Density functional theory investigations of the catalytic mechanism of β-carbonic anhydrase

V Hakkim, V Rajapandian & V Subramanian*
Chemical Laboratory, Central Leather Research Institute, CSIR, Adyar, Chennai 600 020, India
Email: subuchem@hotmail.com/subbu@clri.res.in

Received 18 November 2010; accepted 1 December 2010

Carbonic anhydrase (CA) is an enzyme that catalyses the reversible hydration of carbon dioxide. There are three broad classes α, β, and γ, of CA, divided into three genetically unrelated families, namely, animal, plant, and bacterial CAs, respectively. The active site of this enzyme contains a zinc atom which is necessary for catalysis. In this study, the catalytic mechanism of β-CA has been investigated using its active site model employing DFT based Becke’s three parameter exchange and B3LYP method. It is evident from the results that the activation barrier for the nucleophilic attack is negligible, which is similar to that of α-CA. Furthermore, results show that Asp162-Arg164 dyad and Glu151 residues play a decisive role in the catalysis. Primarily, the catalytic dyad orients the hydroxyl group appropriately to enable nucleophilic attack and stabilizes the negative charge on the bicarbonate.

Keywords: Bioinorganic chemistry, Theoretical chemistry, Catalysis, Reaction mechanisms, Metalloenzymes, Carbonic anhydrase

Studies on metalloenzymes are an important part of structural biology and biochemistry.14 A great deal of literature is available on the structure and mechanistic pathway of catalysis of various enzymes; for example, photosystem II of photosynthesis, cytochrome c oxidase in the respiratory chain, ribonucleotide reductase in DNA synthesis, methane monoxygenase for converting alkanes to alcohols, superoxide dismutase for the conversion of superoxide into oxygen and hydrogen peroxide, and amino peptidases for hydrolysis of peptides have been investigated.5-11 In addition to the sophisticated experimental studies, the potential applications of computer modeling of enzyme reaction mechanism have been illustrated in several studies.12-18

The field of computer modeling of enzymes using molecular mechanics (MM), quantum mechanics (QM), and combined quantum mechanics and molecular mechanics (QM/MM), is known as computational enzymology. This field has adequately matured to reveal how enzymes work.19,20 The success of probing structure and dynamics of enzyme has been illustrated in numerous studies by applying MM based molecular dynamics method.21-24 The QM modeling of the active site provides crucial information on the role of active site in the enzyme reaction mechanism.25 Siegbahn and co-workers26-30 have made seminal contributions to the active site modeling of metalloenzymes. In QM modeling of enzyme active sites, the DFT based B3LYP has been shown to be as a promising method to explore reaction mechanisms and to provide insights into various aspects of catalysis.31-37 Warshel and Levitt38 were the first to investigate the enzyme reaction mechanism using computational approach by combining QM treatment for the active site and MM description for remaining part of the enzymes. Subsequently, several studies have been made to understand the enzyme reaction mechanism using QM/MM methods.39-44 Recently, the QM/MM modeling of enzymes has been reviewed.45,46 This approach can be used to understand and complement results obtained from experiments. Particularly, the atomic level details derived from these studies can be exploited to design new enzymes (new catalysts) and new drug molecules. Furthermore, computational studies on metalloenzymes are useful in engineering and redesigning of active site and to assess the effect of mutation.47,48

Chemical activation of carbon dioxide can help to reduce its concentration in the atmosphere. Transition metal complexes, especially of copper and zinc, as
well as simple salts such as lithium hydroxide monohydrate and sodalime (mixture of sodium and calcium hydroxides) are well known for their assistance in the stoichiometric transformation of carbon dioxide to carbonate salts.  

