

**EXPERIMENTAL AND COMPUTATIONAL
STUDIES OF THE CATALYTIC MECHANISM OF
YEAST CYTOSINE DEAMINASE**

BY

JIFENG WANG

A DISSERTATION

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Biochemistry and Molecular Biology

2011

ABSTRACT

EXPERIMENTAL AND COMPUTATIONAL STUDIES OF THE CATALYTIC MECHANISM OF YEAST CYTOSINE DEAMINASE

By

Jifeng Wang

Yeast cytosine deaminase (yCD) catalyzes the conversion of the prodrug 5-fluorocytosine (5FC) to 5-fluorouracil (5FU), which is a widely used drug in the treatment of cancers. yCD is of significant biomedical interest because of its potential application in combination with 5FC in gene-directed enzyme prodrug therapy (GDEPT).

In this thesis, the catalytic mechanism of yCD was elucidated by both experimental and computational methods. First, the transition state structure was accessed experimentally by kinetic isotope effects (KIEs) studies at natural abundance by ^{13}C nuclear magnetic resonance (NMR) spectroscopy. The results indicate that the reaction proceeds by a stepwise $A_{\text{N}}+D_{\text{N}}$ mechanism, with the transition state in the C4-N4 bond cleavage step. Second, the catalytic mechanism of the deamination of 5-fluorocytosine (5FC) was investigated by an ONIOM (B3LYP/6-31G*:Amber)-electronic embedding scheme and was compared side-by-side with that of cytosine. The higher catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$) for 5FC is likely due to the energetically more favorable intermediates formed during the reaction. The method alone was not suitable to elucidate the $\text{O}^{\text{Zn}}\text{-Zn}$ bond breakage because relatively large rearrangements of the active site

residues are required. Hence, Quantum Mechanics/Molecular Mechanics (QM/MM) MD simulations were set up to incorporate configurational sampling into chemical transformation. In the reaction pathway proposed, it is the C4-N4 bond cleavage step, not the O^{Zn}-Zn bond cleavage step, that is the rate-limiting step for the overall reaction. Functions of conserved residues around the active site were investigated by site-directed mutagenesis, biochemical, NMR, and computational methods. Glu64 is found to be essential in catalysis due to its role as a general base/acid. Steady-state kinetic studies show that mutating Glu64 to Ala causes a dramatic decrease in k_{cat} by a factor of >3000. Glu64 also participates in the stabilization of the intermediates and transition state, suggested by ¹⁹F NMR experiments and ONIOM calculations. Asn51 and Asp155 interact with the substrate through hydrogen bonding, by which the substrate is anchored in the right position during the initial binding and at the transition state. These residues are important for substrate binding and catalysis.

yCD belongs to the cytidine deaminase (CDA) family of purine/pyrimidine deaminases. Enzymes in this family have similar three-dimensional structures with a zinc-containing catalytic apparatus. Studies on yCD not only elucidate the enzymatic mechanism of yCD itself, but also provide significant insights into the catalytic mechanism of enzymes in this family. Moreover, these findings can facilitate antimicrobial drug design and gene directed therapy against cancers.

Copyright by
JIFENG WANG
2011

To Confucius (551 B.C. – 479 B.C.)

ACKNOWLEDGEMENTS

Am I the first PhD ever in my house? Errrr... no. But I am proud to be the second one, in a BIG family of three. *Hei, son, you need to catch up faster.*

Well, I am lucky that my son still regards me as his “best friend” although I could not play with him too often since he was born.

But why he likes me more? (This line is invisible to the wife)

Because I always bought toys for him? I am not sure. *Hei, son, if you get a chance to read this, come and tell me. But I will tell you this: It is you mom who took care of you when I was not around. Let's show our great appreciations to you mom, my best friend!*

Of course, I would like to thank my own mom. And dad. Every little success, from my first step, to this thesis, has endless supports and encouragements from them.

I want to express my sincere appreciations to my advisor, Dr. Honggao Yan, for his guidance in my PhD studies. Dr. Yan gave me a lot of freedom to construct my research and taught me to think independently. I felt more and more confident over these years. The time spent in Room 313 of the Biochemistry building definitely prepared me well to be a good researcher in the future.

Well, Room 313 is not the only room I frequents. In the conference room, I met with my Guidance Committee regularly, from whom, I received great advices and critical reviews. I want to thank all members: Dr. Zachary F. Burton, Dr. Robert I. Cukier, Dr. Robert P. Hausinger, Dr. Aizhuo Liu, and Dr. Claire Vieille. I want to thank Dr. Michael Feig and Dr. William J. Wedemeyer too for their supports and guidance.

Room 313 is and will always be a special place for me. During these years, people working in this room are very supportive with each other. I learnt a little bit this or that from different person. I want to express my gratitude to my people: Dr. Yue Li, Yan Wu, Dr. Lishan Yao, Dr. Zhenwei Lu, Dr. Yi Wang, Dr. Jaroslaw Blaszczyk, Dr. Bin Chen, Rahul Banerjee, Yichen Tang and David Achila. It has been a lot of fun to work with you!

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
ABBREVIATIONS	xv
Chapter 1. Introduction	1
Biomedical Significance and Basic Kinetic Properties of Cytosine Deaminase	1
X-ray Crystallographic Analysis of yCD	3
Reaction Pathway by ONIOM Calculations	5
Zinc-uracil Bond Breakage Calculations	6
Product Release is Rate-limiting in the Activation of 5FC by yCD	8
yCD Dynamics	8
Cytosine Dissociation Pathway by MD Simulations	9
The Scope of this Dissertation	9
Figures	15
References	18
Chapter 2. Probing the Transition State of Yeast Cytosine Deaminase by NMR Measurement of ¹³C Kinetic Isotope Effects	24
Introduction	24
Materials and Methods	26
Results and Discussion	32
Conclusions	39
Tables and Figures	41
References	56
Chapter 3. Reaction Pathway of the Deamination of 5FC: ONIOM Calculation Revisited	62
Introduction	62
Computational Methods	64
Results and Discussion	65
Conclusions	76
Tables and Figures	79
References	91
Chapter 4. QM/MM MD Simulations Shed New Insight on the Catalytic Mechanism	94
Introduction	94

Experimental Procedures	96
Results and Discussion	100
Concluding Remarks	113
Tables and Figures.....	117
References.....	134

Chapter 5. Role of Glutamate 64 in the Activation of the Prodrug 5FC **139**

Introduction.....	139
Experimental Procedures	141
Results	147
Discussion.....	156
Tables and Figures.....	162
References.....	172

Chapter 6. Contributions of Conserved Residues in the Active Site to Substrate Binding and Catalysis **176**

Introduction.....	176
Experimental Procedures	178
Results	184
Discussion.....	193
Conclusions.....	196
Tables and Figures.....	198
References.....	204

Chapter 7. Summary and Future Work **207**

Summary.....	207
Future Work.....	209

LIST OF TABLES

Table 2-1. Reaction extents for yCD catalyzed reaction.....	41
Table 2-2. Intensities of the carbon peaks in the NMR samples for the original cytosine	41
Table 2-3. Intensities of the carbon peaks in the NMR samples for the residual cytosine after deamination by yCD.....	41
Table 2-4. Experimental ^{13}C KIEs.....	42
Table 2-5. Calculated ^{13}C KIEs.....	42
Table 2-6. Bond distances for cytosine and the transition states	43
Table 3-1. Comparison of key coordinates of structures from ONIOM calculations with crystal structures.....	79
Table 4-1. Comparison of changes of key coordinates during the conversion of cytosine to uracil with crystal structures.....	116
Table 5-1. Kinetic constants of the wild-type yCD and the E64 mutant enzymes for the activation of 5FC.....	162
Table 5-2. Inhibitory and binding properties of 5FPy.....	162
Table 6-1. Kinetic constants of the wild type yCD and mutant enzymes for the activation of 5-fluorocytosine	198
Table 6-2. Binding constants of 5FPy with the wild type yCD and mutant enzymes ...	198
Table 6-3. Internal energy and solvation energy components for the interactions between residues and Py.....	199
Table 6-4. Individual energy components by MM-GBSA energy decomposition.....	199

