasible. The scope of these findings are notably limited to

the modeled CO_2 hydration steps of the catalytic cycle, as other steps in the catalytic cycle (e.g. regeneration of the metal-hydroxide or dissociation of the product) may still inhibit or reduce the drive of the reaction. Nevertheless, these results give hope toward the likelihood of synthesizing an active, anaerobic Fe^{2+}-HCAII. When the Fe^{2+} is oxidized to Fe^{3+}, QM-cluster models suggest CO_2 hydration is still kinetically fast, but the poor experimental binding affinity of HCAII to Fe^{3+} will lead to dissociation of the metal and loss of catalytic activity.

Collectively, these QM-cluster models provide atomic-level insight into the changing CO_2 hydration reaction profiles for different HCAII metallovariants. Further investigations into metallovariant proton shuttle and bicarbonate dissociation pathways will improve the knowledge of the overall enzymatic function and help bioengineer enhanced CAs and guide the development of covalent inhibitors to CA.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary information files).

Acknowledgments

The authors thank Professor Joseph Emerson (Mississippi State University), Professor Abby Parrill-Baker (University of Memphis), and Dr Qianyi Cheng (University of Memphis) for inspiration and helpful discussions for this work. This material is based upon work supported by the National Science Foundation (NSF) Graduate Research Fellowship Program under Grant Nos. 1451514, and NSF CAREER BIO-1846408. This work was also supported by start-up funding from the University of Memphis. This support does not necessarily imply endorsement by the University of research conclusions. The authors also thank the High Performance Computing Center and the Computational Research on Materials Institute at the University of Memphis (CROMIUM) for providing generous resources for this research.

ORCID iDs

Thomas J Summers https://orcid.org/0000-0002-4243-6078
Nathan J DeYonker https://orcid.org/0000-0003-0435-2006

References

[1] Supuran C T 2016 Structure and function of carbonic anhydrases *Biochem. J.* **473** 2023–32
[2] Supuran C T 2018 Carbonic anhydrases and metabolism *Metabolites* **8** 25
[3] Krishnamurthy V M, Kaufman G K, Urbach A R, Gitlin I, Gudiksen K L, Weibel D B and Whitesides G M 2008 Carbonic anhydrase as a model for biophysical and physical-organic studies of proteins and protein—ligand binding *Chem. Rev.* **108** 946–1051
[4] Aggarwal M, Boone C D, Kondeti B and McKenna R 2013 Structural annotation of human carbonic anhydrases *J. Enzyme Inhib. Med. Chem.* **28** 267–77
[5] Eriksson A E, Jones T A and Liljas A 1988 Refined structure of human carbonic anhydrase II at 2.0 Å resolution *Proteins Struct. Funct. Bioinform.* **4** 274–82
[6] Lindskog S 1997 Structure and mechanism of carbonic anhydrase *Pharmacol. Ther.* **74** 1–20
[7] Hurst T K, Wang D, Thompson R B and Fierke C A 2010 Carbonic anhydrase II-based metal ion sensing: advances and new perspectives *Biochim. Biophys. Acta* **1804** 393–403
[8] McCall K A and Fierke C A 2004 Probing determinants of the metal ion selectivity in carbonic anhydrase using mutagenesis *Biochemistry* **43** 3979–86
[9] Lionetto M G, Caricato R, Giordano M E and Schettino T 2016 The complex relationship between metals and carbonic anhydrase: new insights and perspectives *Int. J. Mol. Sci.* **17** 127
[10] Kogut K and Rowlett R 1987 A comparison of the mechanisms of CO_2 hydration by native and Co^{2+}-substituted carbonic anhydrase II *J. Biol. Chem.* **262** 16417–24
[11] Bertini I and Luchinat C 1994 The reaction pathways of zinc enzymes and related biological catalysts *Bioinorganic Chemistry* (Mill Valley, CA: University Science Books)
[12] Tripp B C, Bell C B, Cruz F, Krebs C and Ferry J G 2004 A role for iron in an ancient carbonic anhydrase* *J. Biol. Chem.* **279** 6683–7
[13] MacAuley S R, Zimmerman S A, Apolinario E E, Evilia C, Hou Y-M, Ferry J G and Sowers K R 2009 The archetype γ-class carbonic anhydrase (cam) contains iron when synthesized *in vivo Biochemistry* **48** 817–9
[14] Ferry J G 2010 The γ class of carbonic anhydrases *Biochim. Biophys. Acta* **1804** 374–81
[15] Wu Y, Zhao X, Li P and Huang H 2007 Impact of Zn, Cu, and Fe on the activity of carbonic anhydrase of erythrocytes in ducks *Biol. Trace Elem. Res.* **118** 227–32
[16] Merz K M, Hoffmann R and Dewar M J S 1989 The mode of action of carbonic anhydrase *J. Am. Chem. Soc.* **111** 5636–49
[17] Silverman D N and Lindskog S 1988 The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water *Acc. Chem. Res.* **21** 30–36
[18] Lipscomb W N 1983 Structure and catalysis of enzymes *Annu. Rev. Biochem.* **52** 17–34
[19] Bottoni A, Lanza C Z, Miscione G P and Spinelli D 2004 New model for a theoretical density functional theory investigation of the mechanism of the carbonic anhydrase: how does the internal bicarbonate rearrangement occur? *J. Am. Chem. Soc.* **126** 1542–50