β-CA was discovered as a constituent of plant leaf chloroplasts. In addition to the photosynthetic organism, β-CA has been found in *Escherichia coli*, eubacteria, yeast, thermophilic archaeote, *Methanobacterium thermoautotrophicum* and archaeal species. The presence of β-CA in algae has also been reported. Similar to other CAs, this enzyme catalyzes the rapid interconversion of CO$_2$ and H$_2$O to HCO$_3^-$ and H$^+$. It assembles as an octamer with a novel dimer of dimers of dimers arrangement as shown in the Fig. 1. In the active site, zinc ion is located at the interface between two monomers. In contrast to α-CA, the zinc ion is coordinated to Cys160, His220, and Cys223 residues. In α-CA, the zinc ion is coordinated to the His94, His96 and His119 residues, respectively. As shown in the closer view of the active site model in the Fig. 1, Asp162-Arg164 dyad and Gln151 residue are found in the immediate vicinity of the active site. Several experimental studies have been performed on both α- and β-CAs. Although, a number of computational modeling studies have been carried out on the structure and functions of α-CA, similar investigations on the β-CA are limited. Based on the crystal structure and kinetic studies, the minimalistic model of the enzyme reaction is presented in Fig. 2. Previous studies have shown that β-CA is also a fast enzyme with $k_{cat} = 4 \times 10^5$ s$^{-1}$ and $k_{cat}/K_m = 1.8 \times 10^8$ M$^{-1}$s$^{-1}$ for the hydration of carbon dioxide. Since the substrate binding group of β-CA is similar to that of α-CA
active site, both \( \alpha \)- and \( \beta \)-CA are likely to share a common mechanism.\(^{67}\) However, the residues present in the active site are different and hence, it is logical to expect changes in the mechanistic pathway and energetics of \( \beta \)-CA.

Thus, a systematic computational modeling study has been carried out using DFT(B3LYP) method to investigate the hydration of carbon dioxide by employing the active site (cluster model) of \( \beta \)-CA. In this study the structure and energies of (i) formation of encounter complex between active site and substrate \((S)\) (carbon dioxide), (ii) nucleophilic addition of zinc bound hydroxyl to carbon of \( \text{CO}_2 \), (iii) formation of bicarbonate and (iv) exchange of bicarbonate by water at the active site, i.e., regeneration of the active site, have been investigated.

**Computational details and design of active site model**

The model of the enzyme active site was constructed using the X-ray structure (Protein Data Bank (PDB) ID 1EKJ).\(^{58}\) The active site model contains 54 atoms. Since crystal structure does not have water molecules at the active site, two water molecules were added using Gauss View package.\(^{76}\) In the active site, the \( \text{Zn}^{2+} \) ion is coordinated to His220, Cys160, and Cys223, Gln151, Asp162-Arg164 residues and a water molecule. In the current model, the histidine and cystine, aspartate residues were represented as imidazole and methyl thiolate and carboxylate respectively. The total charge and spin multiplicity of the model system are \(-1\) and \(1\) respectively.

As B3LYP method has shown to be a promising tool in exploring the reaction mechanism of transition metal containing enzymes, the geometries of reactants, intermediates, transition states, and products were optimized without any symmetry constraints at the same level of theory using 6-31G* basis set. The active site (reactant), encounter complex, transition state, intermediate and product are designated as \( \text{R, EC, TS, IM and P} \), respectively in the remaining part of the text. Some of the residues which do not participate in the reaction are shaded in Fig. 3. These residues were fixed during the geometry optimization, similar to that of Bottoni \textit{et al.}\(^{75}\) Transition state was analysed by means of frequency calculation. Single point calculations were carried out on all the geometries at B3LYP/6-311+G** level. All calculations were performed using the Gaussian 03 program package.\(^{77}\)

**Results and Discussion**

In this section, a detailed examination of the structures and energetics of hydration of carbon dioxide by the active site model of \( \beta \)-CA is presented.

**Geometrical parameters of the active site**

The optimized geometry of active site \((\text{R})\) is shown in Fig. 4a. The calculated geometrical parameters of \( \text{R} \) are presented in Table 1 along with the experimental values. Overall calculated geometrical parameters are in close agreement with those obtained from the X-ray diffraction. However, deviations are also observed from the calculated geometrical parameters. The O5-N2 and O6-N3 distances are marginally shorter than the corresponding X-ray diffraction structure values. Strong hydrogen bonding network is observed in the active site model. The calculated geometrical parameters reveal that zinc bound hydroxyl ion hydrogen bonds with the oxygen atom of Asp162 residue and water molecules. In addition, hydrogen bonding between the amine hydrogen of Gln151 and water molecule can also be observed.