LIST OF FIGURES

Figure 1-1. Reaction catalyzed by yCD.....	15
Figure 1-2. Crystal structure of yCD.....	16
Figure 1-3. Interactions between the inhibitor DHP and surrounding residues in the active site of yCD.....	17
Figure 2-1. Proposed reaction pathway for yCD-catalyzed deamination of cytosine to uracil.....	44
Figure 2-2. Experimental KIEs mapped on the cytosine structure.....	45
Figure 2-3. Possible transition state structures.....	46
Figure 2-4. Model system for ONIOM calculations.....	47
Figure 2-5. ONIOM optimized transition state structure in the active site of yCD.....	48
Figure 2-6. Molecular electrostatic potential (MEP) surfaces for cytosine, tetrahedral intermediate, uracil, Py, DHP, and three transition states.....	49
Figure 2-7. Proposed reaction mechanism based on KIEs.....	50
Figure 2-8. Coordinates of optimized cytosine.....	51
Figure 2-9. Coordinates of optimized tetrahedral intermediate.....	52
Figure 2-10. Coordinates of optimized transition state 1.....	53
Figure 2-11. Coordinates of optimized transition state 2.....	54
Figure 2-12. Coordinates of optimized transition state 3.....	55
Figure 3-1. Superposition of the ONIOM optimized yCD-DHP complex structure with the crystal structure.....	80
Figure 3-2. Proposed catalytic pathway for the conversion of cytosine (5FC) to Zn-coordinated uracil (5FC).....	81
Figure 3-3. ONIOM optimized structures for complex 1-6 and transition state structures during the deamination of 5FC.....	82
Figure 3-4. ONIOM optimized structures for complex 1'-6' and transition state structures during the deamination of cytosine.....	83

Figure 3-5. Energy profiles for the conversion of cytosine to the Zn-coordinated uracil and the conversion of 5FC to the Zn-coordinated 5FC	84
Figure 3-6. Rotation of the 5FC ring during the deamination of 5FC	85
Figure 3-7. Motion of the active site residues during the deamination of 5FC	86
Figure 3-8. Proposed catalytic pathway for the replacement of the Zn-coordinated uracil or 5FU with water	87
Figure 3-9. ONIOM optimized structures for Complex 7, 8 and the transition state structure for the deamination of 5FC	88
Figure 3-10. ONIOM optimized structures for Complex 7', 8' and the transition state structure for the deamination of cytosine	88
Figure 3-11. Energy profiles for replacement of the Zn-coordinated uracil or 5FU with water	89
Figure 3-12. Conformational change caused by the access of the water towards the zinc atom	90
Figure 4-1. QM partition in the simulation systems for the yCD-DHP complex (top) and with the yCD water/cytosine complex (bottom)	117
Figure 4-2. RMSDs of the backbone C α atoms relative to the crystal structure during the QM/MM MD simulation for the yCD-DHP complex	118
Figure 4-3. Superposition of a representative snapshot during the yCD-DHP complex simulation with the crystal structure	119
Figure 4-4. Bond distances of snapshots along the MD trajectory of the yCD-DHP complex simulation	120
Figure 4-5. Energy profiles for proton transfer	121
Figure 4-6. Proposed reaction pathway for the conversion of cytosine to uracil in the QM/MM MD simulations	122
Figure 4-7. Snapshots of reaction species along the reaction pathway for the conversion of cytosine to uracil in the QM/MM MD simulations	123
Figure 4-8. RMSDs of the backbone C α atoms relative to the crystal structure during the QM/MM MD simulation for the yCD water/cytosine complex	124
Figure 4-9. Rotation of the sidechain of Glu64 (\angle CB-CG-CD-OE2) in the free form yCD and the yCD substrate complex simulations	125

Figure 4-10. Potential of mean force for the proton transfer from the Zinc-coordinated water to O ^{ε2} of Glu64 and to N3 of cytosine.....	126
Figure 4-11. Potential of mean force for nucleophilic attack on C4 of cytosine.....	127
Figure 4-12. Superposition of the average structure of snapshots in the MD trajectory for the complex 4 simulation onto that of the yCD-DHP simulation.....	128
Figure 4-13. Potential of mean force for the C4-N4 bond cleavage.....	129
Figure 4-14. Reaction pathway for the O ^{Zn} -Zn bond breakage.....	130
Figure 4-15. Potential of mean force for proton transfer from N3 to O ^{ε2}	131
Figure 4-16. Potential of mean force for the O ^{Zn} -Zn bond breakage.....	132
Figure 4-17. Overall energy profile for the conversion of cytosine to uracil by yCD...	133
Figure 5-1. ¹⁹ F-NMR spectra of the E64A-5FPy complex and the wild-type yCD-5FPy complex.....	163
Figure 5-2. Proposed reaction pathway from Py to DHP catalyzed by the wild type yCD.....	164
Figure 5-3. ONIOM optimized structures for the conversion of Py to DHP by the wild type yCD.....	165
Figure 5-4. Schematic E ^{ONIOM} and E(High,model) energy profile for the conversion of Py to DHP catalyzed by the wild-type yCD.....	166
Figure 5-5. Proposed reaction pathway from Py to DHP catalyzed by the E64A mutant enzyme.....	167
Figure 5-6. ONIOM optimized structures for the conversion of Py to DHP by the E64A mutant enzyme.....	167
Figure 5-7. Schematic E ^{ONIOM} and E(High,model) energy profile for the conversion from Py to DHP catalyzed by the E64A mutant enzyme.....	168
Figure 5-8. Conversion of cytosine to the zinc-coordinated uracil by the E64A mutant enzyme.....	169
Figure 5-9. ONIOM optimized structures for the conversion from cytosine to the zinc-coordinated uracil.....	170

Figure 5-10. Schematic E ^{ONIOM} and E(High,model) energy profile for the conversion from cytosine to uracil catalyzed by the E64A mutant enzyme	171
Figure 6-1. ¹⁹ F-NMR spectra of the mutant-5FPy complex and wild type yCD-5FPy complex.....	200
Figure 6-2. RMSDs of the backbone C α atoms relative to the starting structure for the yCD Py complex during the 8 ns production MD simulations.....	201
Figure 6-3. Ligand-residue interaction spectrum of the yCD Py complex by MM-GBSA energy decomposition.....	201
Figure 6-4. 2D energy profiles for proton transfer from the Zinc-coordinated water to N3 of Py.....	202
Figure 6-5. Superposition of representative snapshots for the yCD Py complex, N51A Py complex and D155A Py complex.....	203

ABBREVIATIONS

5FC	5-fluorocytosine
5FDHP	5-fluoro-4(<i>R</i>)-hydroxyl-3,4-dihydropyrimidine
5FPy	5-fluoro-pyrimidi-2-one
5FU	5-fluorouracil
ADA	Adenosine Deaminase
Amber	Software Package for Molecular Dynamic Simulation
B3LYP	Becke three parameter, Lee-Yang-Parr
CD	Cytosine Deaminase
CDA	Cytidine Deaminase
CHARMM	Software Package for Molecular Dynamic Simulation
DHP	4(<i>R</i>)-hydroxyl-3,4-dihydropyrimidine
eCD	<i>E. coli</i> Cytosine Deaminase
FID	Free Induction Decay
G03	Software Package for Quantum Mechanics
GBSA	Generalized Born Surface Area
GDA	Guanine Deaminase
GDEPT	Gene Directed Enzyme Prodrug Therapy
IC50	Half Maximal Inhibitory Concentration
k_{cat}	Turnover Number

K_d	Dissociation Constant
K_i	Inhibition Constant
KIE	Kinetic Isotope Effect
K_m	Michaelis Constant
MD	Molecular Dynamics
MM	Molecular Mechanics
MM-PBSA	Molecular Mechanics Poisson Boltzmann Surface Area
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
ONIOM	Our own N-layered Integrated molecular Orbital and molecular Mechanics
PARM99SB	An Amber force field
PMF	Potential of Mean Force
Py	2-pyrimidinone
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuations
TadA	tRNA Adenosine Deaminase
TS	Transition State
WT	Wild Type

yCD

Yeast Cytosine Deaminase

CHAPTER 1

INTRODUCTION

Biomedical Significance and Basic Kinetic Properties of Cytosine Deaminase

Cytosine deaminase (CD) (EC: 3.5.4.1) catalyzes the deamination of cytosine to uracil and plays an important role in the pyrimidine salvage pathway (1) in bacteria and fungi. The enzyme is absent in humans, in which cytidine deaminase is utilized for the pyrimidine conversion instead.

