Structure and Dynamics of the Isozymes II and IX of Human Carbonic Anhydrase

Divya Rai, Satyajit Khatua, and Srabani Taraphder*

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Human carbonic anhydrases (HCAs) are responsible for the pH control and sensing in our body and constitute key components in the central pH paradigm connected to cancer therapeutics. However, little or no molecular level studies are available on the pH-dependent stability and functional dynamics of the known isozymes of HCA. The main objective of this Article is to report the first bench-marking study on the structure and dynamics of the two most efficient isozymes, HCA II and IX, at neutral pH using classical molecular dynamics (MD) and constant pH MD (CpHMD) simulations combined with umbrella sampling, transition path sampling, and Markov state models. Starting from the known crystal structures of HCA II and the monomeric catalytic domain of HCA IX (labeled as HCA IX-c), we have

generated classical MD and CpHMD trajectories (of length 1 μs each). In all cases, the overall stability, RMSD, and secondary structure segments of the two isozymes are found to be quite similar. Functionally important dynamics of these two enzymes have been probed in terms of active site hydration, coordination of the Zn(II) ion to a transient excess water, and the formation of putative proton transfer paths. The most important difference between the two isozymes is observed for the side-chain fluctuations of His-64 that is expected to shuttle an excess proton out of the active site as a part of the rate-determining intramolecular proton transfer reaction. The relative stability of the stable inward and outward conformations of the His-64 side-chain and the underlying free energy surfaces are found to depend strongly on the isozyme. In each case, a lower free energy barrier is detected between predominantly inward conformations from predominantly outward ones when simulated under constant pH conditions. The kinetic rate constants of interconversion between different free energy basins are found to span 10^7-10^8 s^{-1} with faster conformational transitions predicted at constant pH condition. The estimated rate constants and free energies are expected to validate if the fluctuation of the His-64 side-chain in HCA IX may have a significance similar to that known in the multistep catalytic cycle of HCA II.
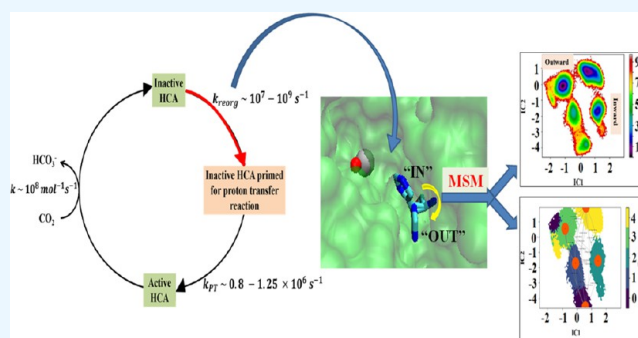
INTRODUCTION

Human carbonic anhydrases (HCAs) are ubiquitous Zn(II)-containing metalloenzymes that catalyze the reversible hydration of CO_2 to HCO_3^- in our body. Of the 15 isozymes of HCA reported to date, HCA II and IX have both been extensively pursued as potential drug targets using structural and kinetic studies. The isozymes II and IX with $k_{cat} \approx 10^6$ s^{-1} are found to have the highest catalytic activity among all HCAs.[1−6] However, these isozymes present major differences at the level of crystallographic structure and pH-dependent activity. While the molecular mechanism associated with the function of HCA II has been extensively investigated using computer simulation, little or no such studies have been carried out addressing the catalytic function of HCA IX. This Article focuses on the implications of some well-known differences between these two isozymes on the molecular mechanism of their action.

Major structural differences are presented by the high-resolution crystallographic structures of HCA II and IX. As

shown in Figure 1a and b, HCA II is a monomeric enzyme found in the cytoplasm,[1−3] while HCA IX is a dimeric, transmembrane glycoprotein.[9] Each monomer of HCA IX has multiple domains[3] consisting of an N-terminal proteoglycan-like (PG) domain, which is unique to HCA IX, a catalytic (CA) domain similar to HCA II, a transmembrane (TM) segment, and an intracytoplasmic (IC) portion.

The compact globular catalytic domain (CA) of HCA IX monomer will be designated as HCA IX-c in the present work. It measures $\sim47 \times 35 \times 42$ Å^3 and is comprised of 259 amino acid residues with a 10-stranded antiparallel β-sheet at the core
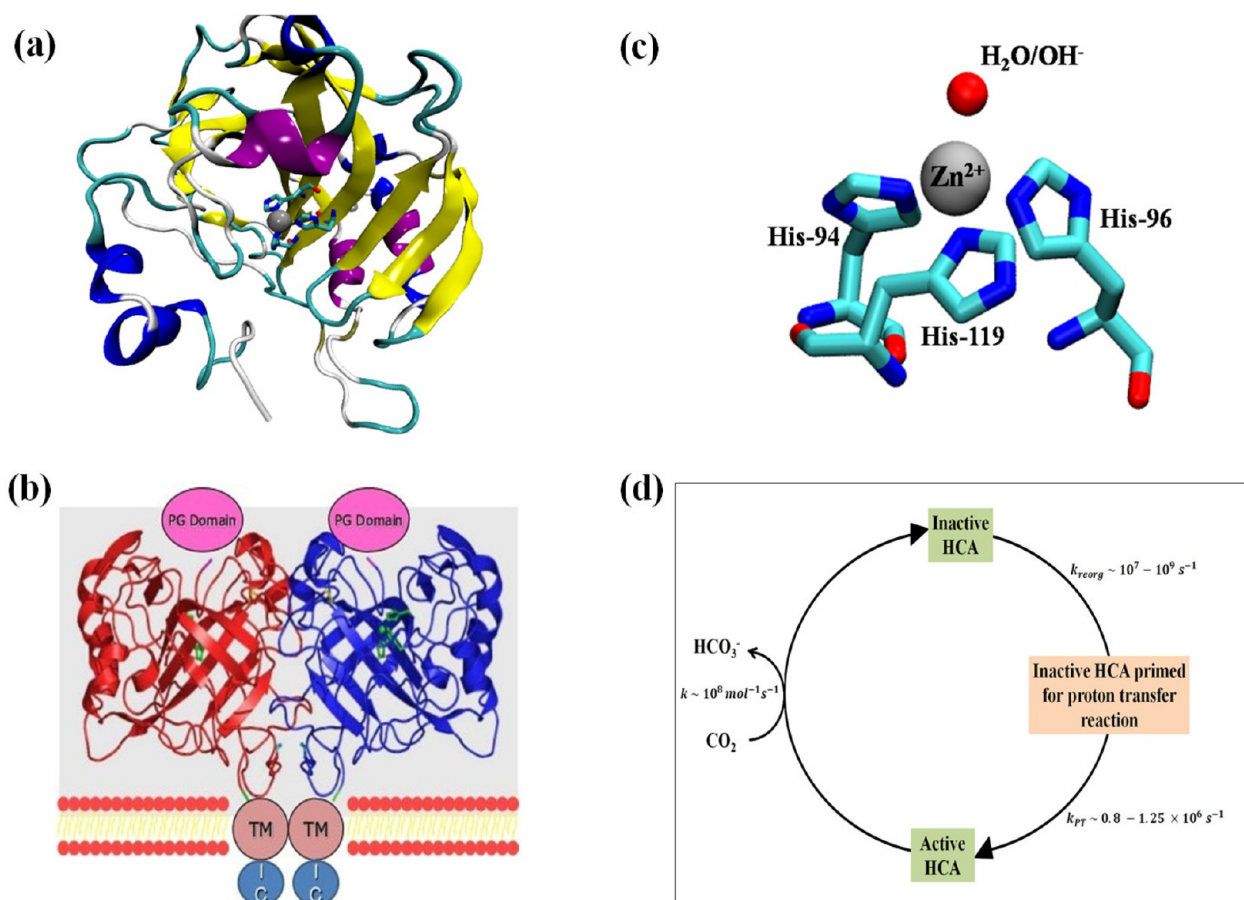
Figure 1. (a) Structure of HCA II derived from high-resolution crystallographic studies (PDB ID: 2ILI).[7] (b) Schematic diagram of the structure of human carbonic anhydrase IX (HCA IX). Figure adapted with permission from ref 8. Copyright 2010 Elsevier. (c) The active site of human carbonic anhydrases showing the Zn(II) ion coordinated to His-94, His-96, His-119, and a water molecule. (d) A general schematic representation of the catalytic mechanism exhibited in HCA along with the associated rate constants.

characteristic of the native fold of other α-HCAs.[4] It presents ~30−40% sequence identity to other catalytic HCA enzymes. While the CA domain of HCA IX has been linked to the growth and survival of tumor cells, kinetic experiments indicate the importance of the unique PG domain in HCA IX leading to better catalytic efficiency at acidic pH ($pK_a \approx 6.49$).[4] Both CA and PG domains are reported to be glycosylated. The dimerization of HCA IX is mediated by the formation of an intermolecular disulfide bond between analogous Cys residues located on two monomeric CA domains.

A complete three-dimensional characterization of the multiple domains of HCA IX is unavailable to date. Instead, the monomeric and dimeric forms of the catalytic (CA) domain have been extensively investigated, providing an interesting comparison with the well-known structure of HCA II. The high-resolution crystal structures of HCA II (PDB ID: 2ILI)[7] and HCA IX-c (PDB ID: 6FE2)[10] reveal the presence of a catalytic Zn(II) ion tetrahedrally coordinated to three histidine residues and a water molecule/hydroxide ion, as shown in Figure 1c. In both cases, the Zn(II) ion is located at the bottom of a large conical active site cavity that spans from the surface to the center of the protein and is limited by two walls, one hydrophobic and the other hydrophilic. Both isozymes are thermally stable with melting temperatures $T_m \approx$ 56−59 °C at pH = 7.0.[11] However, HCA IX-c bears only about 31% sequence identity to HCA II. The upper side of the active site of HCA IX-c is delimited by three Arg residues, Arg-58,

Arg-60, and Arg-130, which are not conserved in other members of the α-CA family.

The widely accepted mechanism of catalysis by HCA II is based on X-ray crystallographic structures and kinetic studies on the wild-type enzyme and its mutational variants.[3] In the hydration direction at physiological pH, the first stage of the catalysis involves a nucleophilic attack by the zinc-bound OH^- ion on the CO_2 molecule located in a nearby hydrophobic pocket producing a HCO_3^- and a water molecule coordinated to the zinc ion. While HCO_3^- leaves the active site, the zinc-bound OH^- ion is regenerated in the second stage of catalysis. At this stage, the zinc-bound water molecule undergoes deprotonation, and the resultant excess proton is transferred to a distal histidine residue (His-64) across the active site through a hydrogen-bonded network of water molecules. The protonated side-chain of His then rotates to shuttle the excess proton out of the active site.[12−15] At low buffer concentration, the intramolecular proton transfer step has been shown to be rate-determining in the catalysis by HCA II.[3] On the basis of the overall similarity in their function and qualitative interpretation of a vast body of experimental data, HCA IX-c is assumed to follow the same catalytic mechanism as has been established using HCA II as the template.[1,3,13−19] However, in view of the noted differences in the overall crystal structures of the two isozymes, it appears pertinent to examine how far this assumption may hold for HCA IX when probed with molecular simulation.

Extensive computer simulation studies on HCA II in water over the past couple of decades[12−16,20] have now culminated into a dynamical perspective of the catalytic mechanism corresponding to the observed turnover rate of $(\sim 0.8-1.25) \times 10^6$ s^{-1}.[3] As shown in Figure 1d, two major dynamical processes at and near the enzyme active site are found to be necessary to replicate the experimental value of k_{cat} in HCA II.

(a) The first is nonreactive dynamical reorganization at and near the active site. The stable tetrahedral coordination sphere around the Zn(II) ion undergoes a transient change to a penta-coordinated trigonal bipyramidal structure by binding to an extra water molecule. This results in the polarization of the O—H bond of the zinc-bound water molecule for its subsequent deprotonation.[18] In addition, dynamical fluctuations of the unprotonated His-64 side-chain ensure that it is predominantly orientated inward, facing the zinc ion, ready to accept the excess proton. The associated rate constants are estimated to be $k_{reorg} \approx 10^7 - 10^9$ s^{-1} as shown in Figure 1d.[12,14−16,18,20]

(b) The second is intramolecular proton transfer between a zinc-bound water molecule and His-64. This process involves translocation of the excess proton through a hydrogen-bonded network of 2−3 active site water molecules. A major activation barrier is encountered in this process that results from dynamical fluctuations of the excess proton between its Eigen and Zundel-like forms, leading to a rate constant of $k_{PT} \approx 1.0 \times 10^6$ s^{-1}.[15]

Reorganization of the enzyme at and near the active site appears to be an integral prerequisite of how the enzyme optimizes the reactive proton transfer step. In the absence of these dynamical changes around the active site, no proton transfer reaction is observed even if hydrogen-bonded pathways are formed by the active site water molecules.[21] It is important to note that all inferences presented above on the catalysis by HCA II have been derived by sampling an ensemble of fluctuating structures of the enzyme at a finite temperature. No molecular level study on the functional dynamics of HCA IX is available. Therefore, it is unclear how any of the dynamical ingredients of catalysis described above would map on to the function of HCA IX.

In distinguishing between HCA II and IX, another factor is now known to play a pivotal role. Unlike the isozyme II, HCA IX is expressed in a limited number of normal tissues.[3,22−25] A high HCA IX expression is linked with hypoxia associated tumor acidosis and the development of the metastatic phenotype.[4,26] It is reported to help in maintaining a low intracellular pH and represents the means by which cancer cells adapt to the toxic conditions of the extracellular milieu.[27,28] HCA IX also serves as a marker for the development of targeted systemic therapies and diagnostic imaging probes in solid and hypoxic tumors.[29] Circular dichroism and kinetic studies[11] have revealed that the catalytic efficiency of HCA IX is correlated to its stability between pH 3.0 and 8.0 but not above pH 8.0. Most HCAs, including HCA II, are unstable below pH 5.0. Recent research on HCAs thus emphasizes the isozyme specific inhibition of HCA IX by novel molecules at nano- to picomolar concentrations[1,30−35] for optimum therapeutic applications. However, a molecular level understanding of the observed pH dependence of activity of HCA II and IX is largely incomplete. The catalytic mechanism as described above has been inferred under charge neutralized conditions, keeping the protonation state of all titrable residues fixed. The free energy and kinetics of His-64 side-chain rotation have been studied classically with and without an

excess proton on it, keeping the protonation state of all other titrable residues (including all other His residues) fixed.[12,15,20] Evidently, the effect of fluctuations in protonation states of all titrable residues has not been investigated for HCA II so far; neither has it been probed at the molecular level for HCA IX.