**Nucleophilic attack**

Optimized geometry of \( \text{R, EC, TS, IM, P} \) is presented in Fig. 4. Crystallographic studies show that \( \text{CO}_2 \) binds to the hydrophobic residues in the active site cleft.\(^{54}\) The first step of the catalysis is the formation of encounter complex between the \( \text{CO}_2 \) with active site of \( \beta \)-CA. The optimized geometry of
EC is given in Fig. 4b. During the formation of EC, the lengthening of the hydrogen bond between the water molecule and N-H moiety of Gln151 residue has been observed. Typically, the O4···H(Gln151) distance in the R and EC is 1.96 and 2.04 Å, respectively. As a result, the water molecule moves away from the Gln151 residue.

Moreover, scrutiny of geometrical parameters of R and EC reveals that marginal increase in Zn-O1 distance is observed in the EC (Zn···O1 distance in R and EC is 1.89 and 1.90, respectively). Similarly, the hydrogen bond distance between H(O1) and O5 in EC decreases by 0.02 Å when compared to that of R. A considerable decrease in the hydrogen bond distance between O1 and H(O4) has also been noted in EC (O1···H(O4) is 1.61 Å in R and 1.51 Å in EC. In the EC, CO2 weakly interacts with O1 (C1···O1 is 4.02 Å), Further, it is anchored by water and –NH2 group of Gln151 (C1···O4 is 2.51 and O3···H(N1) is 2.42 Å).

The optimized structure of the TS formed in the catalytic process is depicted in Fig. 4c, along with the optimized structures of R and IM. Table 1—Geometrical parameters of the active site

<table>
<thead>
<tr>
<th>Bond distances (Å)</th>
<th>Crystal †</th>
<th>Calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zn-SG(C160)</td>
<td>2.29</td>
<td>2.38</td>
</tr>
<tr>
<td>Zn-SG(C223)</td>
<td>2.28</td>
<td>2.36</td>
</tr>
<tr>
<td>Zn-Nε(H220)</td>
<td>2.08</td>
<td>2.22</td>
</tr>
<tr>
<td>Zn-O(W)</td>
<td>-</td>
<td>1.89</td>
</tr>
<tr>
<td>O5-N2</td>
<td>3.06</td>
<td>2.81</td>
</tr>
<tr>
<td>O6-N3</td>
<td>2.78</td>
<td>2.61</td>
</tr>
</tbody>
</table>

† Ref. 58
Important distances. A significant difference in the Zn-O distance can be found from the Fig. 4c when compared to that of R and EC. Typically, the Zn-O distance in TS increases by 0.06 and 0.05 Å with respect to that of R and EC, respectively. The O1···C1 distance in TS is 2.98 Å. A marginal increase in distance between C and O in CO\textsubscript{2} in TS indicates the activation of CO\textsubscript{2}. The hydrogen bonding network in EC undergoes substantial changes in the formation of TS. The hydrogen bond distances Gln151(H(O1)···O5, O4···H(N1), O5···H(N2) and O6···H(N3) in EC are 1.80, 2.04, 1.51 and 1.74 Å, respectively. The corresponding distances in TS are 1.69, 1.90, 1.45 and 1.69 Å respectively.

The optimized geometry of the reaction intermediate (IM) formed during the catalysis is shown in Fig. 4d. Comparison of the geometries of TS and IM reveals that Zn-O1 distance increases by a magnitude of 0.12 Å. The presence of strong bond between O1 and C1 (C1···O1 is 1.55Å) shows the formation of bicarbonate in the IM. In addition, the water molecule forms a hydrogen bond with H(N1) of Gln151. It can be noted from the same that one of the Zn-O distance is shorter than the other and hence, the newly formed bicarbonate behaves as mono-dentate ligand. The negative charge on the bicarbonate is balanced by the imidazole ring.