Bacterial CDs, represented by the *E. coli* enzyme (eCD), and fungal CDs, represented by the yeast enzyme (yCD), have no sequence homology and utilize different metal ions for catalysis, an example of convergent evolution (2). yCD was discovered by Hahn and Schafer in 1925 (3, 4). The basic kinetic properties of yCD were analyzed by Kream and co-workers on a partially purified preparation of the enzyme (5). It was found that yCD could use both cytosine (Figure 1-1) and 5-methylcytosine as substrate. yCD was first purified by Ipata and co-workers (6). Their kinetic studies further demonstrated that yCD specifically catalyzes the deamination of cytosine and 5-methylcytosine ($K_m=2.5$ mM for both) among a number of nucleosides and nucleotides. The reaction rate was pH-dependent, with the optimum at pH 7.0. The reaction was inhibited by cytidine, thymidine, guanosine, and their nucleotides in a noncompetitive manner, whereas many

bases and other nucleosides and nucleotides tested had no inhibitory effects at the concentration of 0.5 mM, including adenine, adenosine, uracil, uridine, guanine, xanthine, hypoxanthine, thymine, AMP, ADP, ATP, UTP, IMP, and ITP. The inhibitory effects by cytidine, thymidine, guanosine, and their nucleotides were suggested to play important roles in the regulation of pyrimidine nucleotide biosynthesis in yeast (6).

Both yCD and eCD also catalyze the deamination of 5-fluorocytosine (5FC) to 5-fluorouracil (5FU) (Figure 1-1). 5FU is a widely used drug to treat breast, rectal, stomach, colon, and pancreatic cancers (7, 8). However, 5FU is highly toxic. The undesired side effects can be minimized (9) by using the prodrug 5FC in combination with a CD in gene-directed enzyme prodrug therapy (GDEPT). The procedure generally includes two steps (10). First, a gene coding a CD is delivered to the tumor cells. Second, the prodrug 5FC is injected into the tumor cells and activated by the CD. Because humans do not have their own CD, 5FC is converted to 5FU only in tumor cells (2). Thus, in principle, only the tumor cells are exposed to the toxicity of 5FU and are killed, and normal cells are not affected.

Both eCD and yCD have been used in GDEPT. eCD is chosen because of its thermal stability (11). However, its application is limited, due to its inefficiency in converting 5FC to 5FU. yCD has higher catalytic efficiency towards 5FC than eCD, though it has a relatively lower thermal stability (12). The k_{cat} of yCD for 5FC is 4 times lower than that of eCD and the K_m for 5FC is 21-fold lower than that of eCD (11, 13-15). Consequently, the catalytic efficiency, k_{cat}/K_m , of yCD is 5 times that of eCD. Its advantage over eCD in GDEPT was demonstrated by *in vivo* experiments, in which

complete tumor regression was observed in 6 of 13 HT29/yCD tumors when treated with 5FC, while none of 10 HT29/eCD tumors were cured (11).

Efforts have been made towards improving the catalytic efficiency and thermostability of yCD (16, 17). yCD has a relatively low k_{cat}/K_m ($1.1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ for 5FC), compared with that of the most proficient enzymes, which is around $10^9 \text{ M}^{-1} \text{ s}^{-1}$ (18). By directed evolution or structure-based protein engineering, variants with high efficiency and thermal stability may be obtained. For example, Stolworthy and coworkers (16) have identified a yCD variant (yCD/D92E) with increased thermostability and comparable activity by substituting the conserved residue Asp92 to a glutamate. Our laboratory also identified several variants which may be more effective than the wild-type enzyme for GDEPT (unpublished data).

In addition to developing improved candidates for gene therapy, a number of kinetic, structural, and computational studies were reported to elucidate the catalytic mechanism of yCD (2, 5, 17, 19-34). yCD belongs to the cytidine deaminase (CDA) family of purine/pyrimidine deaminases. Enzymes in this family have similar structures with a zinc-containing catalytic apparatus (2, 19). Studies on yCD not only elucidate the enzymatic mechanism of yCD itself, but also provide significant insights into the catalytic mechanism of enzymes in this family. Moreover, these findings can facilitate antimicrobial drug design and gene-directed therapy against cancers. Such studies are presented below.

X-ray Crystallographic Analysis of yCD

Structural information at the atomic level is almost a necessity for understanding of the catalytic mechanism of an enzyme. Crystal structures of yCD have been determined at high resolution, including the substrate-free form (PDB id: 1ox7) and the complex form (PDB id: 1p6o and 1uaq) with 2-pyrimidinone (Py), a transition state or reaction intermediate analog (2, 19).

yCD is a homodimeric metalloenzyme (Figure 1-2). Each monomer is tightly packed into a mixed α/β fold, which consists of a central five stranded β -sheet sandwiched by α -helices on both sides. A catalytic zinc atom is found in the active site, covalently bonded with His62 N $^{\delta 1}$, Cys91 S $^{\gamma}$, Cys94 S $^{\gamma}$ and a water molecule in the free form yCD. The water molecule is believed to serve as a nucleophile and is hydrogen bonded to the conserved residue Glu64, which is likely to serve as a proton shuttle during the reaction. The inhibitor is bound in the hydrated form 4(*R*)-hydroxyl-3,4-dihydropyrimidine (DHP) (Figure 1-1) and coordinates with the zinc atom through the 4-hydroxyl group (Figure 1-3). In both the free form and the complex form, the active site is completely covered by the same cluster of residues from the C-terminal α -helix and the loop between $\beta 4$ and αD (F114 loop) and appears to be inaccessible to the substrate. Structural alignment shows that the crystal structures of the free form and the complex form are essentially the same with respect to protein conformation, with a root mean square deviation (RMSD) of 0.23 Å (20) for the backbone C α atoms, which suggests that dynamics may play a role in catalysis because either substrate binding or product release needs the movement of the C-terminal α -helix or the F114 loop, if not both.

The crystal structures are invaluable to computational protein engineering studies (17). A double mutant variant (A23L/I140L) and a triple mutant variant (A23L/I140L/V108I) have been successfully designed (17), possessing higher apparent melting temperature T_m and increased half-life of activity at 50 °C without sacrificing catalytic efficiency.

Reaction Pathway by ONIOM Calculations

An ONIOM approach has been employed to elucidate the catalytic mechanism of yCD for the conversion of cytosine to uracil (20). The ONIOM method in this study is a two-layered hybrid method developed by Morokuma and co-workers (35) and implemented in the Gaussian package (Gaussian 03) (36). Twenty-two amino acid residues were included to model the enzyme in the molecular system. The molecular system was divided into two layers, which were treated with different levels of theory. The inner layer included the substrate cytosine and the zinc-coordinated water, the catalytic zinc ion, the potential proton shuttle Glu64, and residues coordinated with the catalytic zinc, His62, Cys91, and Cys94, and treated by the density functional method B3LYP with the 6-31G** basis set. The remaining residues were included in the outer layer and treated by the semiempirical PM3 method. A reaction pathway was proposed based on the ONIOM calculations. The substrate cytosine is activated via two proton transfer steps. First, the proton is transferred from the zinc-coordinated water to the carboxyl group of Glu64. Second, the same proton is transferred to N3 of cytosine. The energy barriers for these steps are negligible. The complex with the activated substrate is more stable than the original complex by 7.1 kcal/mol at the ONIOM energy level. The

zinc-coordinated hydroxide is well positioned for the nucleophilic attack on C4 of cytosine. The energy barrier is calculated to be 1.0 kcal/mol at the ONIOM level for the formation of the tetrahedral intermediate. The C4-N4 bond is cleaved in the next step, forming two products, ammonia and uracil. It is most likely that ammonia rapidly diffuses into solution and is replaced with a water molecule from the solution, as suggested by binding free energy calculations (21). The other product, uracil, is covalently bonded with the zinc atom. The C4-N4 bond cleavage is suggested to be a rate-limiting step for the deamination of cytosine, with an energy barrier of 7.1 kcal/mol at the ONIOM level.

