Active Species for the Ground-State Complex of Cytidine Deaminase: A Linear-Scaling Quantum Mechanical Investigation

James P. Lewis,*,† Charles W. Carter, Jr.,† Jan Hermans,† Wei Pan,‡ Tai-Sung Lee,‡ and Weitao Yang‡

Contribution from the Department of Biochemistry and Biophysics, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27599-7260, and Department of Chemistry, Duke University, Durham, North Carolina 27708-0354

Received October 8, 1997. Revised Manuscript Received March 18, 1998

Abstract: We present results of large-scale (1330 atoms) linear-scaling quantum mechanical semiempirical (PM3) simulations done to optimize geometries surrounding the active site within the enzyme cytidine deaminase. We make a strong prediction about the structure of the active site for the active species, based on the energetics of the calculated structures and comparisons to X-ray crystallographic data. The lowest energy structure indicates that $Zn-OH^-$ is the active species formed prior to nucleophilic attack of the ligand, that the active species of Glu-104 is with $O^{\epsilon 2}$ protonated and hydrogen-bonded with N^3 of the ligand, and that the C^4 and OH^- atoms are significantly closer than is permitted by their van der Waals radii. In addition, we predict structures corresponding to the low-pH and high-pH states in the active site of the enzyme.

Introduction

Understanding the catalytic mechanisms by which enzymes operate is one of the more challenging issues within molecular biology. Considerable effort has been devoted to study the catalytic mechanisms of the nucleoside deaminases cytidine deaminase (CDA) and adenosine deaminase.[1] The enzyme CDA is an efficient catalyst which accelerates the rate of hydrolytic deamination of cytidine to uridine by roughly 11 orders of magnitude.[2] Understanding the catalytic mechanisms for CDA and determining stable inhibitors could have important potential therapeutic uses, since it is responsible for degradation of several important cytidine-based antitumor agents such as cytidine arabinoside.[3] Available X-ray crystallographic studies include several crystal structures of stable intermediate CDA states achieved using different analogue complexes to represent the ground state,[1l] pre-transition state,[1j] transition state,[1h] and product.[1m] The experimental data from these studies have begun to suggest important subtleties in the configuration of a catalyzed active zinc ion and in conformational differences between complexes. The C^4 position on the pyrimidine ring moves ∼1.5 Å closer to the Zn-activated nucleophile while the amino group moves away from the nucleophile, with the N^4-C^4 distance increasing from 1.33 Å in the ground state to approximately 2.80 Å in the product state.[1m] The enzyme active site apparently changes structure to accommodate each successive ligand in the sequence of structures.

However, many of these subtleties lie at or beyond the limit normally considered to be the noise level of an X-ray diffraction experiment. To verify their significance, higher resolution must be obtained and theoretical tools must be employed to complement the experimental procedures. A quantum mechanical (QM) description of the sequence of structures formed during catalysis has long been a goal to help mechanistic enzymologists understand reaction pathways. Until recently, progress toward this goal has been frustrated by inadequate computational methods and by the lack of a good experimental system with which to calibrate the calculations. The problem is that any theoretical technique employed needs to give a valid description for much of the protein environment surrounding the active site. Recent approaches by others to overcome this problem are discussed in the next section.

The CDA active site includes a Zn ion bound by the thiolate sulfur atoms of the Cys-129 and Cys-132 side chains and by a nitrogen atom of the His-102 side chain (see Figure 1). Bound to this Zn complex is a OH^- or water molecule which displaces the NH_2 group, located at the C^4 position on the cytidine ligand, via nucleophilic hydration of the N^4-C^4 double bond. The product of this displacement is apparently the conventional keto tautomer of uridine. The carboxylate group of Glu 104 near the NH_2 group on the cytidine ligand assists in the reaction both by stabilizing the tetrahedral transition state and by acting as a proton shuttle.[1k−m]

The purpose of calculations presented here is to understand the nature of the enzyme−cytidine ground-state complex more

† University of North Carolina.

‡ Duke University.