The energy profiles obtained for various steps of the catalytic process for the active site model of β-CA is presented in Fig. 5. It can be seen that EC and TS are more stable by 5.78 and is 5.31 kcal/mol respectively, when compared to the reactant. Thus, the nucleophilic attack of the hydroxide on the CO\textsubscript{2} carbon requires only a very low energy barrier (~0.46 kcal/mol). Furthermore, the IM is more stable by a magnitude of 14.67 kcal/mol than the reactant.

**Exchange of bicarbonate by water molecule**

The optimized geometry of the P is shown in Fig. 4e. It may be noted that water molecule in the active site replaces the bicarbonate from zinc and regenerates the active site. The P is more stable than the reactants (R and S) by a magnitude of 36.12 kcal/mol. In P, negative charge on the bicarbonate is stabilized by the Asp162-Arg164 catalytic dyad which is in accordance with the previous experimental findings. In this investigation several attempts have been made to locate TS between the formation of P from IM. However, we could not locate any TS between these steps in the catalytic process. From this active site, the proton from zinc bound water molecule transfers to the external buffer through water mediated bridges, Tyr and His, which regenerates the active site for the subsequent catalytic cycle.

**Role of Asp162-Arg164 dyad and Gln151 in catalysis**

In the case of β-CA, Asp162-Arg164 dyad plays an important role in the catalytic cycle. A similar kind of catalytic dyad is also present in the cadmium containing CA from marine species. These dyads stabilize the charge concentrations, provide an ion pair to polarize third residue, give charges for the transition state stabilization, provide nucleophiles and participate in acid–base catalysis. With a view to assess the interplay of the Asp162-Arg164 dyad and Gln151 residue, the geometry of the active site was optimized without these residues. It can be seen from Fig. 6 that in the absence of these residues, the zinc bound hydroxyl group orients freely due to the absence of any steric field. As evident from the various experimental studies, a specific orientation of this hydroxyl group is necessary for the nucleophilic attack. This evidence clearly reveals the part played by the dyad in the catalysis.

Optimized geometry of simple active site with CO\textsubscript{2} is presented in Fig. 6. It can be found that in the absence of dyad and Gln151 residue, the substrate directly binds with the zinc bound hydroxyl group forming bicarbonate ion. Further, the negative charge on the bicarbonate ion is stabilised by the hydrogen of histidine ligand. These findings clearly elucidate that both catalytic dyad (Asp162-Arg164) and Gln51 residue are crucial for the catalysis.

**Coordination number of zinc in the catalytic cycle**

Coordination number of the metal ion is very important in the catalytic cycle. Evidences from the
Conceptual density functional theory analysis of reactivity

In conceptual DFT, the chemical potential, $\mu$, and global hardness, $\eta$, are defined as the following derivatives: $\mu = (\partial E/\partial N)_N = -\chi$ and $\eta = (\partial^2 E/\partial N^2)_N = (\partial \mu/\partial N)_N$, respectively, where $E$ is the total energy of the system, $N$ the number of electrons, $\nu$ the external potential, and $\chi$ electronegativity. According to Mulliken, $\mu = -\chi = -(1/2)(I + A)$ and $\eta = I - A$ with $I$ and $A$ as the first ionization energy and electron affinity, respectively, which can be approximated by the HOMO and LUMO energies as $I = -\epsilon_{\text{HOMO}}$ and $A = -\epsilon_{\text{LUMO}}$. Softness, $S$, is defined as the inverse of the hardness, $S = (1/\eta)$, and the electrophilicity index, $\omega$, a measure of the lowering of the total binding energy due to the maximal electron acceptance, can be expressed in terms of $\mu$ and $\eta$ as $\omega = (\mu^2/2\eta)$.