It is challenging to understand how the bond between Zn and O4 of uracil is cleaved. Two possible pathways were proposed in this ONIOM study. In one pathway, the bond is directly cleaved. However, the results indicate that this pathway is unlikely, due to the high energy barrier (18.3 kcal/mol). In the second hypothesis, a new mechanism is proposed, in which the oxygen of the water molecule exchanges into uracil. Two enzymes in the CDA family of purine/pyrimidine deaminases, cytidine deaminase and adenosine deaminase, adopt this strategy to break the same bond (37). The rate-limiting step is found in the formation of the gem-diol intermediate from the zinc bound uracil and the water molecule. The energy barrier is 9.3 kcal/mol at the ONIOM level. Although the energy barrier is much lower than that proposed in the first mechanism, it is still the rate-limiting step for the overall reaction of cytosine deamination catalyzed by yCD.

Zinc-uracil Bond Breakage Calculations

In the first hypothesis of the zinc-uracil bond breakage mentioned above, the activation energy seems too high. However, our recent experiments (unpublished data) did not support the second mechanism either, because this oxygen exchange rate is too slow. Direct breakage of the Zn-O4 bond was re-examined by a two-stepped ONIOM-Molecular Dynamics (MD) calculation (22). First, the protein in complex with uracil was simulated by a series of standard MD simulations, with a restraint on the bond distance between Zn and O4 of uracil. The bond distance increased between successive simulations in 0.1 Å increments from 2.0 Å to 3.8 Å. Second, a number of snapshots (5 or 10) were selected from the equilibrated MD ensembles and optimized using the ONIOM scheme as described above (20). The results show that the bond is cleaved with the help of O^{ε1} of Glu64. O^{ε1} replaces O4 of uracil to maintain the zinc coordination. The potential of mean force (PMF) along the bond-breaking process was calculated. The transition state is located at the Zn-O4 distance of approximately 2.6-2.7 Å. The energy barrier was calculated to be only 2.9 kcal/mol. It is much lower than that calculated in the previous ONIOM calculation, in which the energy barrier is 18.3 kcal/mol. The difference is rooted in the fact that in the previous ONIOM calculation, the outer layer is frozen throughout the bond cleavage process, whereas, in this new strategy, the outer layer is allowed to move in response to the bond-breaking process. The Zn-O4 bond cleavage thus is unlikely to be the rate-limiting step in the overall deamination reaction of cytosine.

To explore the dynamic effects on the catalysis, Matsubara and co-workers (23) performed an ONIOM-MD study on the process of the zinc-uracil bond breakage. The results show that the Zn-O4 bond abruptly breaks at 11 ps of simulation. The authors

suggest that the thermal motion of Ile33, which sandwiches uracil with His62, perturbs the dynamics of uracil through steric contact and contributes to the bond breakage.

Product Release is Rate-limiting in the Activation of 5FC by γ CD

The deamination of the anticancer prodrug 5FC by γ CD has been investigated by transient kinetic and NMR studies (25). The results show that product release is the rate-limiting step in the activation of 5FC, with the rate constants of the chemical steps being at least 8 times that of the product release step. NMR spectra for γ CD in the free form and in the complex form with 5FU were acquired and analyzed. The K_d of 5FU with γ CD was estimated to be ~ 20 mM, indicating that the binding affinity of 5FU with γ CD is low. NMR saturation transfer experiments show that 5FU is in slow exchange with its γ CD complex on the NMR time scale. The association and dissociation rate constants are $0.6 \text{ mM}^{-1} \text{ s}^{-1}$ and 14 s^{-1} , respectively.

γ CD Dynamics

As revealed by the crystal structures, the active site is completely covered by the C-terminal α -helix, the F114 loop and loop 1 (residues 53-61). Protein dynamic properties were studied by NMR relaxation and H/D exchange experiments and MD simulations (Yao, unpublished data). 5-fluoro-pyrimidin-2-one (5FPy) is an intermediate analog which can be converted to the hydrated 5-fluoro-4(*R*)-hydroxyl-3,4-dihydropyrimidine (5FDHP) (Figure 1-1). ^{15}N NMR relaxation experiments show that the F114 loop and the C-terminal helix are more rigid in the 5FPy bound form than in the

free form on the ps to ns timescale, whereas the dynamic properties of the 5FU bound form are similar to those of the free form. MD simulations show that the root mean square fluctuations (RMSFs) for the C α atoms are generally larger in these regions than other regions. H/D exchange experiments show that binding of 5FPy increases exchange times of the F114 loop, loop 1 and the C-terminal helix, while binding of 5FU does not increase exchange time too much. These results suggest that these regions are important for the access of the active site.

Cytosine Dissociation Pathway by MD Simulations

To understand how a ligand accesses the sealed active site of yCD, a series of MD simulations were performed on the dissociation process of cytosine (27). Two exit paths were suggested based on the simulations. One path (Path 1) is in between the C-terminal helix and the F114 loop, and the other (Path 2) is in between the C-terminal helix and loop 1. To test which exit path is preferred, two MD simulations were performed; cytosine was pushed out of the active site along Path 1 in one simulation and Path 2 in the other, respectively. In Path 1, the motion associated with cytosine exit is small and rather local, involving only the C-terminal helix and the F114 loop. In Path 2, the conformational changes are much more extensive. The average force is larger for pushing cytosine out along Path 2 than Path 1, suggesting that cytosine and other ligands may exit the active site preferably through Path 1.

The Scope of this Dissertation

In this dissertation, starting from what we have learned from the literature and accumulated work in the laboratory, I carried out a series of experiments and computations to further explore the catalytic mechanism of yCD. The work is presented in detail in Chapter 2-6 and summarized below.

In the first project, presented in Chapter 2, the transition state of yCD was probed by kinetic isotope effects (KIEs) at natural abundance by ^{13}C NMR spectroscopy (38). KIEs reflect the change of reaction rate due to isotopic substitution at a given position. Combining KIEs at different positions, it is possible to understand the reaction mechanism. Instead of being isotopically labeled, the substrate cytosine at natural abundance was used in this study. To determine the transition state structure, a series of conformations was constructed that represented different stages of C4-O4 bond formation or C4-N4 bond cleavage (39), by fixing the C4-O4 bond or the C4-N4 bond at different distances, respectively. Vibrational frequencies were calculated, which, in turn, were used to calculate KIEs. The results indicate that the reaction proceeds by a stepwise $\text{A}_\text{N}+\text{D}_\text{N}$ mechanism, with a transition state in the C4-N4 bond cleavage. The late transition state established in this study is similar to that reported for *E. coli* cytidine deaminase (39) and TadA RNA deaminase (40), but different from the early transition states for adenosine deaminases from human, bovine and *Plasmodium falciparum* (41) and AMP deaminase from *Saccharomyces cerevisiae* (42).

In the second project, presented in Chapter 3, the catalytic mechanism of the deamination of 5FC by yCD was investigated by an ONIOM(QM:MM) method with the electronic embedding (EE) scheme (43). Different from the previous work, the whole protein was included in the model system. The inner layer was treated by the DFT

method B3LYP with the 6-31G* basis set and the outer layer by the Amber force field. The results show that the reaction pathway for the deamination of 5FC is similar to that of cytosine. However, energetically more favorable tetrahedral intermediates are found, which may account for the relatively low K_m of 5FC. This finding may facilitate rational design of enzyme variants specifically towards 5FC, instead of cytosine. In addition, one possible Zn-O4 bond cleavage pathway was proposed. However, configurational sampling of proper geometries along the pathway was not promising. Hence, the reaction pathway was re-investigated by QM/MM MD simulations in Chapter 4.