(1) (a) Cohen, R. M.; Wolfenden, R. *J. Biol. Chem.* **1971**, *246*, 7561. (b) Cohen, R. M.; Wolfenden, R. *J. Biol. Chem.* **1971**, *246*, 7566. (c) Marquez, V. E.; Liu, P. S.; Kelley, J. A.; Driscoll, J. S.; McCormick, J. J. *J. Med. Chem.* **1980**, *23*, 713. (d) Liu, P. S.; Marquez, V. E.; Driscoll, J. S.; Fuller, R. W.; McCormick, J. J. *J. Med. Chem.* **1981**, *24*, 662. (e) Ashley, G. W.; Bartlett, P. A. *J. Biol. Chem.* **1984**, *259*, 13621. (f) Kim, C. H.; Marquez, V. E.; Mao, D. T.; Haines, D. R.; McCormick, J. J. *J. Med. Chem.* **1986**, *29*, 1374 (g) Wilson, D. K.; Rudolph, F. B.; Quiocho, F. A. *Science* **1991**, *252*, 1278. (h) Betts, L.; Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, Jr., C. W. *J. Mol. Biol.* **1994**, *235*, 635. (i) Carlow, D.; Smith, A. A.; Yang, C. C.; Short, S. A.; Wolfenden, R. *Biochemistry* **1995**, *34*, 4220. (j) Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, Jr., C. W., *Biochemistry* **1995**, *34*, 4516. (k) Carlow, D.; Short, S. A.; Wolfenden, R. *Biochemistry* **1996**, *35*, 948. (l) Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, Jr., C. W., *Biochemistry* **1996**, *35*, 1335. (m) Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, Jr., C. W. *Biochemistry* **1997**, *36*, 4768.

(2) Frick, L.; MacNeela, J. P.; Wolfenden, R. *Bioorg. Chem.* **1987**, *15*, 100.

(3) (a) Marquez, V. E. In *Developments in Cancer Chemotherapy*; Glazer, R. I., Ed.; CRC Press: Florida, 1984; pp 91−114 and references cited within. (b) Chandrasekaren, B.; Capizzi, R. L.; Kute, T. E.; Morgan, T.; Dimling, J. *Cancer Res.* **1989**, *49*, 3259.

S0002-7863(97)03522-1 CCC: $15.00
Published on Web 05/20/1998

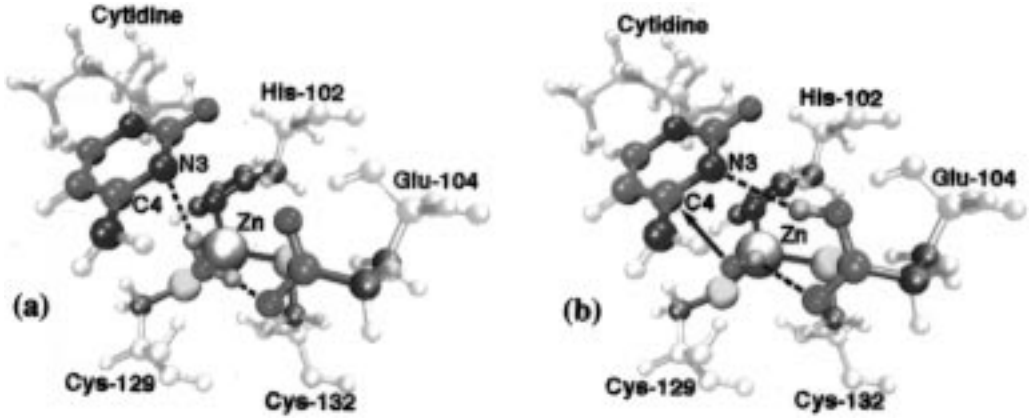

Figure 1. Two structures for the active site of cytidine deaminase, whose geometries were optimized via divide-and-conquer within a semiempirical (PM3) framework. (a) Ground-state substrate complex just before the active species is created by deprotonation of $Zn-H_2O$. (b) This structure has the lowest energy of the structures calculated and, as verified by experimental data, is the active species. A hydrogen-bonding network is formed as a result of deprotonating the substrate H_2O and protonation of the carboxyl oxygen $O^{\epsilon 2}$ of Glu-104.

fully, and specifically to specify the structure of the active species at the initiation of the reaction pathway. This computational investigation addresses two important issues: (1) whether the active species consists of a zinc-coordinated hydroxide ion ($Zn-OH^-$) or a zinc-coordinated water molecule ($Zn-H_2O$); (2) which of the two carboxylate oxygen atoms of Glu-104 is protonated in the active species.