[20] Toba S, Colombo G and Merz K M 1999 Solvent dynamics and mechanism of proton transfer in human carbonic anhydrase II *J. Am. Chem. Soc.* **121** 2290–302
[21] Riccardi D, Yang S and Cui Q 2010 Proton transfer function of carbonic anhydrase: insights from QM/MM simulations *Biochim. Biophys. Acta* **1804** 342–51
[22] Maupin C M and Voth G A 2010 Proton transport in carbonic anhydrase: insights from molecular simulation *Biochim. Biophys. Acta* **1804** 332
[23] Piazzetta P, Marino T and Russo N 2015 Theoretical investigation on the restoring step of the carbonic anhydrase catalytic cycle for natural and promiscuous substrates *Arch. Biochem. Biophys.* **582** 101–6
[24] Miscione G P, Stenta M, Spinelli D, Anders E and Bottoni A 2007 New computational evidence for the catalytic mechanism of carbonic anhydrase *Theor. Chem. Acc.* **118** 193–201
[25] Liang J-Y and Lipscomb W N 1989 Theoretical study of carbonic anhydrase-catalyzed hydration of Co_2: a brief review *Int. J. Quantum Chem.* **36** 299–312
[26] Sola M, Lledos A, Duran M and Bertran J 1992 *Ab initio* study of the hydration of carbon dioxide by carbonic anhydrase. A comparison between the Lipscomb and Lindskog mechanisms *J. Am. Chem. Soc.* **114** 869–77
[27] Piazzetta P, Marino T, Russo N and Salahub D R 2017 The role of metal substitution in the promiscuity of natural and artificial carbonic anhydrases *Coord. Chem. Rev.* **345** 73–85
[28] Marino T, Russo N and Toscano M 2005 A comparative study of the catalytic mechanisms of the zinc and cadmium containing carbonic anhydrase *J. Am. Chem. Soc.* **127** 4242–53
[29] Bertini I and Luchinat C 1983 Cobalt(II) as a probe of the structure and function of carbonic anhydrase *Acc. Chem. Res.* **16** 272–9
[30] Woolley P 1975 Models for metal ion function in carbonic anhydrase *Nature* **258** 677–82
[31] Håkansson K, Carlsson M, Svensson L A and Liljas A 1992 Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes *J. Mol. Biol.* **227** 1192–204
[32] Liang J Y and Lipscomb W N 1986 Theoretical study of the uncatalyzed hydration of carbon dioxide in the gas phase *J. Am. Chem. Soc.* **108** 5051–8
[33] Loferer M J, Tautermann C S, Loeffler H H and Liedl K R 2003 Influence of backbone conformations of human carbonic anhydrase II on carbon dioxide hydration: hydration pathways and binding of bicarbonate *J. Am. Chem. Soc.* **125** 8921–7
[34] Tautermann C S, Loferer M J, Voegele A F and Liedl K R 2003 About the kinetic feasibility of the Lipscomb mechanism in human carbonic anhydrase II *J. Phys. Chem.* B **107** 12013–20
[35] Hartmann M, Merz K M, van Eldik R and Clark T 1998 The important role of active site water in the catalytic mechanism of human carbonic anhydrase II—A semiempirical MO approach to the hydration of CO_2† *Mol. Model. Annu.* **4** 355–65
[36] Zheng Y J and Merz K M 1992 Mechanism of the human carbonic anhydrase II-catalyzed hydration of carbon dioxide *J. Am. Chem. Soc.* **114** 10498–507
[37] Ghiasi M, Gholami S and Nasiri S 2021 QM study of carbon dioxide (CO_2) and carbonyl sulfide (COS) degradation by cluster model of carbonic anhydrase enzyme *Comput. Theor. Chem.* **1199** 113188
[38] Ghiasi M, Goli N E, Gholami S and Supuran C T 2021 QM and QM/MM study on inhibition mechanism of polyphenolic compounds as non-classical inhibitors of α-human carbonic anhydrase (II) *Theor. Chem. Acc.* **140** 141