In the third project, presented in Chapter 4, the reaction mechanism of the deamination of cytosine by yCD was investigated by combined QM/MM MD and free energy simulations (44, 45). First, the substrate cytosine is activated by abstracting one proton from the zinc-coordinated water via $O^{\epsilon 1}$ of Glu64. The zinc-coordinated hydroxide is well positioned for the nucleophilic attack on C4. Second, the tetrahedral intermediate is formed after the nucleophilic attack. The energy barrier is quite low, with a value of 1.4 kcal/mol. Third, the C4-N4 bond is cleaved, which is believed to be the rate-limiting step with an energy barrier of 19.7 kcal/mol. The product ammonia can easily diffuse into the solution, while the other product uracil is covalently bonded with the zinc atom. We propose that one solvent water enters the active site and replaces uracil in coordination with Zn. The energy barrier is about 8.1 kcal/mol. The results show the overall transition state is located in the N4-C4 bond cleavage step, which is in agreement with our KIE experiments.

In the fourth project, presented in Chapter 5, the functional role of Glu64 in the activation of the prodrug 5FC was investigated by site-directed mutagenesis and

biochemical, NMR, and computational studies. In both ONIOM calculations and QM/MM MD simulations, Glu64 was found to play a critical role in yCD catalysis, serving as a proton shuttle in multiple steps during the catalytic cycle. A better understanding of the role of Glu64 will provide more insight not only on the catalytic mechanism of yCD but also on other members of the CDA family. Steady-state kinetic studies show that substituting Glu64 with Ala (E64A) causes a dramatic decrease in k_{cat} and a dramatic increase in K_{m} , with k_{cat} reduced by a factor of >3000 and K_{m} increased by a factor of >100. The substitutions also cause a dramatic decrease in the affinity of the reaction intermediate analogue 5FPy. After initial binding with the wild type enzyme, the inhibitor is hydrated, as suggested by X-ray crystallography. ^{19}F -NMR experiments show that binding of 5FPy causes an upfield shift, indicating that 5FPy is hydrated when bound in the active center of the wild-type enzyme. However, binding with the E64A mutant enzyme causes a downfield shift, indicating that 5FPy remains in the unhydrated form in the complex. ONIOM calculations show that the wild type yCD complex with the hydrated 5FPy is more stable than the initial binding complex, and in contrast, with the E64A mutant enzyme, the hydrated inhibitor is no longer favored and the activation energy is elevated. To further explore the function of Glu64 in catalysis, the deamination of cytosine catalyzed by the E64A mutant enzyme was investigated by ONIOM calculations. The results show that, without the help of Glu64, both proton transfers before and after the formation of the tetrahedral intermediate become partially rate-limiting steps. Taking together all results, we conclude that Glu64 plays critical roles in both binding and catalysis.

In the fifth project, presented in Chapter 6, the functional roles of the active site residues, particularly Asn51 and Asp155, were investigated by site-directed mutagenesis and biochemical, NMR, and computational studies. Steady-state kinetic studies show that both the N51A and D155A variants cause a dramatic decrease in k_{cat} and a dramatic increase in K_{m} , with k_{cat} reduced by a factor of >1000 and K_{m} increased by a factor of >20 . The substitutions also cause a decrease in the binding affinity to 5FPy by a factor of >100 . The MM-PBSA approach (46, 47) was employed to investigate the initial binding of Py with the yCD mutant enzymes. The results show that the binding affinities of Py were reduced by 10.6 and 7.6 kcal/mol with the N51A and D155A enzymes, respectively. Binding of 5FPy with the N51A and D155A mutant enzymes causes a downfield shift in the ^{19}F NMR signal, indicating that the inhibitor is in the unhydrated form. QM/MM MD simulations show that, with the N51A and D155A mutant enzymes, the hydrated form is no longer favored and the activation energy is elevated. The results suggest that similar to Glu64, Asp51 and Glu155 are important not only for the initial binding of the ligand, but also for the stabilization of the transition state.

In conclusion, the catalytic mechanism of yCD was investigated both experimentally and computationally. First, the transition state structure was accessed experimentally by KIE studies. The results indicate that the reaction proceeds by a stepwise $\text{A}_{\text{N}}+\text{D}_{\text{N}}$ mechanism, with the transition state in the C4-N4 bond cleavage. Second, an ONIOM (B3LYP/6-31G*:Amber)-EE scheme was employed to elucidate the catalytic mechanism of the deamination of 5FC and to compare with that of cytosine. The calculations on the pathway for the formation of the zinc-coordinated 5FU are successful.

However, the pathways for the Zn-5FU bond breakage step cannot be properly demonstrated, because a relatively large rearrangement of the active site is required. Hence, QM/MM MD simulations were set up to incorporate configurational sampling. The results show that the C4-N4 bond cleavage is the rate-limiting step for the overall reaction. Functional roles of conserved residues around the active site were investigated by site-directed mutagenesis, biochemical, NMR, and computational methods. Glu64 is found to be essential for catalysis due to its role in proton transfer. It also participates in the binding of ligands. N51 and D155 are important for ligand binding. They also participate in catalysis by stabilizing the transition state.

FIGURES

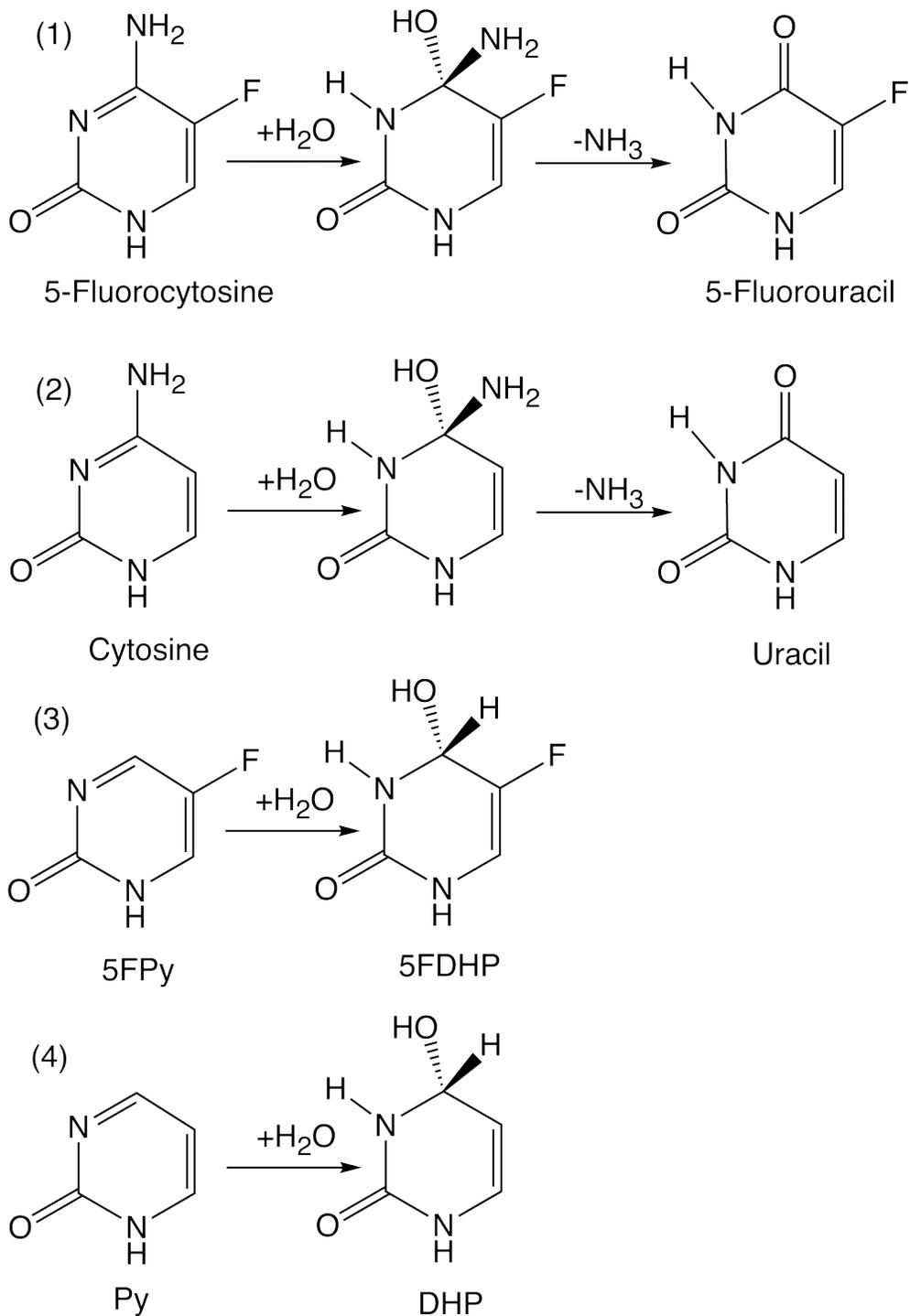


Figure 1-1. Reactions catalyzed by yCD. (1) Deamination of 5FC to 5FU. (2) Deamination of cytosine to uracil. (3) Conversion of 5FPy to 5FDHP. (4) Conversion of Py to DHP.