In this Article, our primary aim is to develop a computer simulation study at a given pH so as to understand the change in the protonation states of titrable residues in each of the two isozymes and investigate how these changes would affect their structure and functionally relevant dynamics. Accordingly, in each isozyme, classical molecular dynamics simulations have been carried out (i) under the conventional charge neutralized condition (MD) and (ii) under the constant pH condition (CpHMD). For HCA IX, only the CA domain of HCA IX (without its unique PG domain) has been studied. This domain corresponds to the structural component common to both isozymes. Neither of the monomeric units of interest is glycosylated in the present work to simplify the comparison. Because classical MD simulations cannot be used to study bond-making/-breaking, classical MD trajectories (of length 1 μs each) have been used to investigate the nonreactive reorganization step as mentioned above. The main objective of the MD- and CpHMD-based analysis is 3-fold: (a) comparison of the crystal structures of HCA II and IX-c to the average ones derived from dynamical trajectories; (b) distribution of water molecules close to the catalytic Zn(II) ion at the respective active sites of the two isozymes; and (c) free energy and kinetics associated with the distal histidine residue (His-64).

Earlier computer simulation studies on HCA IX-c mainly focused on the comparison between the equilibrium structure and dynamics of the enzyme with and without inhibitor(s) in terms of root-mean-square deviation and fluctuation (RMSD, RMSF), radius of gyration (R_g), solvent-accessible surface area (SASA), and secondary structure dynamics. Useful information has been extracted from these studies about binding interactions, affinities, and stabilization of the inhibitor-bound enzyme.[32,36−39] Simulation studies on the folding and unfolding patterns of HCA IX-c under different conditions have also been reported.[40,41] The residues Trp-5, Gln-67, Gln-92, His-94, Glu-106, His-119, Val-131, Thr-200, and Pro-202 have been found to contribute significantly, through $\pi-\pi$, hydrogen bonding, and hydrophobic interactions, in stabilizing the active site of different inhibitor-bound systems.[37,42−44] The HCA IX-c binding site is found to be more spacious because it has a smaller hydrophobic group (Val-131) as compared to a larger hydrophobic group (Phe-131) in HCA II. The shape of the ligand therefore presents a major selectivity criterion in the inhibition.[45−48] Interestingly, the pH dependence of the structure and functional dynamics of HCA II and IX in the absence of any inhibitor has not been probed so far in any of these computer simulation studies.

■ MATERIALS AND METHODS

Crystallographic Structures of HCA II and IX. In the present study, we have employed as input the high-resolution crystal structures of the two isozymes under investigation. For HCA II, the classical MD simulation described in this Article has been initiated from PDB ID 2ILI,[7] while the transition path sampling simulations reported earlier[15,16] used PDB ID 2CBA.[49] Among the large number of structures deposited in the Protein Data Bank for the CA domain of HCA IX (designated as HCA IX-c), the starting structure used is characterized by a single point mutation (C174S) carried out

on the wild-type HCA IX-c (PDB ID: 6FE2).[10] We did not consider several other structures available in the Protein Data Bank for HCA IX-c, such as the one with a sulfonamide inhibitor bound to the metal ion at the active site (PDB ID: 3IAI)[4] or the HCA IX-c mimic (PDB ID: 4ZAO) prepared by site-directed mutagenesis of seven residues using a HCA II cDNA template.[50] Neither did we investigate the high-resolution crystal structure of HCA IX-c (PDB iID: 5DVX) that has been reported recently without any inhibitor attached to it[11] as it comprises a 6-point mutation carried out on the wild-type HCA IX-c to mimic the HCA II surface. When compared to our working structure of HCA IX-c (PDB ID: 6FE2),[10] those corresponding to the PDB IDs 3IAI[4] and 5DVX[11] showed highly similar structures with estimated root-mean-square deviations (RMSDs) for the superimposition of backbone atoms equal to 0.43 and 0.66 Å, respectively. The corresponding values of structural as well as sequence similarity were found to be 96.5% and 93.4%.

Classical Molecular Dynamics Simulation. We have carried out the following simulation studies using a single enzyme molecule (HCA II or IX-c) in water using the molecular dynamics package, AMBER 18:[51] (a) MD, classical MD simulation under charge neutralized conditions where the His-64 side-chain remains neutral all along the trajectory; and (b) CpHMD, classical MD simulation at a constant value of pH using discrete protonation states of titrable residues and explicit water molecules as solvent where we start with His-64 with a (+1) charge on its protonated side-chain.

All equilibrated trajectories reported in this Article are of length 1 μs. In each case, the length of the trajectory is so chosen as to provide converged Markov state models for the subsequent estimation of free energy and kinetics of His-64 fluctuation. For HCA II, the CpHMD trajectory has been generated at a constant pH = 7.0. On the other hand, two CpHMD trajectories of HCA IX-c are investigated at constant values of pH = 4.5 and 7. In the rest of this Article, the classical MD simulation studies under charge neutralized conditions will be referred to simply as MD, while those performed under constant pH conditions will be denoted as CpHMD.

In each case, starting from the high-resolution crystal structure, the system is represented using the Zinc amber force field (ZAFF)[52] for the active site and the empirical *ff14SB* force field[53] for the rest of the protein in the case of HCA IX and *ff12SB* in the case of HCA II. The simulated system contains the starting protein structure solvated in a large cubic box with around 16 004 TIP3P water molecules,[54] which extend 15 Å away from any given protein atom. The periodic boundary condition has been used. The AMBER force field and potential parameters mentioned earlier are used to compute the pair interactions. We perform the minimization run using the steepest descent method and the conjugate gradient method followed by heating to 300 K. Further, NPT equilibration followed by NVT equilibration is carried out. Finally, the system is subjected to 1 μs NVT MD production run at 1 atm and 300 K. The pressure is kept constant at 1 atm utilizing the Langevin piston method with a damping coefficient of 5 ps^{-1}, and the temperature is fixed at 300 K through Langevin damping[55] with a coefficient of 5 ps^{-1}. The list of nonbonded interactions is truncated at 12 Å, and a switching cutoff distance of 10 Å is used for the Lennard-Jones interactions. During the NVT equilibration, all bonds involving hydrogen atoms are constrained using the SHAKE algorithm.[56]

It is noted that the initial structure of HCA IX-c (PDB ID: 6FE2)[10] was deposited at an acidic condition of pH = 4.5. Therefore, while setting up the system for CpHMD simulations, the titratable residues such as Asp, His, Glu, Cys, Lys, and Tyr were chosen to have protonated side-chains. The viability of the structures of HCA IX-c thus sampled has been further probed by carrying out constant pH MD (CpHMD) simulation studies implemented within AMBER at pH = 4.5 and 7.0 for HCA IX-c to generate 1 μs long MD trajectories for each value of pH at 1 atm and 300 K.

We have also used the MultiSeq plugin[57] implemented in Visual Molecular Dynamics (VMD)[58] to compare representative pairs of structures extracted from different equilibrated MD trajectories. MultiSeq[57] is a bioinformatics analysis environment to investigate the homology between a pair of protein structures where a simple distance-based tree reconstruction method is employed with the help of structural measures, Q_H, RMSD, and the percent identity in sequence measurement. The structural measure, Q_H,[59,60] checks for the unnormalized $C_\alpha-C_\alpha$ pair distances that are similar between two aligned systems. For a given pair of structures, $Q_H = 1$ implies perfect alignment, and $Q_H < 0.3$ indicates a poor superposition among the C_α atoms.[60]

The active site volume of both of the enzymes has been estimated by circumscribing the active site residues by a sphere, the center of which is assumed to coincide with the center of mass of C_α atoms of all of the active site residues. The radius of this circle is calculated by the distance between this center of mass and the C_α atom located farthest from it.

Umbrella Sampling and Weighted Histogram Analysis Method. The potential of mean force (PMF) projected along the dihedral angle, $\chi_1 \equiv N-C_\alpha-C_\beta-C_\gamma$, of the His-64 side-chain was constructed using umbrella sampling,[61] employing more than 50 windows along χ_1 followed by weighted histogram analysis (WHAM).[62] For this purpose, each window was sampled for 2 ns at 300 K with a harmonic force constant of 18 kcal mol^{-1} rad^{-2} for HCA IX and 10 kcal mol^{-1} rad^{-2} for HCA II. A tolerance of 10^{-4} has been employed while using WHAM to construct the PMF.

Transition Path Sampling. For the isozyme II, as will be discussed later, the MD trajectory did not have sufficient sampling of the transition of an unprotonated side-chain of HCA II between its inward and outward orientations up to 1 μs. Therefore, for HCA II, we have utilized a transition path ensemble comprised of 1037 classical MD paths of length 1 ns each generated by aimless shooting from the free energy barrier region with respect to χ_1 and forward trajectory propagation until the system reached either state A or B.[15] In our earlier work, this transition path ensemble was used to derive an optimum reaction coordinate using the likelihood maximization method in conjunction with a Bayesian information criterion.[15] In this Article, we repurpose this ensemble to construct the Markov state model of rotation of the unprotonated His-64 side-chain in HCA II.

Free Energy and Kinetics Using Time-Lagged Markov State Models. Unsupervised learning methods of underlying functional dynamics of the enzymes from MD trajectories are now well studied and reviewed at length in the literature.[63−66] We have chosen to employ the Markov state model (MSM) implemented in PyEMMA[67] whereby the conformational space is partitioned into metastable states such that intrastate transitions are fast but interstate transitions are slow. This separation of time scales ensures that the model constructed is
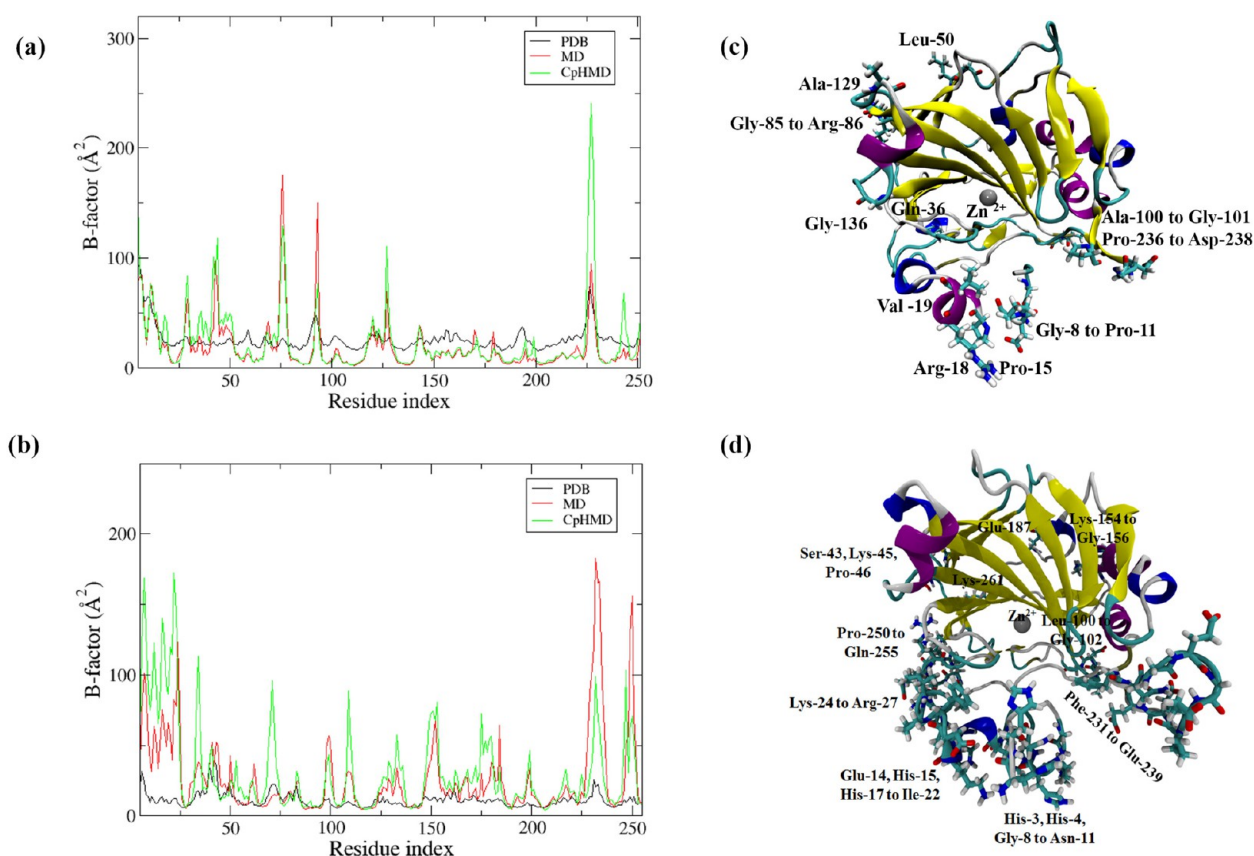
Figure 2. *B*-Factor of the C_α atom of all amino acid residues along 1 μs MD and CpHMD run for (a) HCA IX-c and (b) HCA II. The residues having a *B*-factor value greater than 45 Å^2 are highlighted in licorice representation for (c) HCA IX-c and (d) HCA II.

Markovian; that is, the probability of a transition from state *i* to state *j* depends only on the identity of *i* and not on any of the previously visited states. These probabilities can then be propagated to give long-time scale dynamics:

$$P(n\Delta t) = [T(\Delta t)]^n P(0) \qquad (1)$$

where Δt corresponds to the lag time, $P(n\Delta t)$ is a vector of state populations at time $n\Delta t$, and T is the transition probability matrix. The slowest relaxation time is then obtained as the limiting value of relaxation time against the lag time.