With regard to the first issue, it is commonly asserted that for zinc enzymes the most likely mechanism for hydration involves direct attack of $Zn-OH^-$ rather than $Zn-H_2O$. Much circumstantial evidence suggests participation of $Zn-OH^-$ in catalysis by zinc hydrolases.[4] The experimental data are most compelling for carbonic anhydrase where extensive data tend to rule out mechanisms other than via $Zn-OH^-$; experimental data also support this mechanism for other zinc enzymes such as carboxypeptidase A, alkaline phosphatase, and alcohol dehydrogenase. Ab initio Hartree−Fock calculations for a variety of gas-phase Zn complexes, including $Zn(SCH_3)_2(NH_3)OH^-$ and $Zn(SCH_3)_2(NH_3)H_2O$ which resemble the Zn coordination in the active site of CDA, show that the active species of zinc enzymes most likely consist of $Zn-OH^-$ since these complexes are more nucleophilic than $Zn-H_2O$.[5] While experimental and theoretical data suggest that in general the active species for zinc enzymes is $Zn-OH^-$, no specific conclusions about the zinc enzyme CDA can be made without further specific investigation. Our results provide a theoretical investigation where much of the protein environment surrounding the active site is included.

With regard to the second issue, the pH dependence of activity suggests that the carboxylate group of Glu-104 plays an important role in the catalytic process by serving as a general base to abstract a proton from the $Zn(Cys)_2(His)H_2O$ group. As supported by structural data of CDA transition-state analog[1h] and product[1m] complexes, Glu-104 then acts as a proton shuttle from the $Zn-H_2O$ to the N^3 of the ligand, forming the pre-transition state.[1l] The specific protonated form of Glu-104 in the active species, after the deprotonation of $Zn-H_2O$ and before protonation of N^3 on the ligand, cannot be determined from X-ray diffraction experiments. We have therefore examined complexes containing different protonated forms of Glu-104, each with the deprotonated form of $Zn-H_2O$ to determine the structure of the active species within CDA and to understand the participation of Glu-104 in the mechanism of proton transfer.

Computational Methods

A common approach for including a valid description of the protein environment surrounding the active site has been to combine QM methods with molecular mechanical (MM) force fields.[6] The basic algorithms of QM/MM methods were laid out in the seminal paper of Warshel and Levitt,[6a] and recently, there has been very encouraging progress in the further development of such methods.[7] While QM/MM methods are promising, there remain difficulties in treating the boundaries between the QM and MM subsystems. The computational chemistry community is well aware of the issues, and these difficulties should be overcome in due time. As a possible alternative, we have taken advantage of linear-scaling algorithms which allow larger QM simulations to be performed using purely QM methods. Performing larger QM simulations to account for a significant portion of the protein environment surrounding the active site has long been a goal in quantum chemistry.

Newly developed linear-scaling electronic structure methods allow QM simulations involving thousands of atoms to be performed on a workstation.[8] In the divide-and-conquer (DAC) approach, the system to be calculated is partitioned into smaller subsystems and the electronic

(4) Coleman, J. E. In *Zinc Enzymes*; Maret, W., Zeppezauer, M., Eds.; Birkhäuser: Boston, 1986.

(5) Bertini, I.; Luchinat, C.; Rosi, M.; Sgamellotti, A.; Tarantelli, F. *Inorg. Chem.* **1990**, *29*, 1460.