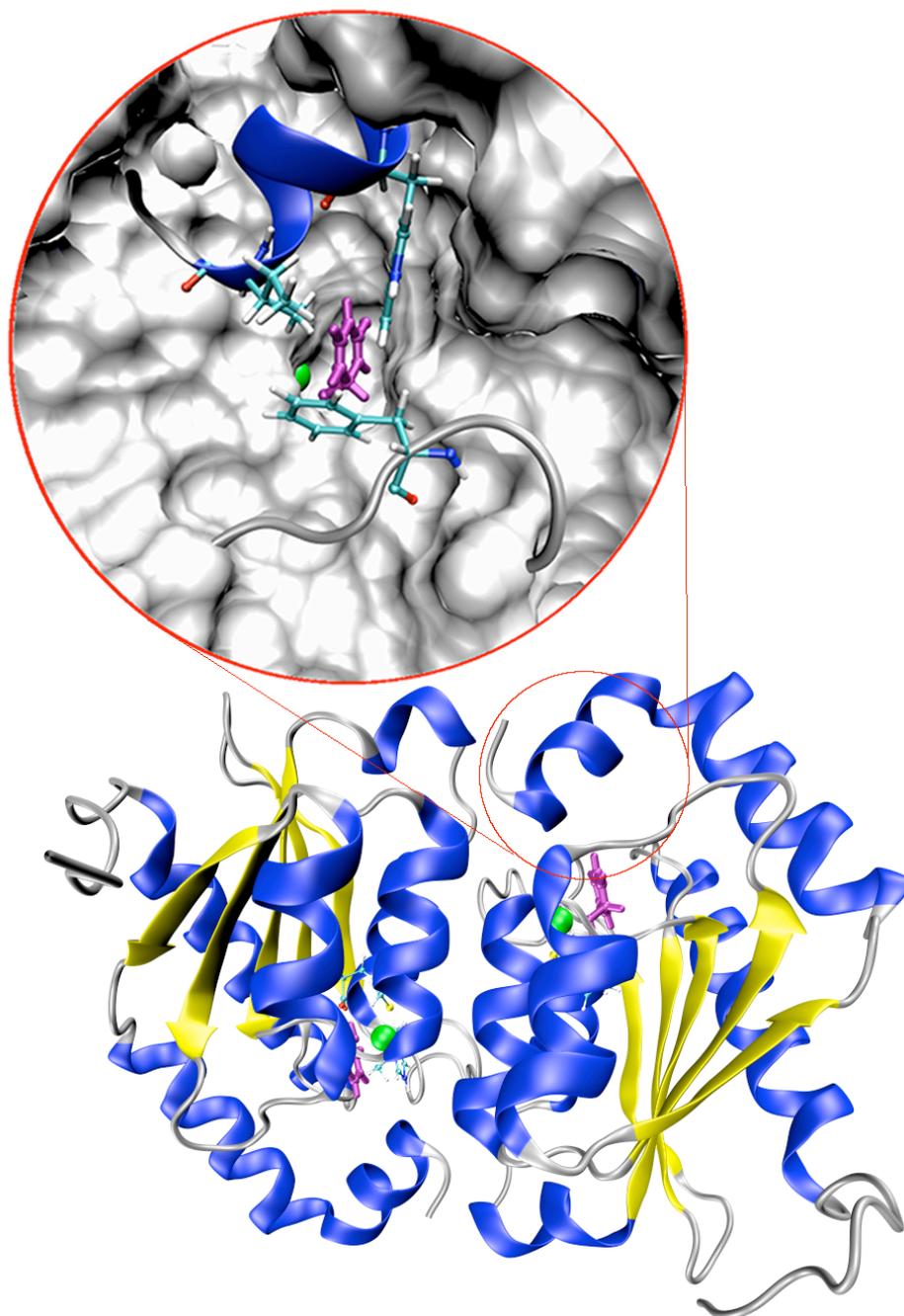


Figure 1-2. Crystal structure of yCD (PDB id: 1p6o). Protein dimer is displayed by ribbon representation with helices colored blue, beta strands colored yellow and the rest colored gray. The zinc atoms are shown as spheres and are colored green. The inhibitor DHP is colored purple. The residues coordinated with the zinc atom are displayed as ball-and-stick. The active site is covered by the C-terminal α -helix and F114 loop, displayed in the red circle. Residues Phe114, Trp152 and Ile156 (displayed in ball and stick) are on top of the inhibitor, covering the active site. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