In the present work, the free energy profile projected along the side-chain dihedral angle, $\chi_1 \equiv N-C_\alpha-C_\beta-C_\gamma$, of the His-64 side-chain clearly indicates the presence of only two stable orientational states, A (outward) and B (inward). Several other collective variables are then tested by computing the respective distribution of their values along the MD trajectory. If a collective variable exhibits distinct peaks and nonoverlapping distribution in states A and B, it is enlisted as an order parameter (OP) for the transition between these states. All OPs thus detected (including χ_1) are then used as a set of input degrees of freedom in the time−structure independent component analysis (tICA).[68,69]

Within the framework of tICA,[68,69] highly autocorrelated linear combinations of the input degrees of freedom are used to extract the slowest-relaxing modes from a time series data set (such as the MD trajectory). Conformations sampled by the MD trajectory are then projected onto the first two tICA components and grouped into kinetically distinct macro-states using the *k*-means clustering algorithm.[70,71] The MSM thus constructed is validated by estimating the implied timescales

(ITS) for each system under investigation. From the computed ITS, the number of slow processes is found to be equal to 5 (HCA IX-c, MD), 4 (HCA IX-c, CpHMD at pH = 7), 4 (HCA II, MD), and 5 (HCA II, CpHMD at pH = 7). The limiting invariant segment of the ITS plots is then used to estimate the lag times needed to build the Markov state models. The implied timescales look only at the eigen values of the matrix T, and further verification is needed for the lag time estimated from ITS. For this purpose, the Chapman−Kolmogorov test[72] was carried out at the estimated lag times by computing the transition probability between metastable states at specified lag times and checking if eq 1 is satisfied, that is, if time evolutions of state populations ($P(n\Delta t)$) obtained from MD simulations agree with the prediction of an MSM via the replication of the transition probability matrix ($[T(\Delta t)]^n P(0)$). Both tests confirmed the choice of the lagtimes as $\Delta t = 100$ ps (HCA IX-c, MD), 15 ps (HCA IX-c, at pH = 7.0 from CpHMD), 240 ps (HCA II, MD), and 35 ps (HCA II, at pH = 7.0 from CpHMD). Finally, the rates of transition between the identified macro-states are calculated as the inverse of the mean first passage time (MFPT).[73]

■ RESULTS AND DISCUSSION

Stability and Fluctuation along Equilibrated Trajectories. Let us first present the analysis of equilibrated trajectories for HCA II and HCA IX-c in water using classical MD and CpHMD simulations at 300 K and 1 atm pressure. The calculated root-mean-square deviation (RMSD) indicates the stability of all simulated systems throughout the respective equilibrated trajectory. The radius of gyration (R_g) is found to
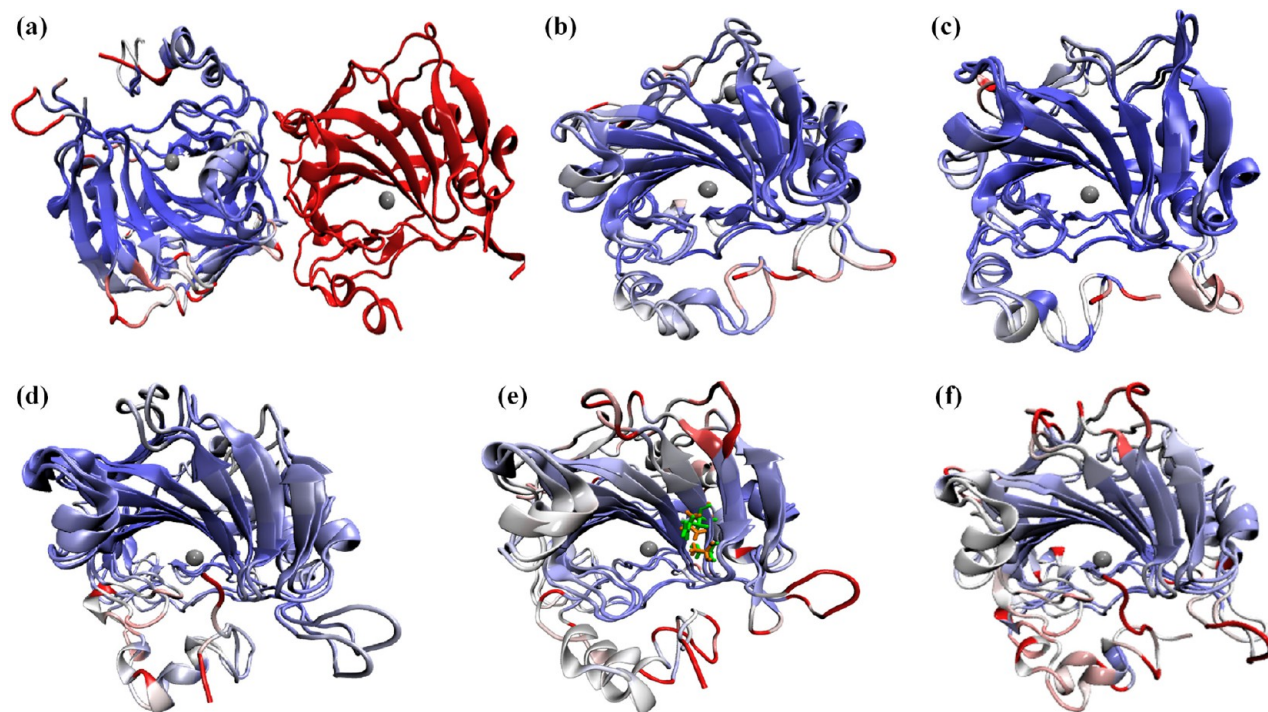
Figure 3. Structural alignment between (a) the crystal structures of HCA IX-c (PDB ID: 6FE2[10]) and HCA II (PDB ID:2ILI[7]) and the representative equilibrated structures of (b) HCA IX-c (CpHMD, pH = 4.5 and 7.0), (c) HCA IX-c (MD and CpHMD, pH = 7.0), (d) HCA II (MD and CpHMD, pH = 7.0), (e) HCA IX-c and HCA II (MD), and (f) HCA IX-c and HCA II (CpHMD, pH = 7.0). An important difference in the amino acid sequence between HCA II and IX-c near His-64 occurs where Ala-65 and Asn-67 in HCA II (orange, licorice) are replaced by Ser-65 and Gln-67 in HCA IX-c (green, licorice). The ice blue regions are structurally conserved in each case, while no correspondence in structural proximity is observed in the red regions.

Table 1. Overall Structural Similarity between the Structures of HCA IX-c and II Derived from the Crystal Structures and at $t = 1 \mu s$ of the Respective MD or CpHMD Trajectories

	pair of HCA structures compared	Q_H
(a)	IX-c crystal (PDB ID: 6FE2);[10] II (PDB ID: 2ILI)[7]	0.74
(b)	IX-c, CpHMD (pH = 4.5); IX-c, CpHMD (pH = 7.0)	0.75
(c)	IX-c, MD; IX-c, CpHMD (pH = 7.0)	0.81
(d)	II, MD; II, CpHMD (pH = 7.0)	0.72
(e)	IX-c, MD; II, MD	0.64
(f)	IX-c, CpHMD (pH = 7.0); II CpHMD (pH = 7.0)	0.60

vary between 17.3 and 17.8 Å for HCA IX-c, indicating the persistence of a compact globular structure. The estimated root-mean-square fluctuation (RMSF) values and the B-factors of C_α atoms of all amino acid residues calculated therefrom show large deviations from their respective crystallographic structures in both isozymes, as shown in Figure 2a and b. As expected, an exact mapping was not observed due to the finite temperature and solvent effects. These analyses identify the flexible regions to be located predominantly on the surface of the enzyme. All residues with higher RMSF values, highlighted in Figure 2, are found to constitute loops and turns and hence do not belong to any well-defined secondary structure. In Figure 2, the N-terminal and C-terminal free ended residues have been excluded to focus on the more relevant fluctuations in each system. Further details are provided in the Supporting Information. It is interesting to note that the same residues are identified as flexible ones by both MD and CpHMD. Yet the

extent of fluctuation is found to be larger along the CpHMD trajectory for both of the isozymes.

Secondary Structures from Equilibrated Structures. We present in Figure 3 the alignment of different pairs of structures (a) derived from HCA IX-c and HCA II from X-ray diffraction studies and from the last frame of the equilibrated MD/CpHMD trajectories of (b) HCA IX-c cpHMD at constant pH = 4.5 and 7.0, (c) HCA IX-c MD and CpHMD at pH = 7.0, and (d) HCA II MD and CpHMD at pH = 7.0. As shown in Figure 3e and f, we have also compared the structures of HCA II and IX-c extracted from the respective MD and CpHMD runs at pH = 7.0. In each case, several structural parameters have been estimated to understand the underlying deviations (details in the Supporting Information).

A quantitative comparison of all pairs of structures shown in Figure 3 has been obtained by estimating the structural homology parameter, Q_H, and the results are shown in Table 1. HCA IX-c prepared at pH = 4.5 and at pH = 7.0 exhibit $Q_H =$ 0.75 implying a high degree of alignment between them. This observation implies the usability of the crystal structure of HCA IX-c (PDB ID: 6FE2)[10] to generate a stable structure at pH = 7.0. Similar high degrees of alignment are observed from the selected structure extracted from the MD and CpHMD simulations for both HCA II and IX-c. It is therefore concluded that there is no major change at the level of the secondary structures when a neutral pH mimic of HCA IX-c and HCA II is constructed by using variable protonation states of the titratable residues such as Asp, His, Glu, Cys, Lys, and Tyr. However, a much poorer superposition of the backbone C_α atoms is predicted upon aligning HCA IX-c and HCA II structures with $Q_H = 0.64$ (MD) and $Q_H = 0.60$ (CpHMD).

Table 2. Different Segments and Their Associated Major Secondary Structure Element Present in the Equilibrated Structures of HCA IX-c and HCA II in Water Determined Using STRIDE[76],[a]

| | | residue number | | | | | |
| | | HCA II | | | HCA IX-c | | |
region	major secondary structure	PDB ID: 2ILI [7]	classical MD	pH = 7.0	PDB ID: 6FE2 [10]	classical MD	pH = 7.0
1	αA	16−19		14−18	13−19	16−18	16−18
2	αB	21−24	21−23	25−28	21−24	21−24	21−24
3	βa	32−33		32−33	32−33	39−40	39−40
4	αC	35−37					
5	βJ	39−40		38−40	39−40		
6	βb	48−50	46−50	48−50	46−50		48−50
7	βB	57−61	56−61	57−61	56−61	56−61	57−61
8	βC	66−69	66−70	66−70	66−70	66−69	66−71
9	βc	78−82	78−82	78−82	78−82	78−80	78−82
10	βD	86−97	88−97	88−97	87−97	87−97	86−97
11	βd	108−109		107−109	108−109	108−109	108−109
12	βE	116−124	116−124	116−124	116−124	116−124	116−124
13	αD	131−134	131−135	131−134		131−133	131−133
14	βF	141−150	141−150	141−150	141−150	141−151	141−151
15	αE	155−166	157−162	154−166	155−162	155−166	165−167
16	βA	172−176	172−178	172−174	172−178	172−177	172−176
17	αF	181−184			181−184	182−184	
18	βH	191−196	191−195	191−196	191−196	191−196	191−197
19	βG	207−212	207−212	207−212	207−212	207−212	205−212
20	βe	216−218	216−218	216−218	216−218	216−219	216−219
21	αG	220−228	220−227	219−227	220−227	220−227	220−228
22	βi	257−258	257−258	257−258	257−258	257−258	257−258

[a]The list of secondary structure elements has been prepared on the basis of the crystal structure of HCA II (PDB ID: 2ILI).[7]
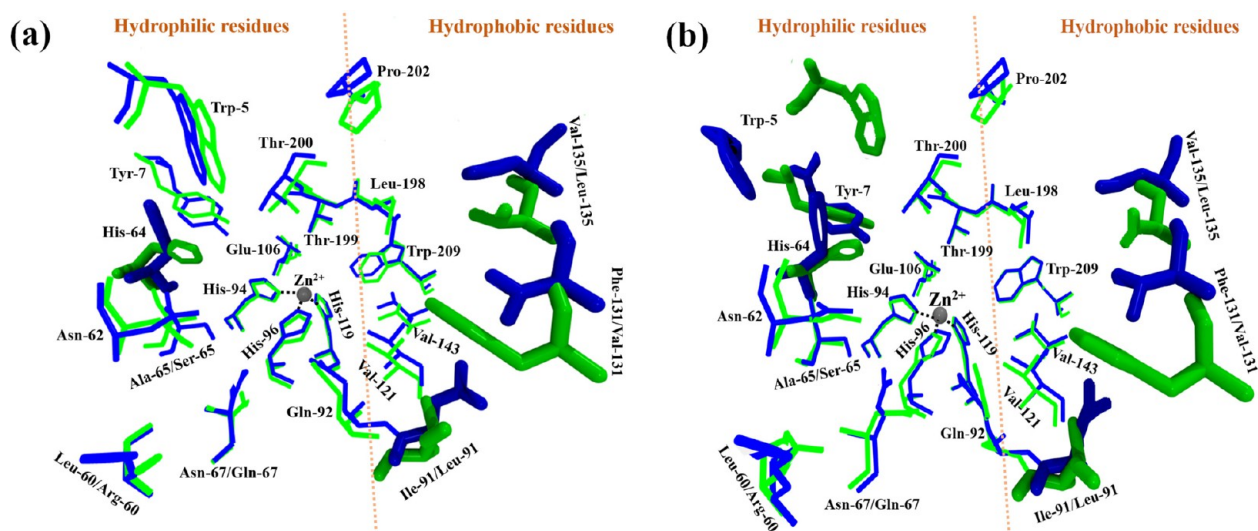

Figure 4. Comparison of the active site structures in HCA IX-c (in blue) and in HCA II (in green) using the positions averaged over a 1 μs long trajectory of (a) MD and (b) CpHMD simulations. Different polar and nonpolar residues are separated by an orange dashed line. The residues assumed to be different are highlighted in licorice, and the others are shown with thick line representation.

Therefore, important reorganizations appear to distinguish the chosen structures of HCA IX-c and II.