(6) (a) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227. (b) Singh, U. C.; Kollman, P. *J. Comput. Chem.* **1986**, *7*, 718. (c) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700. (d) Gao, J.; Xia, X. *Science* **1992**, *258*, 631.

(7) (a) Warshel, A. *Curr. Opin. Struct. Biol.* **1992**, *2*, 230. (b) Stanton, R. V.; Hartsough, D. S.; Merz Jr., K. M. *J. Phys. Chem.* **1993**, *97*, 11868. (c) Wei, D.; Salahub, D. R. *Chem. Phys. Lett.* **1994**, *224*, 291. (d) Thery, V.; Rinaldi, D.; Rivail, J.-L. *J. Comput. Chem.* **1994**, *15*, 269. (e) Gao, J. *Review in Computational Chemistry*; VCH: New York, 1995; Vol. 7, p 119. (f) Thompson, M. A.; Schenter, G. K. *J. Phys. Chem.* **1995**, *99*, 6374. (g) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170. (h) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959. (i) Assfeld, X.; Rivail, J.-L. *Chem. Phys. Lett.* **1996**, *263*, 100. (j) Merz Jr., K. M.; Banci, L. *J. Am. Chem. Soc.* **1997**, *119*, 863. (k) Peräkyla, M.; Kollmann, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 1189.

(8) (a) Baroni, S.; Giannozzi, P. *Europhys. Lett.* **1992**, *17*, 547. (b) Galli, G.; Parinello, M. *Phys. Rev. Lett.* **1992**, *69*, 3547. (c) Mauri, F.; Galli, G.; Car, R. *Phys. Rev. B* **1993**, *47*, 9973. (d) Li, X.-P.; Nunes, W.; Vanderbilt, D. *Phys. Rev. B* **1993**, *47*, 10891. (e) Daw, M. S. *Phys. Rev. B* **1993**, *47*, 10899. (f) Ordejón, P.; Drabold, D. A.; Grumbach, M. P.; Martin, R. M. *Phys. Rev. B* **1993**, *48*, 14646. (g) Stechel, E. B.; Williams, A. P.; Feibleman, P. J. *Phys. Rev. B* **1993**, *49*, 3898. (h) Drabold, D. A.; Sankey, O. F. *Phys. Rev. Lett.* **1993**, *70*, 3631. (i) Kohn, W. *Chem. Phys. Lett.* **1993**, *208*, 167. (j) Goedecker, S.; Colombo, L. *Phys. Rev. Lett.* **1994**, *73*, 122. (k) Hierse, W.; Stechel, E. B. *Phys. Rev. B* **1994**, *50*, 17811. (l) Ordejón, P.; Drabold, D. A.; Martin, R. M.; Grumbach, M. P. *Phys. Rev. B* **1995**, *51*, 1456. (m) Hernandez, E.; M. J. Gillan. *Phys. Rev. B* **1995**, *51*, 10157. (n) Kim, J.; Mauri, F.; Galli, G. *Phys. Rev. B* **1995**, *52*, 1640. (o) Stewart, J. P. P. *Int. J. Quantum Chem.* **1996**, *58*, 133.

Table 1. Comparisons of Certain Distances (Å) and Energies (kcal/mol) for Different Optimized Cytidine Deaminase Subsystems[a]

system	energy (kcal/mol)	$d_{Zn-O(OH)}$	$d_{O(OH)-C4(CTD)}$	$d_{OE2(Glu-104)-N3(CTD)}$	$d_{O(THR127)-N4(CTD)}$
experiment[11]		1.89	3.03	3.23	2.74
s0, Zn−H_2O, Glu-104 unprotonated	−6007.16	2.28	3.47	3.62	3.15
s1, Zn−OH^-, $O^{\epsilon 2}$ of Glu-104 protonated	**−6073.11**	**1.99**	**3.26**	**3.62**	**3.075**
s2, Zn−OH^{2-}, $O^{\epsilon 2}$ of Glu-104 protonated	−6026.87	1.97	3.50	3.61	3.12
s3, Zn−OH^-, $O^{\epsilon 1}$ of Glu-104 protonated	−6007.71	1.97	3.36	3.73	3.14
s4, Zn−OH^{2-}, $O^{\epsilon 1}$ of Glu-104 protonated	−6021.12	1.97	3.30	3.76	3.14
t1, Zn−OH^-, Glu-104 unprotonated	−5951.59	1.94	3.47	3.62	3.15
t2, Zn−OH^{2-}, Glu-104 unprotonated	−5967.70	1.88	3.50	3.64	3.57
u1, Zn−H_2O, $O^{\epsilon 2}$ of Glu-104 protonated	−6027.40	2.40	3.25	3.52	2.76