[39] Piazzetta P, Marino T and Russo N 2014 Promiscuous ability of human carbonic anhydrase: QM and QM/MM investigation of carbon dioxide and carbodiimide hydration *Inorg. Chem.* **53** 3488–93
[40] Kim J K, Lee C, Lim S W, Adhikari A, Andring J T, McKenna R, Ghim C-M and Kim C U 2020 Elucidating the role of metal ions in carbonic anhydrase catalysis *Nat. Commun.* **11** 4557
[41] Håkansson K, Wehnert A and Liljas A 1994 X-ray analysis of metal-substituted human carbonic anhydrase II derivatives *Acta Crystallogr.* D **50** 93–100
[42] Di Paola L, De Ruvo M, Paci P, Santoni D and Giuliani A 2013 Protein contact networks: an emerging paradigm in chemistry *Chem. Rev.* **113** 1598–613
[43] Doncheva N T, Assenov Y, Domingues F S and Albrecht M 2012 Topological analysis and interactive visualization of biological networks and protein structures *Nat. Protocols* **7** 670–85
[44] Avvaru B S, Arenas D J, Tu C, Tanner D B, McKenna R and Silverman D N 2010 Comparison of solution and crystal properties of Co(II)–substituted human carbonic anhydrase II *Arch. Biochem. Biophys.* **502** 53–59
[45] Word J M, Lovell S C, Richardson J S and Richardson D C 1999 Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation *J. Mol. Biol.* **285** 1735–47
[46] Fisher S Z, Aggarwal M, Kovalevsky A Y, Silverman D N and McKenna R 2012 Neutron diffraction of acetazolamide-bound human carbonic anhydrase II reveals atomic details of drug binding *J. Am. Chem. Soc.* **134** 14726–9
[47] Domsic J F, Avvaru B S, Kim C U, Gruner S M, Agbandje-mckenna M, Silverman D N and McKenna R 2008 Entrapment of carbon dioxide in the active site of carbonic anhydrase II *J. Biol. Chem.* **283** 30766–71
[48] Håkansson K and Wehnert A 1992 Structure of cobalt carbonic anhydrase complexed with bicarbonate *J. Mol. Biol.* **228** 1212–8
[49] Word J M, Lovell S C, LaBean T H, Taylor H C, Zalis M E, Presley B K, Richardson J S and Richardson D C 1999 Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms *J. Mol. Biol.* **285** 1711–33
[50] Cheng Q, DeYonker N J, Summers T J, Agbaglo D A, Suhagia T and Berquist E 2022 *RINRUS*, GitHub repository (available at: https://github.com/natedey/RINRUS)
[51] Kitajima N, Hikichi S, Tanaka M and Morooka Y 1993 Fixation of atmospheric carbon dioxide by a series of hydroxo complexes of divalent metal ions and the implication for the catalytic role of metal ion in carbonic anhydrase. Synthesis, characterization, and molecular structure of [LM(OH)]n (n = 1 or 2) and LM(.Mu.-CO_3)ML (M(II) = Mn, Fe, Co, Ni, Cu, Zn; L = HB(3,5-Iso-Pr2pz)3) *J. Am. Chem. Soc.* **115** 5496–508
[52] Ragsdale S W 2009 Nickel-based enzyme systems *J. Biol. Chem.* **284** 18571–5
[53] Clark P A and Wilcox D E 1989 Magnetic properties of the nickel enzymes urease, nickel-substituted carboxypeptidase A and nickel-substituted carbonic anhydrase *Inorg. Chem.* **28** 1326–33
[54] Frisch M J *et al* 2016 *Gaussian 16* (Wallingford, CT: Gaussian, Inc)
[55] Becke A D 1993 Density-functional thermochemistry III. The role of exact exchange *J. Chem. Phys.* **98** 5648–52
[56] Lee C, Yang W and Parr R G 1988 Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density *Phys. Rev.* B **37** 785–9
[57] Petersson G A and Al-Laham M A 1991 A complete basis set model chemistry. II. Open-shell systems and the total energies of the first-row atoms *J. Chem. Phys.* **94** 6081–90