The corresponding variations of secondary structure elements between these two systems present in their respective crystal structures as well as equilibrated structures obtained from classical MD and CpHMD simulations have been summarized in Table 2. The residues Ala-65 and Asn-67 have been reported to be coupled to the fluctuation of His-64 in HCA II.[14] These residues have been replaced by Ser-65 and Gln-67 in HCA IX, resulting in a small variation between βB

and βC as highlighted in Figure 3e. The central hydrophobic β-core, comprised of βA−βJ and the other relatively smaller β-structures (βa−βe), remains mostly unchanged along with the α and 3−10 helical structures (αA−αG). The most notable changes are observed on the surface of the protein only.[74,75]

Comparison of Average Active Site Structures. As shown in Figure 4, we have checked the structural integrity of the active site in HCA IX-c and HCA II, aligning their averaged structures obtained by sampling the respective simulated trajectories. It is evident from the MD studies that polar
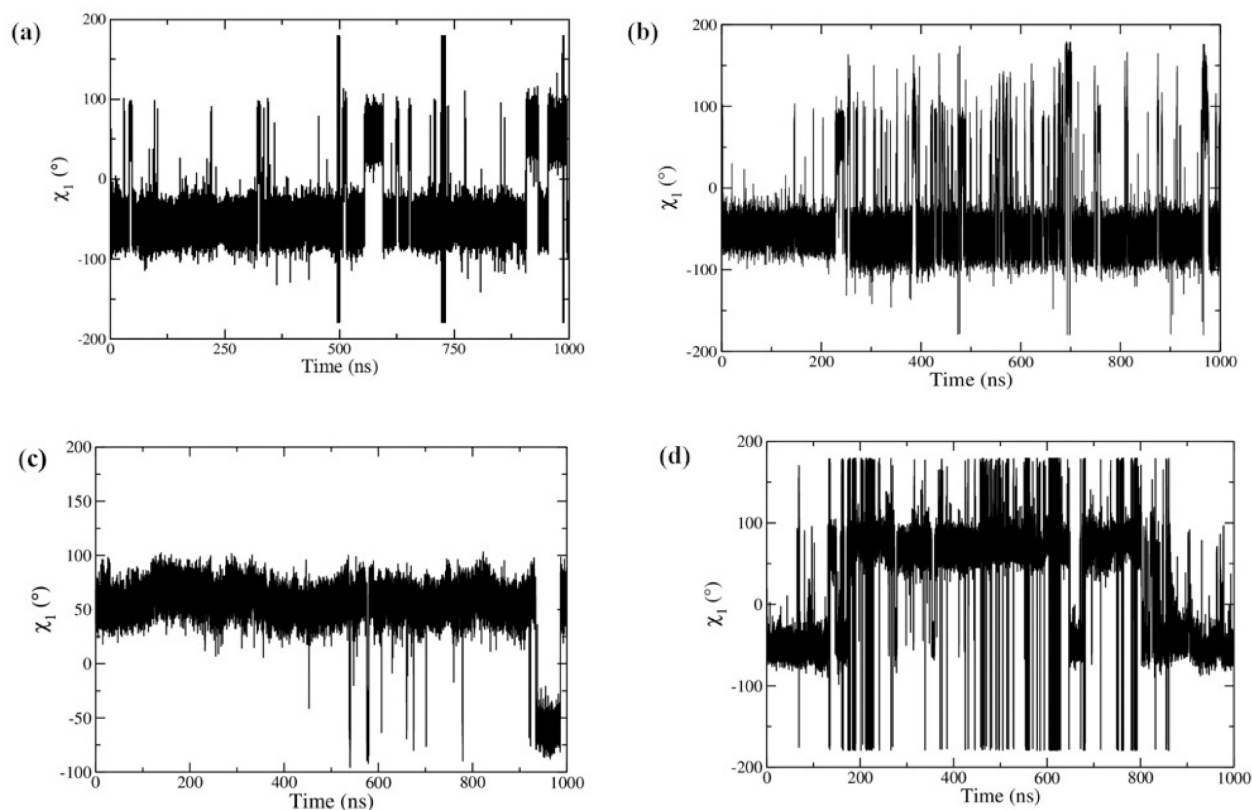
Figure 5. Side-chain dihedral χ_1 ($N-C_\alpha-C_\beta-C_\gamma$) fluctuations of His-64 of (a) HCA IX-c (MD), (b) HCA IX-c (CpHMD at pH = 7.0), (c) HCA II (MD), and (d) HCA II (CpHMD at pH = 7.0) along the respective trajectories of length 1 μs each.
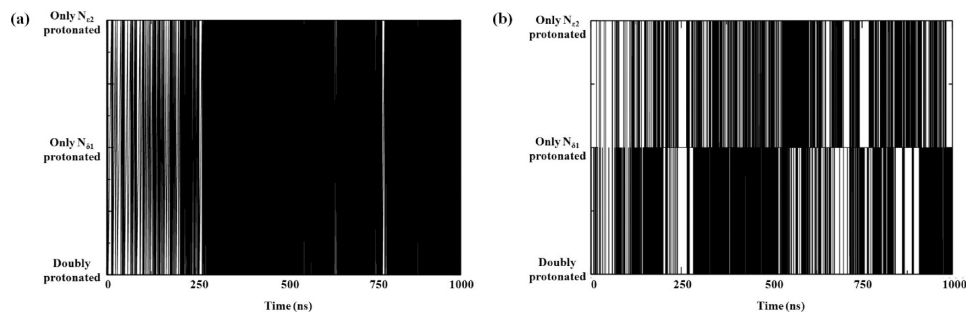

Figure 6. Fluctuation of the protonation state adopted by the His-64 side-chain along the CpHMD trajectory at pH = 7.0 for (a) HCA IX-c and (b) HCA II.

residues at the active site remain mostly invariant with a marked exception of side-chain orientation of His-64. Similar to the crystal structure, the average active site structure detects a predominant outward orientation of His-64 in HCA IX-c. In contrast, an inward orientation has been observed in the case of HCA II. Other major differences in the active site have been shown by the residues 91, 131, and 135 at the hydrophobic side of the active site wall.

The residues highlighted in Figure 4 have been identified along both MD and CpHMD simulations of HCA II and IX-c. The CpHMD simulation of HCA IX-c at pH = 7 presents an additional higher degree of fluctuation by the N-terminus residues (including catalytically important Trp-5 and Tyr-7) as compared to those observed along the corresponding CpHMD trajectory of HCA II. Such disorder is expected at the N-terminus that is not tethered to any support. Importantly, irrespective of the simulation method employed, the average active site structure in each isozyme is found to remain stable

even with the variations in protonation states of all titrable residues. A similar figure has been presented in ref 1 comparing the active sites of HCA II and IX-c on the basis of their crystallographic structures only. Figure 4, however, not only shows the average structure as sampled from dynamical trajectories but also highlights the effects of variable protonation states of titrable residues on the average structure, resulting in small local changes at the active site.

Comparison of Functionally Relevant Dynamics. Although HCA II and IX-c exhibit similar trends at the level of overall stability, secondary structure elements, and active site structure, several local differences are observed, which also depend on the method of simulation used. We next evaluate the effect of such differences on the various dynamical factors that are expected to be important in the catalysis of the two HCA isozymes being studied at neutral pH.

Dynamics of the His-64 Side-Chain. Conformational dynamics of the side-chain of His-64, observed along the 1

Table 3. Predicted pK_a Values of the Side-Chains of Different Titrable Residues Present in HCA IX-c and HCA II Determined by Sampling in Each Case a 1 μs Long CpHMD Trajectory at pH = 7

HCA IX		HCA II	
residue	pK_a	residue	pK_a
Tyr-7	12.071	His-4	6.039
Asp-10	4.351	His-10	6.191
Asp-32	0.699	Glu-14	5.238
Glu-48	5.812	His-15	5.869
Glu-56	5.851	His-17	5.595
His-64	7.312	Asp-19	4.171
Glu-78	6.064	Glu-26	4.150
Glu-87	6.211	Asp-32	3.176
Tyr-88	12.15	Asp-34	5.539
Glu-106	3.642	His-36	6.782
His-107	5.980	Asp-41	3.756
Glu-110	5.472	Asp-52	3.134
His-112	7.301	His-64	6.336
Glu-117	4.203	Glu-69	3.993
His-122	5.972	Asp-71	3.375
Asp-132	3.618	Asp-72	4.766
Glu-133	4.807	Asp-75	4.090
Glu-149	4.891	Asp-85	4.857
Glu-150	5.387	Asp-101	3.826
Glu-152	5.443	Glu-106	1.778
Glu-153	5.533	His-107	6.531
Tyr-157		Asp-110	4.322
Glu-158	5.828	Glu-117	4.751
Glu-165	5.343	His-122	4.419
Glu-166	4.781	Asp-130	4.014
Glu-169	5.312	Asp-139	5.543
Glu-170	4.757	Asp-162	4.133
Glu-173	3.906	Asp-165	3.683
Asp-180	3.852	Asp-175	4.430
Asp-188	4.568	Asp-180	5.210
Tyr-191		Glu-187	5.509
Tyr-194	9.471	Asp-190	4.483
Glu-195	5.041	Glu-205	5.032
Lys-221	9.990	Glu-214	5.382
His-224	6.105	Glu-221	4.002
Asp-227	5.522	Glu-234	4.262
Asp-238	5.147	Glu-236	5.106
Glu-257	5.710	Glu-238	4.468
		Glu-239	6.011
		Asp-243	4.819

μs long classical MD and CpHMD trajectories for HCA IX-c, have been presented in Figure 5a and b. It is important to note that, in our CpHMD studies, the His side-chain is initially positively charged, with both N-atoms in the imidazolium group being protonated. However, in the dynamics reported for HCA IX-c MD, there is no such extra charge on its side-chain. The observed transitions are broadly categorized to take place between two orientational states of the side-chain as follows with respect to the Zn(II) ion that is located at the bottom of the conical active site cavity: inward conformation, where the side-chain of His-64 is pointing toward the Zn(II) ion; and outward conformation, where the side-chain of His-64 is pointing toward the bulk solvent, facing away from the Zn(II) ion.

Identification of these two states is based on the observation of two side-chain conformations of His-64 in the high-resolution crystal structure of HCA II.[7,49] However, in the crystal structure of HCA IX-c (PDB: 6FE2),[10] His-64 was found only in an outward orientation where the distance from the Zn(II) ion to $N_{\delta 1}$ of His-64 ($r_{Zn-N_{\delta 1}}$) was estimated to be 10.96 Å, and its side-chain dihedral angles are $\chi_1 \equiv N-C_\alpha-C_\beta-C_\gamma = -47.00°$ and $\chi_2 \equiv C_\alpha-C_\beta-C_\gamma-N_{\delta 1} = -64.02°$. As evident from Figure 5a and b, in both HCA IX-c mimics, the His-64 side-chain undergoes repeated transitions between its inward and outward orientations. The frequency of transition is estimated to be higher from the CpHMD trajectory. The associated fluctuations in the protonation state of His-64 side-chain along the respective CpHMD trajectories are presented in Figure 6. A frequent variation of the protonation states of the titrable residues at and near the active site is expected to exert a marked effect on the frequency with which the orientation of His-64 side-chain changes. We note that the titrable residues Tyr-7, Glu-170, and Glu-106 are identified within 5 Å for $N_{\delta 1}$ or $N_{\epsilon 2}$-atoms along the CpHMD trajectory of HCA IX-c at pH = 7.0 with decreasing frequencies of occurrence in the order shown. A similar list for HCA II (CpHMD, pH = 7.0) includes Tyr-7, Glu-106, His-4, and Lys-170. As mentioned earlier, the lower number of transition events observed along the MD trajectory of HCA II is statistically insufficient for our subsequent construction of the Markov state model.

Multiple Protonation States of Titrable Residues. The dynamics of the His-64 side-chain evidently depend on the fluctuating electrostatic environment presented by the surrounding protein matrix. To understand this effect, various protonation states and their respective populations were sampled throughout the course of the CpHMD simulations of both HCA IX-c and II. The estimated values of pK_a for all titrable residues in these systems have been highlighted in Table 3. The associated equilibrium populations of underlying protonation states have been summarized in Table 4. Even up to 1 μs, a couple of Tyr residues did not have a measurable sampling of their alternative protonation states and have therefore been omitted in Table 3. The populations of most of the other titrable residues are found to be more than 90% in one protonation state. The most notable exceptions are Glu-78, Glu-87, His-107, His-112, His-122, and His-224 (HCA IX-c); His-10, His-15, His-17, His-36, and His-107 (HCA II). Since electrostatic interactions are long ranged, the dynamics of population of different protonation states are expected to affect the catalytic centers even if they are not in the immediate vicinity of any of these residues. As a result of these variations, the side-chain pK_a values of His-64 are found to be 7.312 in HCA IX-c and 6.336 in HCA II. Accordingly, along the respective 1 μs-long CpHMD trajectory, the side-chain of His-64 remains mostly protonated (~67%) in HCA IX-c, whereas the unprotonated state (with only one H-atom connected to either $N_{\delta 1}$ or $N_{\epsilon 2}$) is predominantly populated (~82%) in HCA II. Furthermore, when similar calculations were carried out for HCA IX-c at low acidic pH (pH = 4.5 and 3.0), the His-64 side-chain was found to remain protonated for ~99.9% of all structures sampled along the corresponding CpHMD trajectory. These results clearly establish the accuracy of the description of protonation states as adopted in our simulation studies.