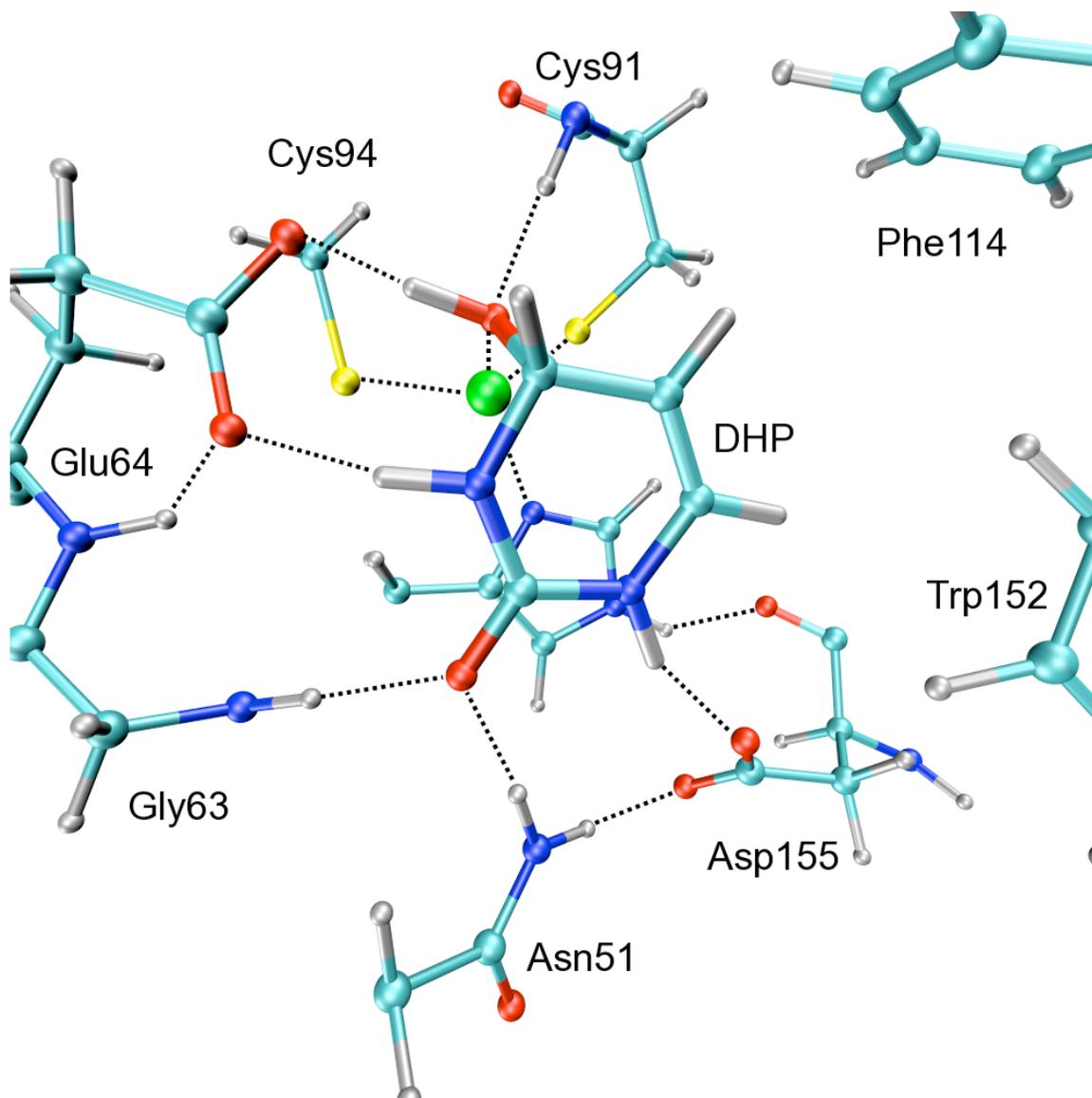


Figure 1-3. Interactions between the inhibitor DHP and surrounding residues in the active site of yCD (PDB id: 1p6o). Hydrogen atoms were added using InsightII. The zinc atom is coordinated with His62, Cys91, Cys94 and the 4-hydroxyl group of DHP.

REFERENCES

REFERENCES

1. Kurtz, J., Exinger, F., Erbs, P., and Jund, R. (1999) New insights into the pyrimidine salvage pathway of *Saccharomyces cerevisiae*: requirement of six genes for cytidine metabolism, *Curr Genet* 36, 130-136.
2. Ko, T., Lin, J., Hu, C., Hsu, Y., Wang, A., and Liaw, S. (2003) Crystal structure of yeast cytosine deaminase - Insights into enzyme mechanism and evolution, *J Biol Chem* 278, 19111-19117.
3. Hahn, A., and Schafer, L. (1925) Uber das Verhalten von Pyrimidinderivaten in den Organismen, *Z Biol* 511-514.
4. O'Donovan, G. A., and Neuhard, J. (1970) Pyrimidine metabolism in microorganisms, *Bacteriol Rev* 34, 278-343.
5. Kream, J., and Chargaff, E. (1952) On the cytosine deaminase of yeast, *J Am Chem Soc* 74, 5157-5160.
6. Ipata, P., Marmocch, F., Magni, G., Feliciol, F., and Polidoro, G. (1971) Bakers yeast cytosine deaminase - Some enzymic properties and allosteric inhibition by nucleosides and nucleotides, *Biochemistry* 10, 4270-4276.
7. Aghi, M., Hochberg, F., and Breakefield, X. (2000) Prodrug activation enzymes in cancer gene therapy, *J Gene Med* 2, 148-164.
8. Greco, O., and Dachs, G. U. (2001) Gene directed enzyme/prodrug therapy of cancer: Historical appraisal and future prospectives, *J Cell Physiol* 187, 22-36.
9. Erbs, P., Regulier, E., Kintz, J., Leroy, P., Poitevin, Y., Exinger, F., Jund, R., and Mehtali, M. (2000) *In vivo* cancer gene therapy by adenovirus-mediated transfer of a bifunctional yeast cytosine deaminase/uracil phosphoribosyltransferase fusion gene, *Cancer Res* 60, 3813-3822.
10. Dachs, G., Hunt, M., Syddall, S., Singleton, D., and Patterson, A. (2009) Bystander or no bystander for gene directed enzyme prodrug therapy, *Molecules* 14, 4517-4545.
11. Kievit, E., Bershad, E., Ng, E., Sethna, P., Dev, I., Lawrence, T., and Rehemtulla, A. (1999) Superiority of yeast over bacterial cytosine deaminase for enzyme/prodrug gene therapy in colon cancer xenografts, *Cancer Res* 59, 1417-1421.
12. Senter, P., Su, P., Katsuragi, T., Sakai, T., Cosand, W., Hellstrom, I., and Hellstrom, K. (1991) Generation of 5-fluorouracil from 5-fluorocytosine by monoclonal-antibody cytosine deaminase conjugates, *Bioconjugate Chem* 2, 447-451.
13. Kievit, E., Nyati, M., Ng, E., Stegman, L., Parsels, J., Ross, B., Rehemtulla, A., and Lawrence, T. (2000) Yeast cytosine deaminase improves radiosensitization and bystander effect by 5-fluorocytosine of human colorectal cancer xenografts, *Cancer*

Res 60, 6649-6655.

14. Mahan, S., Ireton, G., Knoeber, C., Stoddard, B., and Black, M. (2004) Random mutagenesis and selection of *Escherichia coli* cytosine deaminase for cancer gene therapy, *Protein Eng Des Sel* 17, 625-633.
15. Mahan, S. D., Ireton, G. C., Stoddard, B. L., and Black, M. E. (2004) Alanine-scanning mutagenesis reveals a cytosine deaminase mutant with altered substrate preference, *Biochemistry* 43, 8957-8964.
16. Stolworthy, T., Korkegian, A., Willmon, C., Ardiani, A., Cundiff, J., Stoddard, B., and Black, M. (2008) Yeast cytosine deaminase mutants with increased thermostability impart sensitivity to 5-fluorocytosine, *J Mol Biol* 377, 854-869.
17. Korkegian, A., Black, M., Baker, D., and Stoddard, B. (2005) Computational thermostabilization of an enzyme, *Science* 308, 857-860.
18. Fersht, A. (1999) Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding. W. H. Freeman, New York.
19. Ireton, G., Black, M., and Stoddard, B. (2003) The 1.14 angstrom crystal structure of yeast cytosine deaminase: Evolution of nucleotide salvage enzymes and implications for genetic chemotherapy, *Structure* 11, 961-972.
20. Sklenak, S., Yao, L., Cukier, R., and Yan, H. (2004) Catalytic mechanism of yeast cytosine deaminase: An ONIOM computational study, *J Am Chem Soc* 126, 14879-14889.
21. Yao, L., Sklenak, S., Yan, H., and Cukier, R. (2005) A molecular dynamics exploration of the catalytic mechanism of yeast cytosine deaminase, *J Phys Chem B* 109, 7500-7510.
22. Yao, L., Yan, H., and Cukier, R. (2006) A combined ONIOM quantum chemical-molecular dynamics study of zinc-uracil bond breaking in yeast cytosine deaminase, *J Phys Chem B* 110, 26320-26326.
23. Matsubara, T., Dupuis, M., and Aida, M. (2008) An insight into the environmental effects of the pocket of the active site of the enzyme. Ab initio ONIOM-molecular dynamics (MD) study on cytosine deaminase, *J Comput Chem* 29, 458-465.
24. Hsu, Y., Hu, C., Lin, J., and Liaw, S. (2003) Crystallization and preliminary crystallographic analysis of yeast cytosine deaminase, *Acta Crystallogr D* 59, 950-952.
25. Yao, L., Li, Y., Wu, Y., Liu, A., and Yan, H. (2005) Product release is rate-limiting in the activation of the prodrug 5-fluorocytosine by yeast cytosine deaminase, *Biochemistry* 44, 5940-5947.