[58] Hehre W J, Ditchfield R and Pople J A 1972 Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules *J. Chem. Phys.* **56** 2257–61
[59] Couty M and Hall M B 1996 Basis sets for transition metals: optimized outer p functions *J. Comput. Chem.* **17** 1359–70
[60] Wadt W R and Hay P J 1985 *Ab initio* effective core potentials for molecular calculations. Potentials for main group elements Na to Bi *J. Chem. Phys.* **82** 284–98
[61] Grimme S, Ehrlich S and Goerigk L 2011 Effect of the damping function in dispersion corrected density functional theory *J. Comput. Chem.* **32** 1456–65
[62] Barone V and Cossi M 1998 Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model *J. Phys. Chem.* A **102** 1995–2001
[63] Cossi M, Rega N, Scalmani G and Barone V 2003 Energies structures, and electronic properties of molecules in solution with the C-PCM solvation model *J. Comput. Chem.* **24** 669–81
[64] Keum C, Kim M-C and Lee S-Y 2015 Effects of transition metal ions on the catalytic activity of carbonic anhydrase mimics *J. Mol. Catal. Chem.* **408** 69–74
[65] Liao R-Z and Thiel W 2013 Convergence in the QM-only and QM/MM modeling of enzymatic reactions: a case study for acetylene hydratase *J. Comput. Chem.* **34** 2389–97
[66] Kulik H J, Zhang J, Klinman J P and Martínez T J 2016 How large should the QM region be in QM/MM calculations? The case of catechol O-methyltransferase *J. Phys. Chem.* B **120** 11381–94
[67] Hu L, Eliasson J, Heimdal J and Ryde U 2009 Do quantum mechanical energies calculated for small models of protein-active sites converge? *J. Phys. Chem.* A **113** 11793–800
[68] Summers T J, Cheng Q, Palma M A, Pham D-T, Kelso D K, Webster C E and DeYonker N J 2021 Cheminformatic quantum mechanical enzyme model design: a catechol-o-methyltransferase case study *Biophys. J.* **120** 3577–87
[69] Cui Q and Karplus M 2003 Is a proton wire concerted or stepwise? A model study of proton transfer in carbonic anhydrase *J. Phys. Chem.* B **107** 1071–8
[70] Mikulski R, West D, Sippel K H, Avvaru B S, Aggarwal M, Tu C, McKenna R and Silverman D N 2013 Water networks in fast proton transfer during catalysis by human carbonic anhydrase II *Biochemistry* **52** 125–31
[71] Jiao D and Rempe S B 2012 Combined density functional theory (DFT) and continuum calculations of pKa in carbonic anhydrase *Biochemistry* **51** 5979–89
[72] Hakkim V and Subramanian V 2010 Role of second coordination sphere amino acid residues on the proton transfer mechanism of human carbonic anhydrase II (HCA II) *J. Phys. Chem.* A **114** 7952–9
[73] Maupin C M, McKenna R, Silverman D N and Voth G A 2009 Elucidation of the proton transport mechanism in human carbonic anhydrase II *J. Am. Chem. Soc.* **131** 7598–608
[74] Andreini C, Bertini I, Cavallaro G, Holliday G L and Thornton J M 2008 Metal ions in biological catalysis: from enzyme databases to general principles *J. Biol. Inorg. Chem.* **13** 1205–18
[75] Bauer R, Limkilde P and Johansen J T 1976 Low and high PH form of cadmium carbonic anhydrase determined by nuclear quadrupole interaction *Biochemistry* **15** 334–42
[76] Alterio V, Langella E, De Simone G and Monti S M 2015 Cadmium-containing carbonic anhydrase CDCA1 in marine diatom Thalassiosira weissflogii *Mar. Drugs* **13** 1688–97
[77] Amata O, Marino T, Russo N and Toscano M 2011 Catalytic activity of a ζ-class zinc and cadmium containing carbonic anhydrase. Compared work mechanisms *Phys. Chem. Chem. Phys.* **13** 3468–77
[78] Lindskog S and Nyman P O 1964 Metal-binding properties of human erythrocyte carbonic anhydrases *Biochim. Biophys. Acta* **85** 462–74