Table 4. Equilibrium Populations of Every State for Every Titratable Residue Present in HCA IX-c and HCA II Determined in Each Case Using a 1 μs Long CpHMD Trajectory at pH = 7[a]

| | HCA IX-c | | | | | | HCA II | | | | |
| | population | | | | | | population | | | | |
residue	state 0	state 1	state 2	state 3	state 4	residue	state 0	state 1	state 2	state 3	state 4
Tyr-7	0.999991	0.000009				His-4	0.098544	0.755239	0.146217		
Asp-10	0.997760	0.001001	0.000030	0.001173	0.000037	His-10	0.134398	0.758692	0.106910		
Asp-32	0.999999	0.0	0.0	0.000001	0.0	Glu-14	0.982993	0.007946	0.000544	0.007947	0.000570
Glu-48	0.939042	0.023304	0.004866	0.022996	0.009793	His-15	0.072982	0.627641	0.299377		
Glu-56	0.933797	0.031956	0.001433	0.030599	0.002215	His-17	0.037882	0.516441	0.445677		
His-64	0.672272	0.299792	0.027936			Asp-19	0.998518	0.001127	0.000010	0.000336	0.000009
Glu-78	0.896206	0.036180	0.009059	0.042440	0.016115	Glu-26	0.998590	0.000578	0.000046	0.000716	0.000070
Glu-87	0.860215	0.024763	0.045803	0.024321	0.044900	Asp-32	0.999850	0.000065	0.000004	0.000077	0.000004
Tyr-88	0.999993	0.000007				Asp-34	0.966569	0.000316	0.000000	0.000377	0.032738
Glu-106	0.999562	0.000084	0.000016	0.000261	0.000078	His-36	0.377154	0.556422	0.066424		
His-107	0.087119	0.450464	0.462418			Asp-41	0.999430	0.000286	0.000000	0.000284	0.000000
Glu-110	0.971183	0.013728	0.000288	0.014520	0.000282	Asp-52	0.999864	0.000055	0.000001	0.000080	0.000000
His-112	0.666817	0.194621	0.138563			His-64	0.178158	0.808703	0.013139		
Glu-117	0.998407	0.001593	0.0	0.0	0.0	Glu-69	0.999017	0.000510	0.000010	0.000461	0.000002
His-122	0.085700	0.760399	0.153902			Asp-71	0.999763	0.000143	0.000000	0.000093	0.000001
Asp-132	0.999585	0.000151	0.000002	0.000108	0.000154	Asp-72	0.994200	0.003264	0.000023	0.002484	0.000029
Glu-133	0.993630	0.002205	0.001021	0.002201	0.000944	Asp-75	0.998771	0.000573	0.000009	0.000642	0.000005
Glu-149	0.992286	0.003635	0.000213	0.003622	0.000246	Asp-85	0.992850	0.002183	0.000010	0.004946	0.000011
Glu-150	0.976205	0.011962	0.000124	0.011582	0.000128	Asp-101	0.999330	0.000302	0.000005	0.000348	0.000015
Glu-152	0.973037	0.012016	0.001457	0.011845	0.001646	Glu-106	0.999994	0.000006	0.000000	0.000000	0.000000
Glu-153	0.967020	0.015552	0.002005	0.013875	0.001549	His-107	0.253662	0.739068	0.007270		
Tyr-157	1.000000	0.0				Asp-110	0.997905	0.000982	0.000204	0.000869	0.000040
Glu-158	0.936886	0.024064	0.006092	0.023618	0.009341	Glu-117	0.994393	0.001514	0.000002	0.004090	0.000001
Glu-165	0.978446	0.010385	0.000619	0.010226	0.000325	His-122	0.002618	0.947544	0.049838		
Glu-166	0.994001	0.002776	0.000133	0.002784	0.000307	Asp-130	0.998968	0.000518	0.000013	0.000488	0.000013
Glu-169	0.979879	0.009284	0.000636	0.009607	0.000595	Asp-139	0.966239	0.017303	0.000214	0.015000	0.001244
Glu-170	0.994318	0.002201	0.000373	0.002209	0.000899	Asp-162	0.998643	0.000660	0.000013	0.000680	0.000004
Glu-173	0.999194	0.000417	0.000010	0.000367	0.000012	Asp-165	0.999518	0.000224	0.000003	0.000246	0.000009
Asp-180	0.999290	0.000424	0.000002	0.000261	0.000024	Asn-175	0.997317	0.001183	0.000032	0.001434	0.000034
Asp-188	0.996314	0.003195	0.000043	0.000410	0.000038	Asp-180	0.984058	0.004356	0.000017	0.011535	0.000034
Tyr-191	1.000000	0.0				Glu-187	0.968707	0.011409	0.003988	0.012482	0.003414
Tyr-194	0.996632	0.003368				Asp-190	0.996970	0.001450	0.000011	0.001554	0.000015
Glu-195	0.989128	0.005660	0.000092	0.004936	0.000185	Glu-205	0.989352	0.003500	0.000387	0.005008	0.001753
Lys-221	0.998978	0.001022				Glu-214	0.976474	0.009208	0.002160	0.010318	0.001840
His-224	0.113033	0.599674	0.287293			Glu-221	0.998997	0.000527	0.000017	0.000442	0.000017
Asp-227	0.967829	0.009360	0.002986	0.009809	0.010017	Glu-234	0.998176	0.000899	0.000057	0.000827	0.000041
Asp-238	0.986179	0.008855	0.000224	0.004515	0.000227	Glu-236	0.987400	0.005664	0.000892	0.005258	0.000786
Glu-257	0.951212	0.022644	0.001764	0.022190	0.002192	Glu-238	0.997070	0.001449	0.000049	0.001350	0.000082
						Glu-239	0.906978	0.038040	0.006949	0.041616	0.006417
						Asp-243	0.993456	0.003355	0.000067	0.003043	0.000079

[a]The definition of different protonation states are available in AMBER CpHMD documentation.[51]

Hydration and Proton Transfer Paths at the Active Site. Figure 7 highlights the volume slice(s) representing 50% occupancy of the 10−11 active site water molecules located within 8 Å from the Zn(II) ion in HCA II and IX-c. In both isozymes, the classical MD simulations under charge neutral conditions predict the active site water molecules to be more disordered, while the hydration sites are found to be more localized in comparison as sampled along the CpHMD trajectories. These active site water molecules are expected to constitute the hydrogen-bonded network that serves as the crucial proton path between the Zn(II)-bound water molecule (Zn-wat) and the His-64 side-chain in its inward orientation. The average number of water molecules present in the active site from the classical MD simulation was found to be ~18 (HCA IX-c) and ~22 (HCA II) when His-64 is present in its inward conformation. When His-64 is in its outward conformation, the average number of active site water molecules turns out to be ~19 (HCA IX-c) and ~23 (HCA II). However, CpHMD simulations yielded the same to be ~16 (HCA IX-c) and ~19 (HCA II) in the inward orientation of His-64 and ~17 (HCA IX-c) and ~21 (HCA II) in the latter's outward conformation. In Figure 8a, a complete proton transfer path is evident between Zn-wat and the inward conformation of the His-64 side-chain mediated by the water molecules labeled as W1, W2, W3a, and W3b. However, no such path leading up to the His-64 side-chain (in its outward conformation) is detected in the crystal structure of HCA IX-c as shown in Figure 8b, indicating the proton transfer path to be inherently dynamic in nature. We have further examined two representative structures sampled from the 1 μs long classical
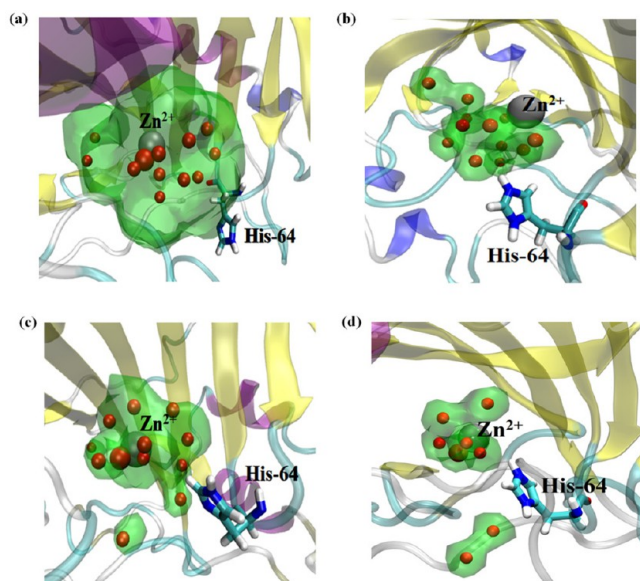
Figure 7. Hydration of the active site of (a) HCA IX-c (MD), (b) HCA IX-c (CpHMD at pH = 7.0), (c) HCA II (MD) showing the isosurfaces, and (d) HCA II (CpHMD at pH = 7.0) with 50% occupation probability of water molecules present within 8 Å of the Zn(II) ion.
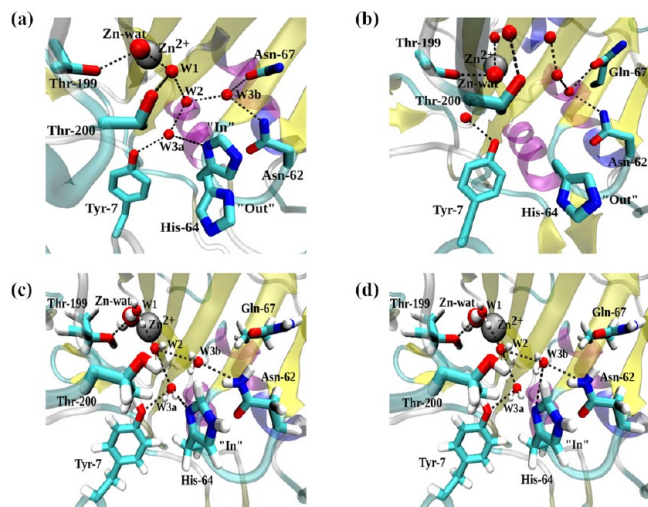

Figure 8. Putative proton transfer paths connecting Zn-water and the catalytically important histidine residue (His-64) at the active site of HCA II and HCA IX-c. The upper panel shows the hydrogen-bonded network of the water molecules as obtained from the high-resolution crystal structure of (a) HCA II (PDB ID: 2ILI)[7] and (b) HCA IX-c (PDB ID: 6FE2).[10] The lower panel demonstrates the two proton transfer paths present in the two representative structures of HCA IX-c extracted from its MD trajectory with a minimum potential energy and having the His-64 side-chain pointed inward.

MD trajectory of HCA IX-c, which correspond to minimum potential energy. Unlike the crystal structure, the His-64 side-chain is found to have undergone sufficient reorientation so as to allow the formation of a complete proton transfer path linking it to the Zn-wat. For ease in comparison with HCA II, we have labeled the analogous active site water molecules participating in the hydrogen-bonded network as W1, W2, W3a, and W3b. We, therefore, conclude that transient proton paths in HCA IX-c may be formed when inward His-64 orientation is populated. This mandates a further probe into

the relative population of the two conformations of the His-64 side-chain in equilibrium at neutral pH.

The equilibrium distribution of water molecules around the Zn(II) ion has been presented in Figure 9 in terms of the radial distribution function, $g(d)$, where d signifies the distance separating the O-atom of a water molecule and the Zn(II) ion. As discussed before, the water structure is found to be more localized along the CpHMD trajectory in both isozymes, the effect being more pronounced in the case of HCA II. Interestingly, the first peak in $g(d)$ represents the coordination between Zn(II) and zinc-bound water; its location shows the average distance of separation, and its height signifies the strength of their interaction. Both of these parameters are found to depend crucially on the simulation condition used. The location of the first coordination shell, that is, the presence of zinc-bound water, is more ordered in HCA II (CpHMD). We have also observed transient changes in the number of water molecules within 1.95 Å of the catalytic Zn(II) ion. For HCA IX-c, ∼0.01% of the structures obtained from classical MD and ∼0.5% of the CpHMD structures show a transient penta-coordinated environment around Zn(II) with the presence of an excess water (Ex-wat). HCA II, however, exhibits a more significant population of the same (∼53.5% from classical MD and ∼1.42% from CpHMD trajectories). The penta-coordinated state is so short-lived in HCA IX-c that it does not bring in any marked change in the average coordination number, CN, of the Zn(II) ion. We find CN = ∼1.0 (MD) and ∼1.004 (CpHMD) for HCA IX-c. However, the values CN = ∼1.53 (MD) and ∼1.01 (CpHMD) have been calculated in the case of HCA II. It will indeed be interesting to resolve the importance of a transient penta-coordinated state of Zn(II) ion on the catalysis by HCA IX-c. Further studies employing QM-MM-MD simulation interfaced with some enhanced sampling method need to be carried out, which is beyond the scope of this Article. Illustrative examples of the penta-coordination environment at the active site of HCA IX-c and HCA II are shown in Figure 9c,d, which have been extracted from the respective MD trajectories.

The active site volume has been found to be 31 043.61 Å^3 for HCA IX-c and ∼24 292.33 Å^3 for HCA II. The associated changes in active site due to the His-64 side-chain fluctuations are well-reflected in the change of ∼618 Å^3 active site volume of HCA IX-c and a comparatively lesser pronounced change of ∼110 Å^3 in HCA II.

Stable Side-Chain Orientations of the His Side-Chain. We next set to identify the stable orientations of the His side-chain in both of the isozymes by plotting the potential-of-mean-force (PMF) along the side-chain dihedral angle χ_1. As shown in Figure 10, we have used both MD and CpHMD simulations to sample each window in the associated umbrella sampling simulations. In addition, the variation of the number of stable states with pH has also been investigated by carrying out CpHMD simulations at pH = 4.5 and 7.0 for HCA IX-c and at pH = 7.0 for HCA II. It is noted that the details of the PMF profile depend crucially on the pH employed. However, irrespective of the type of simulation performed or the protonation state of the His side-chain, only two stable conformations are detected. These two conformations could be mapped on to the ones obtained earlier for HCA II as the outward ($\chi_1 \lesssim 0°$) and inward ($\chi_1 \gtrsim 0°$) orientations of the His-64 side-chain. The differences in free energies at the minima $\Delta F = F_{\text{inward}} - F_{\text{outward}}$ and the free energy barriers,
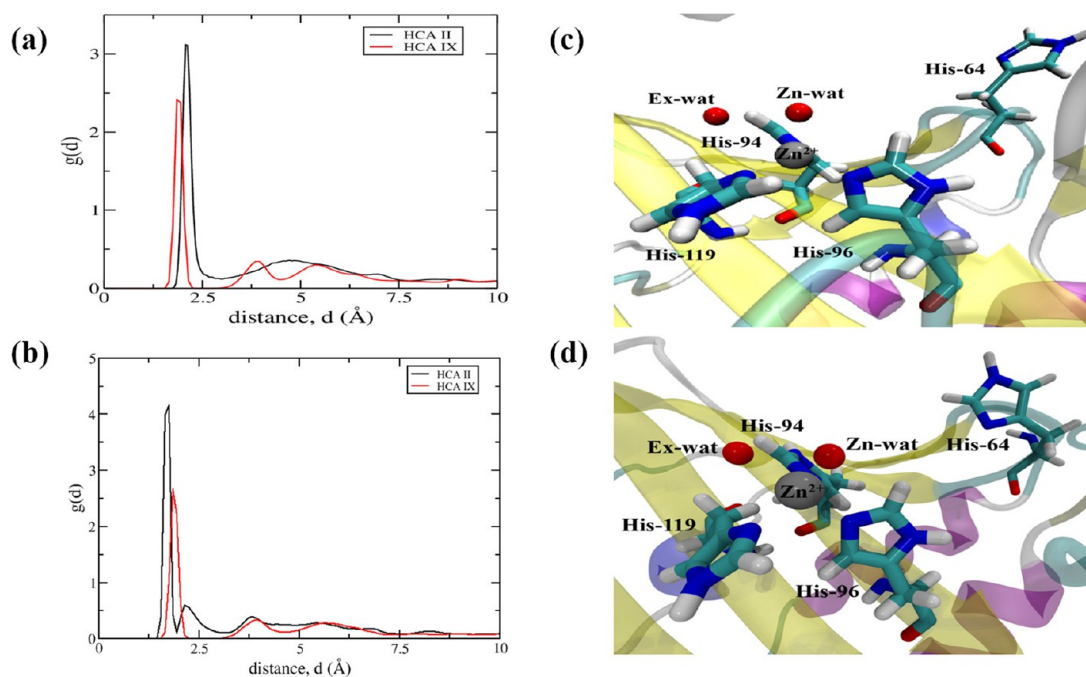
Figure 9. Radial distribution function, $g(d)$, at a distance d between the zinc ion and the O-atom of a neighboring water molecule estimated of HCA IX-c and HCA II for the 1 μs trajectory obtained from (a) classical MD and (b) CpHMD. The active site structure represents the penta-coordinated state around the Zn(II) ion (c) for HCA IX-c and (d) for HCA II. Also shown are zinc in VDW (gray), Zn-wat and Ex-wat in cpk (red), and amino acid residues in licorice representation.
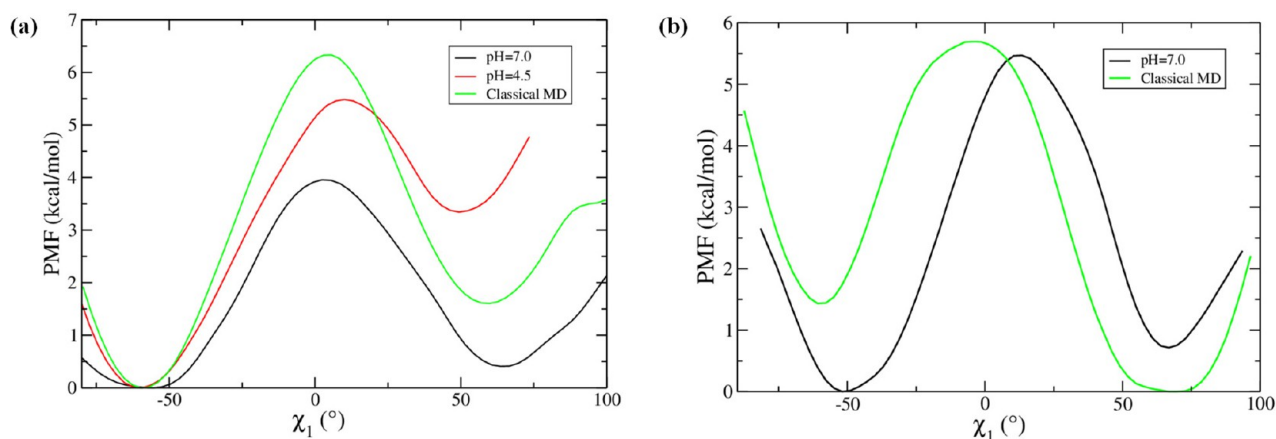