To describe regioselectivity tendencies of individual atoms in molecules, local descriptors were employed. The first well-known example of such a category is called the Fukui function, $f_k(r) = q_k(N+1) - q_k(N)$, in the finite difference approximation is a measure of nucleophilic attack, where $q_k(N)$ is the gross atomic charge for atom $k$ with $N$ electrons, obtained from a population analysis (such as NBO analysis). For systems that can donate electrons, the condensed Fukui index, $f_k^c = q_k(N+1) - q_k(N-1)$. For radical attack reactions, $f_k^c = (1/2)(f_k^+ + f_k^-)$, where $q_k(N+1)$, $q_k(N)$, and $q_k(N-1)$ are the gross NBO population on atom $k$ in a molecule with $N+1$ (anion state), $N$ (neutral state), and $N-1$ (cation state) electrons, respectively. The local component of the softness reactivity indicator, the local softness, introduced by Parr and Yang, is $s(r) = (\partial^2 \rho(r)/\partial \mu_N) = (\partial \rho(r)/\partial N)(\partial N/\partial \mu_N)$, has been widely used. Similar to the Fukui function, within the finite difference approximation, the

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Substrate</th>
<th>Reactant</th>
<th>EC</th>
<th>TS</th>
<th>IM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_k^+$</td>
<td>0.04461</td>
<td>-0.00005</td>
<td>0.01838</td>
<td>0.05458</td>
<td>0.00675</td>
</tr>
<tr>
<td>$f_k^-$</td>
<td>-0.00005</td>
<td>0.00012</td>
<td>-0.07450</td>
<td>0.07101</td>
<td>-0.05973</td>
</tr>
<tr>
<td>$f_k^\omega$</td>
<td>0.04467</td>
<td>-0.00006</td>
<td>0.09288</td>
<td>-0.13298</td>
<td>0.11431</td>
</tr>
</tbody>
</table>

Fig. 6—Schematic representation of the optimized geometries of minimalistic model with substrate. [(a) reactant; (b) complex].
condensed form of these three local softness indicators for any particular atom k can be written as $s_k^+ = S_f^k$ (for nucleophilic attack), $s_k^- = S_f^k$ (for electrophilic attack), and $s_k^0 = S_f^k$ (for radical attack).

Another kind of reactivity index, called the dual descriptor, has been proposed\(^3\), which serves as “an indicator for both the nucleophilic and electrophilic regions of a molecule”. Morell, Grand, and Toro-Labbe proposed the first dual descriptor using the cross-term third-order derivative $f^{(2)}(r) = (\partial^2/\partial N(\delta E/\delta N))_{N'} = (\partial^2/\partial N(\partial^2 E/\partial N^2))_{N'} = (\delta f(r)/\delta N)$, and under the finite difference approximation, one gets $f^{(2)}(r) = f^+(r) - f^-(r)$. The dual descriptor, $f^{(2)}(r)$, will be positive in electrophilic regions and negative in nucleophilic regions.

Global descriptors from conceptual DFT including hardness, softness, and electrophilicity index for the reaction coordinates are tabulated in Table 2. Since hardness is an indication of molecular stability, the smaller the hardness, the less stable is the system. The hardness values (Table 2) clearly shows that the stability of the complex increases from Ec to product.

Table 3 shows local reactivity descriptor and dual descriptor from conceptual DFT for a few key atomic sites of the system. The $f^{(2)}$ value for the carbon substrate is positive in R, EC, TS indicates the possibility for the nucleophilic attack on the substrate by zinc bound –OH. Thus, dual descriptors are useful to clearly identify the various reactive sites in the bio-catalysis.

Conclusions

In this study, DFT(B3LYP) calculations have been carried out on the reaction mechanism of $\beta$–CA. It is found from the results that the activation barrier for the nucleophilic attack is negligible which is akin to that for $\alpha$–CA. Evidences show that Asp162–Arg164 dyad orients the hydroxyl group and stabilises the negative charge on the bicarbonate. The anchoring of substrate by Gln151 explains the necessity of the above residue in the nucleophilic attack. Throughout the catalysis, the coordination number of zinc in the $\beta$–CA is four. In addition, it is observed from the results that bicarbonate binds with $\beta$ through the monodentate mode. Conceptual DFT reactivity descriptors clearly describe the nucleophilic attack on the substrate.

Acknowledgement

This study is supported by the grant from the Department of Science and Technology (DST), Government of India, New Delhi, India, under New Initiatives in Bioinorganic Chemistry.

References