[a] The different structures represent different protonation states of Zn−H_2O (either H^1 or H^2 is deprotonated and Glu-104 (either $O^{\epsilon 1}$ or $O^{\epsilon 2}$ is protonated). Structure s1, shown in Figure 1b, is prominently the lowest energy structure and compares significantly to the experimental distances listed. Structures t1 and t2 correspond to the high pH state, while structure u1 corresponds to the low-pH state.

degrees of freedom are localized, eliminating the need for diagonalization of a global matrix.[9] This algorithm has been implemented into a semiempirical quantum chemistry framework so that electronic structure calculations can be extended to macromolecular systems.[10] In addition, given the nature of DAC, the algorithm is quite suitable for parallel machine architectures, which we have used in the work presented here.[11] In this work, we performed several geometry optimizations of subsystems of CDA using the parallel implementation of the DAC approach within a semiempirical framework.

This DAC semiempirical method was used with the PM3 Hamiltonian[12] for all calculations reported within this paper. The PM3 parametrization is shown to work well for Zn complexes,[12b] and recently for simulations of the enzyme carbonic anhydrase which has a very similar hydrolytic mechanism and Zn tetrahedral coordination as found in cytidine deaminase, although typically semiempirical methods give bond lengths on the order of 5% too large.[7j] Macromolecules were divided into subsystems consisting of a single amino acid or nucleotide residue. Basis functions for projection of the Fock matrix for each subsystem were chosen to be the set of basis functions on all atoms within 5.0 Å of any subsystem atom. A 7.0 Å cutoff was used in the computation of density matrix elements. This protocol has been shown to give accuracy on the order of 10^{-3} eV/atom. The convergence criterion for the SCF procedure was 10^{-5} eV/atom.

In our calculations, we have created a 1330-atom subsystem which surrounds the active site (any residue within 8.0 Å of the ligand) of the enzyme CDA and includes contributions from the other monomer; therefore, the size of this subsystem consists of approximately 30% of a monomer of the enzyme. Chain ends in the subsystem were terminated by addition of acetyl or *N*-methyl groups. From this subsystem, eight geometries of different starting arrangements were optimized, keeping backbone atoms fixed, with a 0.2 kcal/(mol Å) rms and 2.4 kcal/(mol Å) maximum tolerance in the gradients (Table 1). Keeping the backbone atoms fixed is justified for this preliminary study by the fact that backbone atoms in several refined CDA complexes coincide with rms deviations of ∼0.15 Å for backbone atoms, and the only significant changes occur near the active site.[1m]

All of the optimized geometries of our large subsystems were obtained from gas-phase calculations, although we should stress that all crystallographic water molecules within the chosen radius were included in simulations of this subsystem. Given that the active site of CDA is not solvent-accessible,[1h] the gas-phase approximation is a physically accurate description. Ab initio work on Zn complexes indicate that a reaction in an enzyme-active cavity may actually be better approximated by a gas-phase model rather than by a model reaction in solution.[5] This point is definitely more valid for CDA than other zinc enzymes where the active site, such as in carbonic anhydrase, is partly solvent-accessible. Since our calculations include much of the protein environment surrounding the active site, our results are more physically reasonable than gas-phase calculations of a small subsystem where only a few primary components of the active site are included.