Figure 10. Potential of mean force (PMF), projected along χ_1, for the side-chain reorientation of His-64 in (a) HCA IX-c and (b) HCA II using umbrella sampling and WHAM.

$\Delta F^{\ddagger}$, for the outward to inward orientational transition in all five simulations are summarized in Table 5.

In all three cases shown for HCA IX-c, the outward conformation is found to be more stable and located around the same position along χ_1. This is indicated by the positive free energy difference, $\Delta F = F_{inward} - F_{outward}$. ΔF is found to be maximum at pH = 4.5, thereby corroborating the detection of outward conformation only in the crystal structure of HCA IX-c recorded at the same pH. Unlike the outward conformer, the location and relative stability of the inward conformation are found to show a strong dependence on the conditions used in the respective simulations, as indicated in Table 5. The inward conformer in His-64 is found to be substantially more stabilized as the pH is changed from 4.5 to 7.0, thereby reducing the value of ΔF to 0.40 kcal mol^{-1}. Because the free energy barrier separating the two orientations is quite high,

both conformations of the His-64 side-chain are likely to be detected in HCA IX-c if crystallized at neutral pH.

In earlier simulation studies on HCA II using the classical MD method,[12,15] the outward conformer is more stable in HCA II when both N-atoms on the side-chain are protonated ($\Delta F = 3.2$ and $\Delta F^{\ddagger} = 6.2$ kcal mol^{-1}).[12] In contrast, the inward conformation was found to be more populated if the side-chain was neutral. The rationale behind these preferences appears to involve the stabilization of a neutral side-chain in close vicinity of the positively charged Zn(II) ion at the bottom of the active site. The results from classical MD simulations on HCA II, shown in Table 5, indicate a preferential population of the inward orientation of the neutral side-chain as before. However, the CpHMD studies predict a much lower discrimination of the two orientational states at pH = 7.0 with a smaller value of $\Delta F = 0.71$ kcal mol^{-1} with a high free

Table 5. Free Energy Change ($\Delta F = F_{inward} - F_{outward}$) and Free Energy Barrier, $\Delta F^{\ddagger}$, along the Side-Chain Dihedral Angle, χ_1, for Its Side-Chain Rotation from Outward to Inward Orientations[a]

system	ΔF (kcal mol^{-1})	$\Delta F^{\ddagger}$ (kcal mol^{-1})
	HCA IX-c	
MD	1.60	6.33
CpHMD (pH = 4.5)	3.34	5.48
CpHMD (pH = 7.0)	0.40	3.90
	HCA II	
MD	−1.43	5.69
CpHMD (pH = 7.0)	0.71	5.47

[a]Only the $N_{\delta1}$-atom of the side-chain is protonated in HCA IX-c (MD and CpHMD at pH = 4.5) and in HCA II (MD and CpHMD at pH = 7.0). Both N-atoms of the side-chain are protonated in HCA IX-c (CpHMD, pH = 7.0).

Table 6. Definition of Primary Order Parameter (OP) and the Stable States A and B for HCA IX-c and HCA II

system	HCA IX-c	HCA II
primary OP	$\chi_1 \equiv N-C_{\alpha}-C_{\beta}-C_{\gamma}$ of His-64	$\chi_1 \equiv N-C_{\alpha}-C_{\beta}-C_{\gamma}$ of His-64
state A (outward)	$-100° < \chi_1 < -20°$	$-100° < \chi_1 < -7°$
state B (inward)	$20° < \chi_1 < 100°$	$-7° < \chi_1 < 100°$

energy barrier of 5.47 kcal mol^{-1} separating them. The different residues surrounding the His-64 (listed in Table 3) and the corresponding changes in their protonation states (as shown in Table 4) induce a change in the electrostatic environment. This leads to a dominant presence of the outward conformation of the His-64 side-chain, which is well

Table 7. Definition of Secondary OPs for HCA IX-c and HCA II[a]

system	selected secondary OPs
HCA IX-c (classical MD)	χ_2 of His-64
	d_1
	d_2
	d_3
HCA IX-c (CpHMD)	χ_2 of His-64
	d_1
	d_2
HCA II (classical MD)	d_1
	d_2
HCA II (CpHMD)	d_1
	d_2

[a]The side-chain dihedral angle of His-64 used is defined as $\chi_2 \equiv C_{\alpha}-C_{\beta}-C_{\gamma}-N_{\delta1}$. The three pair distances used are d_1 between Zn(II) and His64 ($N_{\delta1}$), d_2 between Tyr7(O)−His64($N_{\delta1}$), and d_3 between Trp5 ($C_{\delta1}$)−Asn62 ($N_{\delta2}$).

reflected in the ΔF value for the CpHMD simulated HCA II system. Similar to the observation for HCA IX-c, our results corroborate the report from the high-resolution crystal structures of HCA II at neutral pH, detecting both orientations of the His-64 side-chain.

It is evident from the discussion above that χ_1 may be used as the primary order parameter in our study as it unambiguously distinguishes between the two stable states in both HCA IX-c and HCA II (also reported earlier by Roy et al.[20]). We also introduce a generalized definition of the stable conformational states A (outward) and B (inward) of His-64 as ranges of χ_1 values approximating thermally accessible
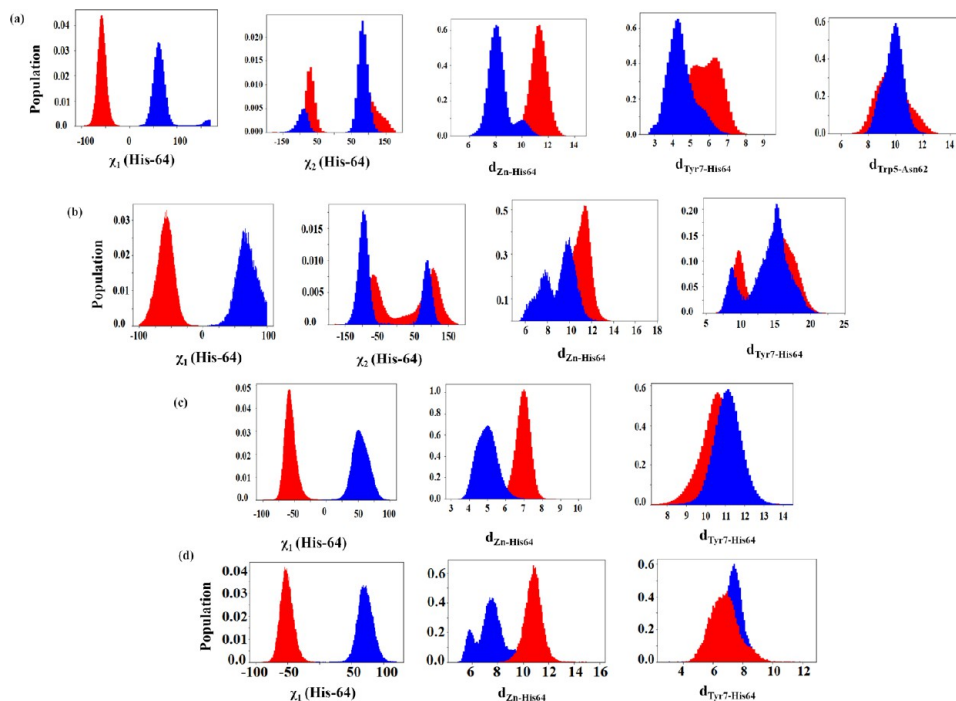

Figure 11. Population distribution plots of (a) χ_1 of His-64, χ_2 of His-64, d_1, d_2, and d_3 at state A (highlighted in red) and state B (highlighted in blue) for HCA IX-c obtained from classical MD simulations, (b) χ_1 of His-64, χ_2 of His-64, d_1 and d_2 at state A (highlighted in red) and state B (highlighted in blue) for HCA IX-c obtained from the CpHMD simulation at pH = 7, (c) χ_1 of His-64, d_1 and d_2 at state A (highlighted in red) and state B (highlighted in blue) for HCA II from the classical MD trajectory, and (d) χ_1 of His-64, d_1 and d_2 at state A (highlighted in red) and state B (highlighted in blue) for HCA II from the CpHMD trajectory at pH = 7.
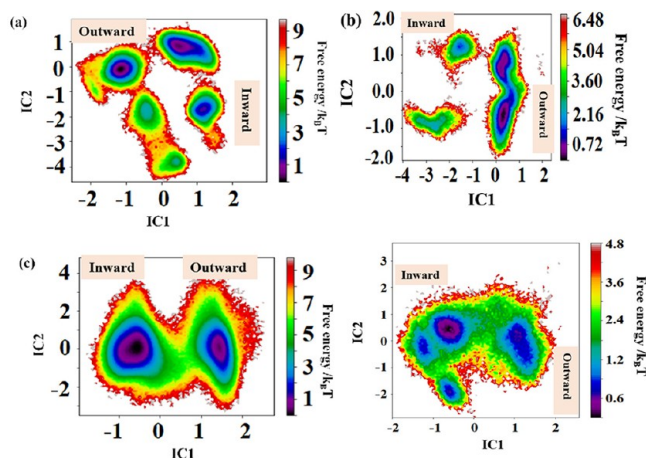
Figure 12. Unweighted free energy landscape obtained from tICA for (a) HCA IX-c (MD), (b) HCA IX-c (CpHMD, pH = 7.0), (c) HCA II (transition path ensemble), and (d) HCA II (CpHMD, pH = 7.0).
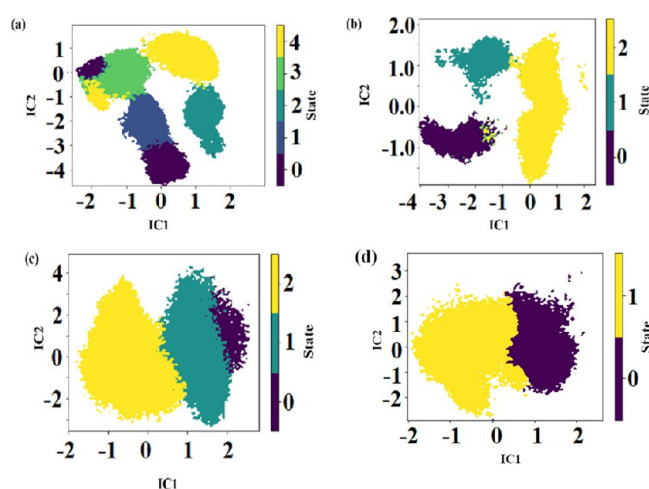

Figure 13. Clustering of the free energy profile into separate macrostates of (a) HCA IX-c from the classical MD simulation, (b) HCA IX-c from the CpHMD simulation at pH = 7, (c) HCA II, and (d) HCA II (CpHMD, pH = 7.0).

regions of the two minima in all five simulations of the two isozymes. Details of the states A and B for HCA IX-c and HCA II have been presented in Table 6. It is also clear that additional order parameters may play a crucial part in

Table 9. Rates Calculated Using Mean First Passage Times (MFPT) for HCA II (in s^{-1}) for (a) Classical MD and (b) CpHMD

		state		
	state	0	1	2
(a)	0	0	7.27×10^8	8.69×10^8
	1	2.60×10^6	0	6.92×10^8
	2	2.59×10^6	3.01×10^8	0

		state	
	state	0	1
(b)	0	0	1.72×10^8
	1	9.14×10^7	0

providing a better description of the free energy and kinetics associated with the side-chain rotation of His-64.

Markov State Models of Conformational Dynamics of His-64 in HCA IX-c and HCA II. We have used the MD and CpHMD (at pH = 7.0) trajectories of HCA IX-c to construct the Markov state models corresponding to His-64 side-chain rotation. As mentioned earlier, the MSM for HCA II with an unprotonated His-64 side-chain has been constructed using the associated transition path ensemble comprised of unbiased MD trajectories. However, a 1 μs long CpHMD trajectory (at pH = 7.0) has been used to construct the corresponding MSM for the transition of the His-64 side-chain in HCA II. In each case, we have investigated several collective variables (CVs) as the system moves between the states A (outward) and B (inward). These CVs include the backbone and side-chain dihedral angles of all amino acid residues experimentally known to be important in the enzyme function and some selected distance parameters relevant to the problem. A list of these collective variables is included in the Supporting Information. As shown in Figure 11a, the primary OP, χ_1, clearly exhibits distinct distributions at states A and B. Apart from the primary OP, the second side-chain dihedral angle, χ_2, of His-64, along with three chosen distances, d_1, d_2, and d_3, are found to have a similar distinctive distribution along HCA IX-c (MD). These distances are defined as (i) d_1, between the Zn(II) and the $N_{\delta 1}$ of His-64; (ii) d_2, between the $N_{\delta 1}$ of His-64 and the O−H of Tyr-7; and (iii) d_3, between the $C_{\delta 1}$ of TRP-5 and the $N_{\delta 2}$ of Asn-62. Similarly, the secondary OPs are detected for HCA IX-c from CpHMD at pH = 7.0 and HCA II (transition path ensemble and CpHMD at pH = 7.0). The respective distributions are shown in Figure 11b−d. A full list

Table 8. Kinetic Rate Constants (in s^{-1}) for HCA IX-c from (a) Classical MD and (b) CpHMD Simulation at pH = 7.0

		state				
	state	0	1	2	3	4
(a)	0	0	1.14×10^8	1.28×10^7	1.23×10^8	6.87×10^7
	1	8.27×10^6	0	1.17×10^7	2.55×10^8	7.52×10^7
	2	5.70×10^6	1.48×10^7	0	3.63×10^7	7.13×10^7
	3	5.94×10^6	2.07×10^7	1.11×10^7	0	8.39×10^7
	4	5.57×10^6	1.60×10^7	1.23×10^7	6.02×10^7	0

		state		
	state	0	1	2
(b)	0	0	9.47×10^7	4.19×10^8
	1	4.72×10^7	0	6.74×10^8
	2	4.14×10^7	7.88×10^7	0

Table 10. Region-wise List of Interacting Residues within 5 Å of Distal His-64 Obtained from CpHMD Trajectories (at pH = 7.0) of HCA IX-c and HCA II

HCA IX-c			HCA II	
region 0	region 1	region 2	region 0	region 1
Tyr-7	Tyr-7	Trp-5	His-4	His-4
Asn-61	Trp-16	Arg-6	Trp-5	Trp-5
Asn-62	Asn-62	Tyr-7	Gly-6	Gly-6
Gly-63	Gly-63	Asn-61	Tyr-7	Tyr-7
Ser-65	Ser-65	Asn-62	Gly-8	Asn-11
Trp-97	Val-66	Gly-63	Asn-11	Trp-16
Glu-170	His-94	Ser-65	Gly-12	Asn-61
Leu-241	Leu-95	Val-66	Pro-13	Asn-62
Asn-244	His-96	Leu-241	Asn-61	Gly-63
	GL4-98	Asn-244	Asn-62	Ala-65
	Thr-200		Gly-63	Phe-66
	Leu-241		Ala-65	Gln-92
	Asn-244		Phe-66	His-94
			His-96	Phe-95
			Lys-170	His-96
			Thr-200	Trp-97
			Phe-231	GL4-106
			Met-241	Lys-170
			Asn-244	Thr-199
				Thr-200
				Pro-201
				Pro-202
				Phe-231
				Met-241
				Asn-244
				Trp-245
				Arg-246

of secondary order parameters (OPs) for all three simulations has been presented in Table 7.

Using their respective order parameters as inputs, we performed tICA for which the corresponding free energy profiles are shown in Figure 12. The projection of underlying free energy for reorientation of the His-64 side-chain along the first two independent components (IC_1 and IC_2) is shown in Figure 12. As expected, several substates could be resolved in the two-dimensional free energy profiles that could not be traced in our earlier one-dimensional study, as shown in Figure 10. The labels "inward" and "outward" have been assigned to the regions comprised predominantly ($\gtrsim$90%) of the corresponding side-chain orientation of His-64. It is evident from Figure 12a and b that the details of underlying free energy surfaces, as sampled along the MD and CpHMD trajectories of HCA IX-c, are quite different not only in terms of the location of free energy basins but also in the magnitudes of barriers separating them. A similar but less marked distinction is observed while considering the corresponding free energy surfaces of HCA II as shown in Figure 12c and d.

The kinetically accessible macro-states for all four trajectories are presented in Figure 13a–d that are identified within the respective MSM. The resultant kinetic rate constants (estimated as the inverse of the respective mean first passage times between these states) are presented in Table 8 for HCA IX-c and in Table 9 for HCA II. Analysis of these results clearly indicates that the slowest transition of a neutral His-64 side-chain in HCA IX-c may be as slow as 6×10^6 s^{-1}, while frequent updating of its protonation state may lead to a

much faster reorientation with a rate constant of 4.14×10^7 s^{-1}. In the case of HCA II, the sensitivity of the rate of side-chain rotation is found to be less significant with the slowest transition with $k = 9.14 \times 10^7$ s^{-1} being recorded from predominantly inward to predominantly outward orientation in CpHMD simulation at pH = 7.0. The other transitions (MD or CpHMD) are found to be marginally faster with $k \approx 10^8$ s^{-1}. It may be noted that if the side-chain of His-64 is protonated in an otherwise neutral MD simulation of HCA II, the corresponding rate constant for the inward to outward orientation may be as fast as 3.2×10^9 s^{-1}.[15] Therefore, the present results, as summarized in Tables 8 and 9, are consistent with our earlier findings.

We have further mapped and presented in Table 10 the region-wise interactions of neighboring residues around catalytically important His-64 in the CpHMD trajectories (at pH = 7.0) of HCA IX-c and HCA II. Apart from some differences in the N-terminal end and important hydrophilic residues, the rotating His-64 side-chain mostly interacts with its immediate neighbors such as Tyr-7, Asn-61, Asn-62, Gly-63, Ser/Ala-65, Leu/Met-241, and Asn-244. It is also noted that protonated Glu residues, GL4-98 for HCA IX-c and GL4-106 for HCA II at region 1, participate in determining the major population of the inward conformer of the His-64 side-chain. Interactions of the rotating His-64 side-chain with Trp-5 and Thr-200 are found while sampling the classical MD trajectory but not along the CpHMD path.

■ CONCLUSION

Although HCAs are long known to be important in pH control and sensing, it is quite surprising that no constant pH simulation study has been reported so far on any of the 15 isozymes known. The results presented in this Article furnish the first ever calibration of the simulation setup for HCA IX-c using both MD and CpHMD simulation studies. The overall structure and stability of the two isozymes are found to be largely similar, irrespective of the simulation method used. However, several local reorganizations are found to be sensitive to the fluctuations of several residues, including the titrable ones. This Article also provides the first ever insight into several dynamical aspects of catalysis by HCA IX-c with a detailed comparison with HCA II. The crucial reorganization step prior to the actual proton transfer step is found to involve important changes in the active site hydration, formation of proton transfer paths, and the rate at which the key His-64 residue would fluctuate. The crystal structure of HCA IX-c detects only one outward side-chain conformation of His-64 that is well understood to be inefficient for the proton shuttle.[77] Our results establish that rapid rotations between its inward and outward orientations are possible. This corroborates the experimentally observed high turnover value of HCA IX-c.

Each of the simulated trajectories in the present work has been extended up to 1 μs. In view of the noted kinetic rate constant, $k_{reorg} \approx 10^7-10^9$ s^{-1} in HCA II, these trajectories are expected to capture the effects of fast reorganizations in both isozymes. Accordingly, no enhanced sampling method (except for HCA II under charge neutralized conditions) was needed to construct the MSMs and estimate the associated free energy and kinetics. However, the estimated pK_a for the His-64 side-chain in HCA II is substantially lower as compared to the earlier reported value of 7.2 ± 0.1 obtained using nonlinear fitting of the ^{18}O-exchange kinetics experiment data on the

catalysis by wild-type HCA II.[77] It, therefore, appears that enhanced sampling methods such as replica exchange CpHMD may be necessary for our future goal of estimating the pK_a values of all titrable residues that can potentially couple to the reactive motions across the active sites of HCA II and IX-c. The in-depth bench-marking presented in this Article in terms of equivalent simulations performed for both HCA II and IX-c also gives us the confidence to extend these studies to investigate the pH-dependent stability of these two isozymes. It is also important to note that we have not carried out any QM−MM MD-based analysis as our focus in this Article was on the nonreactive dynamics at and around the active site of the two isozymes. It will certainly be interesting to probe how the enzyme structure and dynamics would couple to the actual proton transfer step, and we hope to report on this in the future. Finally, as shown in our previous study[19] on the most accurate reaction coordinate for multiple steps of catalysis by HCA II using charge neutralized systems, the kinetic rate constants estimated for different elementary steps help in validating alternative mechanistic pathways for different routes involving conformational fluctuations and water reorganization spanning multiple length and timescales. Although one such microkinetic scheme could be optimized at neutral pH for HCA II, we are far from compiling one for HCA IX-c, where we also need to carry out a molecular simulation of the actual intramolecular proton transfer step at the catalytic region. We strongly believe that this Article will be useful in triggering a wide range of activities in the investigations on these therapeutically important isozymes of HCA.

ASSOCIATED CONTENT

Data Availability Statement

The trajectories were generated by using the AMBER 18 MD suite (https://ambermd.org/). The Markov state models were built with PyEMMA 2 (http://emma-project.org/latest/). Because of the large volume of data corresponding to long time scale simulations, full trajectories will be provided upon request.

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c03356.

> RMSD, RMSF, R_g, B-factor, secondary structure classifications, order parameters (OPs), time−structure independent components (TICs) correlation, autocorrelation functions, implied time scale (ITS), and Chapman−Kolmogorov (CK) test (PDF)

AUTHOR INFORMATION

Corresponding Author

Srabani Taraphder − Department of Chemistry, Indian Institute of Technology, Kharagpur 721302, India; orcid.org/0000-0003-3330-6033; Email: srabani@chem.iitkgp.ac.in

Authors

Divya Rai − Department of Chemistry, Indian Institute of Technology, Kharagpur 721302, India

Satyajit Khatua − Department of Chemistry, Indian Institute of Technology, Kharagpur 721302, India; orcid.org/0000-0002-2734-3965

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.2c03356

Author Contributions

D.R. carried out the simulations on HCA IX-c and S.K. performed the same for HCA II. Both D.R. and S.K. contributed equally to the research and analysis. D.R., S.K., and S.T. designed the project and wrote the paper.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Matulis, D. *Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding*; Springer International Publishing: New York, 2019.

(2) Supuran, C.; Nocentini, A. *Carbonic Anhydrases: Biochemistry and Pharmacology of an Evergreen Pharmaceutical Target*; Elsevier Science: New York, 2019.

(3) Frost, S., McKenna, R. *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications; Subcellular Biochemistry*; Springer: Netherlands, 2013.

(4) Alterio, V.; Hilvo, M.; Di Fiore, A.; Supuran, C. T.; Pan, P.; Parkkila, S.; Scaloni, A.; Pastorek, J.; Pastorekova, S.; Pedone, C.; et al. Crystal structure of the catalytic domain of the tumor-associated human carbonic anhydrase IX. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16233−16238.

(5) Opavský, R.; Pastoreková, S.; Zelník, V.; Gibadulinová, A.; Stanbridge, E. J.; Závada, J.; Kettmann, R.; Pastorek, J. Human MN/CA9Gene, a Novel Member of the Carbonic Anhydrase Family: Structure and Exon to Protein Domain Relationships. *Genomics* **1996**, *33*, 480−487.

(6) Supuran, C. T.; Winum, J.-Y. Carbonic anhydrase IX inhibitors in cancer therapy: an update. *Future Med. Chem.* **2015**, *7*, 1407−1414.

(7) Fisher, S. Z.; Maupin, C. M.; Budayova-Spano, M.; Govindasamy, L.; Tu, C.; Agbandje-McKenna, M.; Silverman, D. N.; Voth, G. A.; McKenna, R. Atomic crystal and molecular dynamics simulation structures of human carbonic anhydrase II: insights into the proton transfer mechanism. *Biochemistry* **2007**, *46*, 2930−2937.

(8) De Simone, G.; Supuran, C. T. Carbonic anhydrase IX: Biochemical and crystallographic characterization of a novel antitumor target. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 404−409.

(9) Pastorek, J.; Pastoreková, S.; Callebaut, I.; Mornon, J.; Zelník, V.; Opavský, R.; Zaťovicová, M.; Liao, S.; Portetelle, D.; Stanbridge, E. Cloning and characterization of MN, a human tumor-associated protein with a domain homologous to carbonic anhydrase and a putative helix-loop-helix DNA binding segment. *Oncogene* **1994**, *9*, 2877−2888.

(10) Kazokaitė, J.; Niemans, R.; Dudutienė, V.; Becker, H. M.; Leitāns, J.; Zubrienė, A.; Baranauskienė, L.; Gondi, G.; Zeidler, R.; Matulienė, J.; et al. Novel fluorinated carbonic anhydrase IX inhibitors reduce hypoxia-induced acidification and clonogenic survival of cancer cells. *Oncotarget* **2018**, *9*, 26800−26816.

(11) Mahon, B. P.; Bhatt, A.; Socorro, L.; Driscoll, J. M.; Okoh, C.; Lomelino, C. L.; Mboge, M. Y.; Kurian, J. J.; Tu, C.; Agbandje-

McKenna, M.; et al. The Structure of Carbonic Anhydrase IX Is Adapted for Low-pH Catalysis. *Biochemistry* **2016**, *55*, 4642−4653.

(12) Maupin, C. M.; Voth, G. A. Preferred Orientations of His64 in Human Carbonic Anhydrase II. *Biochemistry* **2007**, *46*, 2938−2947.

(13) Maupin, C. M.; McKenna, R.; Silverman, D. N.; Voth, G. A. Elucidation of the Proton Transport Mechanism in Human Carbonic Anhydrase II. *J. Am. Chem. Soc.* **2009**, *131*, 7598−7608.