Results and Discussion

Initially, we optimized the geometry for a subsystem (s0) containing Zn−H_2O in the active site and with the carboxylate group of Glu-104 unprotonated to give the results shown in Figure 1a. This represents the complex just before the active species is created by deprotonation of Zn−H_2O. Results of geometry optimization for this subsystem indicate that the substrate water is stabilized by the carboxylate group of Glu-104 via hydrogen-bonding. Activation of CDA above pH ∼5 suggests that Zn−H_2O transfers a proton to the carboxylate group of Glu-104; however, two different carboxylate oxygen sites ($O^{\epsilon 1}$ and $O^{\epsilon 2}$) could accept a proton, and two possible protons could be donated by the substrate water. Four additional subsystems (s1, s2, s3, and s4) were created representing the configurations of these possible protonation arrangements, and their geometries were optimized.

The energies of these five optimized structures (s0, s1, s2, s3, and s4) are recorded in Table 1. One structure (s1), shown in Figure 1b, is lower in energy and is prominently the lowest energy structure than the other four structures by at least 46.0 kcal/mol. In this structure, a hydrogen-bonding network is formed among the hydroxide proton of Zn−OH^-, the protonated $O^{\epsilon 2}$ of Glu-104, and the N^3 position of the cytidine substrate. Crystallographic evidence suggests that the proton from the protonated $O^{\epsilon 2}$ of Glu-104 is later transferred to N^3 of the ligand in the catalytical mechanism, and a pre-transition state is formed.[1l] This hydrogen-bonding network gives credence to the conclusion that this structure is lower in energy, since the higher energy structures cannot form such a nicely formed hydrogen-bonding network.

To confirm that the lowest energy structure is the structure of the active species, three main criteria, determined on the basis of crystallographic data, should be considered. First, after deprotonation of Zn−H_2O occurs, a stronger Zn−OH^- bond is formed, yielding a decrease in the Zn−O bond length ($d_{Zn-O(OH)}$). Second, the distance between the oxygen on the substrate water and the C^4 position ($d_{O(OH)-C4(CTD)}$), where nucleophilic attack occurs on the cytidine ligand, should decrease as the nucleophilic attack is initiated and as the active species is formed. Third, recent crystallographic results show that the carbonyl oxygen of Thr-127 holds the leaving group (NH_2) in place as it is cleaved from the substrate.[1m] In the active species this distance between the carbonyl oxygen and N^4 of the cytidine ligand ($d_{O(THR127)-N4(CTD)}$) is expected to remain constant or decrease

(9) (a) Yang, W. *Phys. Rev. Lett.* **1991**, *66*, 438. (b) Zhao, Q.; Yang, W. *J. Chem. Phys.* **1995**, *102*, 9598.

(10) (a) Yang, W.; Lee, T.-S. *J. Chem. Phys.* **1995**, *103*, 5674. (b) Lee, T.-S.; York, D. M.; Yang, W. *J. Chem. Phys.* **1996**, *105*, 2744.

(11) Pan, W.; Lee, T.-S.; Yang, W. *J. Comput. Chem.*, in press.

(12) (a) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221. (b) Stewart, J. J. P. *J. Comput. Chem.* **1991**, *12*, 320.

slightly, placing a strain on the N^4-C^4 bond of the cytidine ligand. During nucleophilic attack, or as $d_{O(OH)-C4(CTD)}$ progressively decreases, this strain on the N^4-C^4 bond increases, allowing the C^4-O bond to strengthen.

In reference to these three criteria, distances for the optimized structures s0−s4 ($d_{Zn-O(OH)}$, $d_{O(OH)-C4(CTD)}$, and $d_{O(THR127)-N4(CTD)}$) are recorded in Table 1. Overall, calculated distances are on the order of 5% too large, inherent in semiempirical methods as previously discussed. The significance of the results listed in this table is discussed as follows. First, as expected, the distance $d_{Zn-O(OH)}$ is longer for the s0 structure, containing Zn−H_2O, compared to any of the structures containing Zn−OH^- (s1−s4). The decrease from 2.28 to ∼1.98 Å is attributable to the deprotonation of the substrate water, yielding a stronger covalent Zn−O bond. Second, the distance $d_{O(OH)-C4(CTD)}$ decreases as the system goes from the Zn−H_2O structure (s0) to the lower energy Zn−OH^- structure (s1). The decrease (∼0.21 Å) is quite significant. In the other structures (s2, s3, and s4), which are higher in energy, $d_{O(OH)-C4(CTD)}$ also decreases, but the decrease is the greatest in the lowest energy structure. Third, the distance $d_{O(THR127)-N4(CTD)}$ decreases most significantly as the system goes from the Zn−H_2O structure (s0) to the lowest energy Zn−OH^- structure (s1), with hardly any decrease for the other Zn−OH^- structures.

Subsystems with protonating arrangements corresponding to low- and high-pH states were also considered. The low-pH structure consists of a three-proton arrangement containing Zn−H_2O with the $O^{\epsilon 2}$ of Glu-104 protonated (u1), while the two high-pH structures consist of a one-proton arrangement containing Zn−OH^- with Glu-104 unprotonated (t1 and t2). Note that these charged systems (t1, t2, and u1) are contained in different classes and the energies are not comparable to the normal pH structures (s0−s4); however, our results do provide a prediction for the structure of the active site at low pH and high-pH. In the t1 and t2 high pH structures, the distance $d_{O(OH)-C4(CTD)}$ does not decrease as the system is changed from the original Zn−H_2O structure (s0), indicating that the nucleophilic attack has not been initiated. These structures are therefore eliminated as possible candidates for the structure of the active species. In addition, we predict that at high pH the distance $d_{Zn-O(OH)}$ decreases by approximately 0.05 Å compared to the structure of the active species at normal pH. The u1 structure looks promising as a structure of the active species, given that the distances $d_{O(OH)-C4(CTD)}$ and $d_{O(THR127)-N4(CTD)}$ decrease as predicted by crystallographic data; however, we predict that at low pH the Zn−O bond length should be quite large, as if the substrate H_2O is not even bound to the Zn atom. While the high-pH form of the enzyme is unlikely to be experimentally obtainable, the low-pH state of the crystallized enzyme should be achievable, providing an excellent experimental test of predictions arising from the calculations.

Considering all calculated structures, not only is the s1 structure lower in energy compared to the other two-proton structures (s0, s2, s3, and s4), but it also correlates quite well in detail with the geometry observed at the active site of the enzyme in the X-ray crystal structure.[1l] Both the crystal structure[1l] and our calculations for the s1 structure indicate that the oxygen of the hydroxide and C^4 of the cytidine ligand are significantly closer than the van der Waals limit. Moreover, the short $OH-C^4$ distance implies, in turn, that the nucleophilic attack is being initiated as the ground-state Zn−OH^- structure is formed. Evidence both from the energetics and from comparison to crystallographic data support the fact that the s1 structure closely resembles the structure of the active species.

Conclusions

We demonstrate that the DAC approach within a semiempirical framework is quite efficient for studying the zinc enzyme CDA. We predict the structure of the active species by comparing geometry-optimized structures and demonstrating that the lowest energy structure is closest to experiment in comparison to the other geometry-optimized structures. From the lowest energy structure, our results indicate that Zn−OH^- is the active species formed prior to nucleophilic attack of the ligand, and that in the active species Glu-104 is protonated at $O^{\epsilon 2}$ and hydrogen-bonded with N^3 of the ligand, which is also supported experimentally from crystallographic geometry details. Having determined the structure of the active species, simulations of analogue complexes of the pre-transition,[1j] transition,[1h] and product states[1m] in relation to the simulations of the substrate complex will help more fully understand the structural and energetic evolution along the reaction pathway.

Acknowledgment. This is a publication from the UNC/Duke Computational Structural Biology Resource (NIH, National Center for Research Resources, Grant No. RR08102). We thank Dean Carlow for many useful discussions in relation to these results. W.Y. also acknowledges support from the National Science Foundation and the Alfred P. Sloan Foundation. The North Carolina Supercomputing Center provided CPU time on the Cray T3E.

JA973522W