(14) Roy, A.; Taraphder, S. Identification of Proton-Transfer Pathways in Human Carbonic Anhydrase II. *J. Phys. Chem. B* **2007**, *111*, 10563−10576.

(15) Paul, S.; Paul, T. K.; Taraphder, S. Reaction Coordinate, Free Energy, and Rate of Intramolecular Proton Transfer in Human Carbonic Anhydrase II. *J. Phys. Chem. B* **2018**, *122*, 2851−2866.

(16) Paul, S.; Taraphder, S. Determination of the Reaction Coordinate for a Key Conformational Fluctuation in Human Carbonic Anhydrase II. *J. Phys. Chem. B* **2015**, *119*, 11403−11415.

(17) Taraphder, S.; Maupin, C. M.; Swanson, J. M. J.; Voth, G. A. Coupling Protein Dynamics with Proton Transport in Human Carbonic Anhydrase II. *J. Phys. Chem. B* **2016**, *120*, 8389−8404.

(18) Paul, T. K.; Taraphder, S. Coordination Dynamics of Zinc Triggers the Rate Determining Proton Transfer in Human Carbonic Anhydrase II. *ChemPhysChem* **2020**, *21*, 1455−1473.

(19) Paul, T. K.; Taraphder, S. Nonlinear Reaction Coordinate of an Enzyme Catalyzed Proton Transfer Reaction. *J. Phys. Chem. B* **2022**, *126*, 1413−1425.

(20) Roy, A.; Roy, A.; Taraphder, S. Transition path sampling study of the conformational fluctuation of His-64 in human carbonic anhydrase II. *J. Phys. Chem. B* **2009**, *113*, 12555−12564.

(21) Paul, T. K.; Taraphder, S. Molecular modelling of two coordination states of Zn(II) ion at the active site of human carbonic anhydrase II. *Chem. Phys.* **2021**, *549*, 111281.

(22) Supuran, C.; De Simone, G. *Carbonic Anhydrases as Biocatalysts: From Theory to Medical and Industrial Applications*; Elsevier Science: New York, 2015.

(23) Supuran, C. T. Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. *Nat. Rev. Drug Discovery* **2008**, *7*, 168−181.

(24) Angeli, A.; Del Prete, S.; Alasmary, F. A.; Alqahtani, L. S.; AlOthman, Z.; Donald, W. A.; Capasso, C.; Supuran, C. T. The first activation studies of the η-carbonic anhydrase from the malaria parasite Plasmodium falciparum with amines and amino acids. *Bioorg. Chem.* **2018**, *80*, 94−98.

(25) Capasso, C.; Supuran, C. T. An overview of the alpha-, beta- and gamma-carbonic anhydrases from Bacteria: can bacterial carbonic anhydrases shed new light on evolution of bacteria? *J. Enzyme Inhib. Med. Chem.* **2015**, *30*, 325−332.

(26) Tafreshi, N. K.; Lloyd, M. C.; Bui, M. M.; Gillies, R. J.; Morse, D. L. *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*; Springer: Netherlands, 2014; pp 221−254.

(27) Kato, Y.; Ozawa, S.; Miyamoto, C.; Maehata, Y.; Suzuki, A.; Maeda, T.; Baba, Y. Acidic extracellular microenvironment and cancer. *Cancer Cell Int.* **2013**, *13*, 89.

(28) Švastová, E.; Hulíková, A.; Rafajová, M.; Zaťovičová, M.; Gibadulinová, A.; Casini, A.; Cecchi, A.; Scozzafava, A.; Supuran, C. T.; Pastorek, J.; et al. Hypoxia activates the capacity of tumor-associated carbonic anhydrase IX to acidify extracellular pH. *FEBS Lett.* **2004**, *577*, 439−445.

(29) Pastorek, J.; Pastorekova, S.; Zatovicova Cancer-Associated, M. Carbonic Anhydrases and Their Inhibition. *Curr. Pharm. Des.* **2008**, *14*, 685−698.

(30) Supuran, C. T. Carbonic anhydrase inhibitors: an update on experimental agents for the treatment and imaging of hypoxic tumors. *Expert Opin. Invest. Drugs* **2021**, *30*, 1197−1208.

(31) Nerella, S. G.; Singh, P.; Arifuddin, M.; Supuran, C. T. Anticancer carbonic anhydrase inhibitors: a patent and literature update 2018−2022. *Expert Opin. Ther. Pat.* **2022**, *32*, 833−847.

(32) Abdelrahman, M. A.; Ibrahim, H. S.; Nocentini, A.; Eldehna, W. M.; Bonardi, A.; Abdel-Aziz, H. A.; Gratteri, P.; Abou-Seri, S. M.; Supuran, C. T. Novel 3-substituted coumarins as selective human

carbonic anhydrase IX and XII inhibitors: Synthesis, biological and molecular dynamics analysis. *Eur. J. Med. Chem.* **2021**, *209*, 112897.

(33) Sharonova, T.; Paramonova, P.; Kalinin, S.; Bunev, A.; Gasanov, R.; Nocentini, A.; Sharoyko, V.; Tennikova, T. B.; Dar'in, D.; Supuran, C. T.; et al. Insertion of metal carbenes into the anilinic N-H bond of unprotected aminobenzenesulfonamides delivers low nanomolar inhibitors of human carbonic anhydrase IX and XII isoforms. *Eur. J. Med. Chem.* **2021**, *218*, 113352.

(34) Mboge, M. Y.; Chen, Z.; Wolff, A.; Mathias, J. V.; Tu, C.; Brown, K. D.; Bozdag, M.; Carta, F.; Supuran, C. T.; McKenna, R.; et al. Selective inhibition of carbonic anhydrase IX over carbonic anhydrase XII in breast cancer cells using benzene sulfonamides: Disconnect between activity and growth inhibition. *PloS one* **2018**, *13*, e0207417.

(35) Mboge, M. Y.; Mahon, B. P.; Lamas, N.; Socorro, L.; Carta, F.; Supuran, C. T.; Frost, S. C.; McKenna, R. Structure activity study of carbonic anhydrase IX: Selective inhibition with ureido-substituted benzenesulfonamides. *Eur. J. Med. Chem.* **2017**, *132*, 184−191.

(36) Mahmud, S.; Rahman, E.; Nain, Z.; Billah, M.; Karmakar, S.; Mohanto, S. C.; Paul, G. K.; Amin, A.; Acharjee, U. K.; Saleh, M. A. Computational discovery of plant-based inhibitors against human carbonic anhydrase IX and molecular dynamics simulation. *J. Biomol. Struct. Dyn.* **2021**, *39*, 2754−2770.

(37) Chahal, V.; Nirwan, S.; Pathak, M.; Kakkar, R. Identification of potent human carbonic anhydrase IX inhibitors: a combination of pharmacophore modeling, 3D-QSAR, virtual screening and molecular dynamics simulations. *J. Biomol. Struct. Dyn.* **2020**, 1−16.

(38) Gopinath, P.; Kathiravan, M. Molecular insights of oxadiazole benzene sulfonamides as human carbonic anhydrase IX inhibitors: Combined molecular docking, molecular dynamics, and 3D QSAR studies. *J. Indian Chem. Soc.* **2022**, *99*, 100339.

(39) John, A.; Sivashanmugam, M.; Umashankar, V.; Natarajan, S. K. Virtual screening, molecular dynamics, and binding free energy calculations on human carbonic anhydrase IX catalytic domain for deciphering potential leads. *J. Biomol. Struct. Dyn.* **2017**, *35*, 2155−2168.

(40) Idrees, D.; Rahman, S.; Shahbaaz, M.; Haque, M. A.; Islam, A.; Ahmad, F.; Hassan, M. I. Estimation of thermodynamic stability of human carbonic anhydrase IX from urea-induced denaturation and MD simulation studies. *Int. J. Biol. Macromol.* **2017**, *105*, 183−189.

(41) Prakash, A.; Idrees, D.; Haque, M. A.; Islam, A.; Ahmad, F.; Hassan, M. I. GdmCl-induced unfolding studies of human carbonic anhydrase IX: a combined spectroscopic and MD simulation approach. *J. Biomol. Struct. Dyn.* **2017**, *35*, 1295−1306.

(42) Giovannuzzi, S.; Capasso, C.; Nocentini, A.; Supuran, C. T. Continued Structural Exploration of Sulfocoumarin as Selective Inhibitor of Tumor-Associated Human Carbonic Anhydrases IX and XII. *Molecules* **2022**, *27*, 4076.

(43) Koruza, K.; Lafumat, B.; Nyblom, M.; Mahon, B.; Knecht, W.; McKenna, R.; Fisher, S. Structural comparison of protiated, H/D-exchanged and deuterated human carbonic anhydrase IX. *Acta Crystallogr., Sect. D: Struct. Biol.* **2019**, *75*, 895−903.

(44) Bua, S.; Lomelino, C.; Murray, A. B.; Osman, S. M.; ALOthman, Z. A.; Bozdag, M.; Abdel-Aziz, H. A.; Eldehna, W. M.; McKenna, R.; Nocentini, A.; et al. "A Sweet Combination": Developing Saccharin and Acesulfame K Structures for Selectively Targeting the Tumor-Associated Carbonic Anhydrases IX and XII. *J. Med. Chem.* **2020**, *63*, 321−333.

(45) Aggarwal, M.; Boone, C. D.; Kondeti, B.; McKenna, R. Structural annotation of human carbonic anhydrases. *J. Enzyme Inhib. Med. Chem.* **2013**, *28*, 267−277.

(46) Alterio, V.; Di Fiore, A.; D'Ambrosio, K.; Supuran, C. T.; De Simone, G. Multiple binding modes of inhibitors to carbonic anhydrases: how to design specific drugs targeting 15 different isoforms? *Chem. Rev.* **2012**, *112*, 4421−4468.

(47) Supuran, C. T. Structure-based drug discovery of carbonic anhydrase inhibitors. *J. Enzyme Inhib. Med. Chem.* **2012**, *27*, 759−772.

(48) Zakšauskas, A.; Čapkauskaitė, E.; Paketurytė-Latvė, V.; Smirnov, A.; Leitans, J.; Kazaks, A.; Dvinskis, E.; Stančaitis, L.;

Mickeviči ūtė, A.; Jachno, J.; et al. Methyl 2-Halo-4-Substituted-5-Sulfamoyl-Benzoates as High Affinity and Selective Inhibitors of Carbonic Anhydrase IX. *Int. J. Mol. Sci.* **2022**, *23*, 130.

(49) Häkansson, K.; Carlsson, M.; Svensson, L.; Liljas, A. Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes. *J. Mol. Biol.* **1992**, *227*, 1192−1204.

(50) Pinard, M. A.; Aggarwal, M.; Mahon, B. P.; Tu, C.; McKenna, R. A sucrose-binding site provides a lead towards an isoform-specific inhibitor of the cancer-associated enzyme carbonic anhydrase IX. *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **2015**, *71*, 1352−1358.

(51) Case, D.; Babin, V.; Berryman, J.; Betz, R.; Cai, Q.; Cerutti, D.; Cheatham, T., III; Darden, T.; Duke, R.; Gohlke, H.; et al. *Amber 14*; University of California: San Francisco, CA, 2014.

(52) Peters, M. B.; Yang, Y.; Wang, B.; Füsti-Molnár, L.; Weaver, M. N.; Merz, K. M. Structural Survey of Zinc-Containing Proteins and Development of the Zinc AMBER Force Field (ZAFF). *J. Chem. Theory Comput.* **2010**, *6*, 2935−2947.

(53) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.

(54) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(55) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N′-methylamide. *Biopolymers* **1992**, *32*, 523−535.

(56) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **2001**, *22*, 501−508.

(57) Roberts, E.; Eargle, J.; Wright, D.; Luthey-Schulten, Z. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinf.* **2006**, *7*, 382.

(58) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.

(59) O'Donoghue, P.; Luthey-Schulten, Z. On the Evolution of Structure in Aminoacyl-tRNA Synthetases. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 550−573.

(60) Eastwood, M. P.; Hardin, C.; Luthey-Schulten, Z.; Wolynes, P. G. Evaluating protein structure-prediction schemes using energy landscape theory. *IBM J. Res. Dev.* **2001**, *45*, 475−497.

(61) Torrie, G.; Valleau, J. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(62) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(63) Glielmo, A.; Husic, B. E.; Rodriguez, A.; Clementi, C.; Noé, F.; Laio, A. Unsupervised Learning Methods for Molecular Simulation Data. *Chem. Rev.* **2021**, *121*, 9722−9758.

(64) Husic, B. E.; Pande, V. S. Markov State Models: From an Art to a Science. *J. Am. Chem. Soc.* **2018**, *140*, 2386−2396.

(65) Pande, V. S.; Beauchamp, K.; Bowman, G. R. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods* **2010**, *52*, 99−105.

(66) Chodera, J. D.; Noé, F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **2014**, *25*, 135−144.

(67) Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernández, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J.-H.; Noé, F. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **2015**, *11*, 5525−5542.

(68) Pérez-Hernández, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noé, F. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys.* **2013**, *139*, 015102.

(69) Schultze, S.; Grubmüller, H. Time-Lagged Independent Component Analysis of Random Walks and Protein Dynamics. *J. Chem. Theory Comput.* **2021**, *17*, 5766−5776.

(70) Forgy, E. Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* **1965**, *21*, 768−780.

(71) MacQueen, J. B. Some Methods for Classification and Analysis of MultiVariate Observations. *Proc. 5th Berkeley Symp. Math. Stat. Prob.* **1967**, 281−297.

(72) Prinz, J.-H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schütte, C.; Noé, F. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **2011**, *134*, 174105.

(73) Polizzi, N. F.; Therien, M. J.; Beratan, D. N. Mean First-Passage Times in Biology. *Isr. J. Chem.* **2016**, *56*, 816−824.

(74) Eriksson, A. E.; Jones, T. A.; Liljas, A. Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 274−282.

(75) Halder, P.; Taraphder, S. Identification of putative unfolding intermediates of the mutant His-107-tyr of human carbonic anhydrase II in a multidimensional property space. *Proteins: Struct., Funct., Bioinf.* **2016**, *84*, 726−743.

(76) Heinig, M.; Frishman, D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* **2004**, *32*, W500−W502.

(77) Zheng, J.; Avvaru, B. S.; Tu, C.; McKenna, R.; Silverman, D. N. Role of Hydrophilic Residues in Proton Transfer during Catalysis by Human Carbonic Anhydrase II. *Biochemistry* **2008**, *47*, 12028−12036.