26. Kornblatt, M., and Tee, O. (1986) Inhibition of yeast cytosine deaminase by 5-bromo-2-pyrimidinone and its covalent hydrate, *Eur J Biochem* *156*, 297-300.
27. Yao, L., Yan, H., and Cukier, R. (2007) A molecular dynamics study of the ligand release path in yeast cytosine deaminase, *Biophys J* *92*, 2301-2310.
28. Liu, A., Li, Y., Yao, L., and Yan, H. (2006) Simultaneous NMR assignment of backbone and side chain amides in large proteins with IS-TROSY, *J Biomol NMR* *36*, 205-214.
29. Liu, A., Yao, L., Li, Y., and Yan, H. (2007) TROSY of side-chain amides in large proteins, *J Magn Reson* *186*, 319-326.
30. Liu, A., Wang, J., Lu, Z., Yao, L., Li, Y., and Yan, H. (2008) Hydrogen-bond detection, configuration assignment and rotamer correction of side-chain amides in large proteins by NMR spectroscopy through protium/deuterium isotope effects, *ChemBioChem* *9*, 2860-2871.
31. Liu, A., Lu, Z., Wang, J., Yao, L., Li, Y., and Yan, H. (2008) NMR detection of bifurcated hydrogen bonds in large proteins, *J Am Chem Soc* *130*, 2428-2429.
32. Xu, Q., Guo, H., Gorin, A., and Guo, H. (2007) Stabilization of a transition-state analogue at the active site of yeast cytosine deaminase: Importance of proton transfers, *J Phys Chem B* *111*, 6501-6506.
33. Yao, L., Liu, A., Li, Y., Cukier, R., and Yan, H. (2004) Dynamic properties of yeast cytosine deaminase as revealed by NMR spectroscopy and molecular dynamic simulations, *Faseb J* *18*, C280.
34. Yao, L., Cukier, R., and Yan, H. (2007) Catalytic mechanism of guanine deaminase: An ONIOM and molecular dynamics study, *J Phys Chem B* *111*, 4200-4210.
35. Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., and Morokuma, K. (1996) ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and $\text{Pt}(\text{P}(t\text{-Bu})_3)_2 + \text{H}_2$ oxidative addition, *J Phys Chem* *100*, 19357-19363.
36. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J.

- V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2004) Gaussian 03, Revision D.01. Gaussian, Inc., Wallingford CT.
37. Shih, P., and Wolfenden, R. (1996) Enzyme-substrate complexes of adenosine and cytidine deaminases: Absence of accumulation of water adducts, *Biochemistry* 35, 4697-4703.
 38. Singleton, D. A., and Thomas, A. A. (1995) High-precision simultaneous determination of multiple small kinetic isotope effects at natural abundance, *J Am Chem Soc* 117, 9357-9358.
 39. Snider, M. J., Reinhardt, L., Wolfenden, R., and Cleland, W. W. (2002) ^{15}N kinetic isotope effects on uncatalyzed and enzymatic deamination of cytidine, *Biochemistry* 41, 415-421.
 40. Luo, M., and Schramm, V. L. (2008) Transition state structure of *E. coli* tRNA-specific adenosine deaminase, *J Am Chem Soc* 130, 2649-2655.
 41. Luo, M., Singh, V., Taylor, E. A., and Schramm, V. L. (2007) Transition-state variation in human, bovine, and *Plasmodium falciparum* adenosine deaminases, *J Am Chem Soc* 129, 8008-8017.
 42. Merkler, D. J., Kline, P. C., Weiss, P., and Schramm, V. L. (1993) Transition-state analysis of AMP deaminase, *Biochemistry* 32, 12993-13001.
 43. Vreven, T., Byun, K. S., Komáromi, I., Dapprich, S., Montgomery, Morokuma, K., and Frisch, M. J. (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM, *J Chem Theory Comput* 2, 815-826.
 44. de M Seabra, G., Walker, R. C., Elstner, M., Case, D. A., and Roitberg, A. E. (2007) Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the amber molecular dynamics package, *J Phys Chem A* 111, 5655-5664.
 45. Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S., and Seifert, G. (1998) Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties, *Phys Rev B* 58, 7260-7268.
 46. Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E. (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models, *Acc Chem Res* 33, 889-897.
 47. Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A., and Case, D. A. (1998)

Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate–DNA helices, *J Am Chem Soc* 120, 9401-9409.

CHAPTER 2

PROBING THE TRANSITION STATE OF YEAST CYTOSINE DEAMINASE BY NMR MEASUREMENT OF ¹³C KINETIC ISOTOPE EFFECTS

INTRODUCTION

Yeast cytosine deaminase (yCD) catalyzes the conversion of cytosine to uracil in the salvage pathways of pyrimidine ribonucleotides. yCD is of great biomedical interest because it is capable of catalyzing the conversion of the prodrug 5FC to the anti-cancer drug 5FU (1, 2). Combination of yCD and 5FC is a promising strategy in gene directed enzyme prodrug therapy (GDEPT) (1-11), in which the administered nontoxic 5FC is converted to the toxic and lethal 5FU in cancer cells.

X-ray crystal structures of yCD have been solved at high resolution either in the free form or the complex form with the inhibitor Py (12, 13). The crystal structure of yCD in complex with Py (Figure 1-3) reveals that the inhibitor is in the hydrated form DHP (12, 13), suggesting that during deamination of cytosine, the substrate may also go through a similar tetrahedral intermediate. Based on this assumption, a complete reaction pathway for deamination of cytosine was proposed by a two-layered quantum mechanics calculation using the ONIOM method (14). The calculations suggest that deamination of cytosine is a step-wise reaction via a tetrahedral intermediate, which is stabilized by

Glu64 and other surrounding residues by a set of hydrogen bonds and π -packing interaction. The formation of the tetrahedral intermediate is facile and the C4-N4 bond cleavage during the decomposition of the tetrahedral intermediate is the rate-limiting step. The transition state was located when the C4-N4 bond distance is 2.10 Å.

The computational study stimulated our interests in finding the transition state of yCD experimentally by analysis of kinetic isotope effects (KIEs). KIEs reflect rate changes of a reaction for isotopically different substrates and have been a powerful probe to understanding of enzymatic mechanisms (15-24). Combing multiple KIEs makes it possible to determine the transition state structure. For example, ^{15}N KIEs have been measured for the deamination of cytidine catalyzed by *E. coli* cytidine deaminase (CDA) (25), one homolog of yCD in the CDA family of purine/pyrimidine deaminases. The results indicate a mechanism involving formation of a tetrahedral intermediate followed by cleavage of the C6-N6 bond. *E. coli* tRNA-specific adenosine deaminase (26), another homolog of yCD, was also found by KIE studies to adopt a stepwise reaction with a late transition state in the N-C bond cleavage. On the contrary, adenosine deaminases from human, bovine and *Plasmodium falciparum* (27) and AMP deaminase from *Saccharomyces cerevisiae* (28), as revealed by KIE studies, have transition states in the formation of the tetrahedral intermediate. Thus, although these enzymes share the same catalytic apparatus, the mechanisms could be different.

In this study, ^{13}C KIEs were measured by NMR spectroscopy for yCD-catalyzed deamination of cytosine at natural abundance (ca. 1.1% for ^{13}C), using the method developed by Singleton and co-workers (29). Hence, the measured KIEs reflected the competitive nature of substrates with different isotopes on the specificity constant

k_{cat}/K_m (30). This method has been successfully applied on non-enzymatic catalyzed (31-37) and enzymatic catalyzed reactions (15, 38). The transition state of yCD was determined by combining ^{13}C KIEs and quantum mechanics calculations. The results indicate that the reaction proceeds by a stepwise $\text{A}_\text{N}+\text{D}_\text{N}$ mechanism, with a transition state during the C4-N4 bond cleavage.

MATERIALS AND METHODS

Materials. Cytosine (>99%) and uracil (>99%) were purchased from Sigma. Dimethyl sulfoxide- D_6 (DMSO-D_6) (99.9%) and D_2O (99%) were purchased from Cambridge Isotope Laboratories. Chromium(III) acetylacetonate ($\text{Cr}(\text{acac})_3$) (99.99%) was purchased from Aldrich. yCD was produced in *E. coli* and purified as previously described (39).

Reaction Setup for KIE Measurement. Four individual reactions for the deamination of cytosine by yCD were set up on a 300 mL scale. For each reaction, the system underwent the following steps. First, cytosine was dissolved in 300 mL 50 mM potassium phosphate buffer (pH 7.0) to a concentration of 15.0 mM. The reaction was initialized by adding 30 μl of 60 μM yCD and kept at 25 $^\circ\text{C}$ until the extent of reaction reached about 90%. The reaction was quenched by boiling for 10 min and adjusting the solution pH to 2.5.

Extents of Reaction. Extents of reaction were monitored by following the absorption change of cytosine at 270 nm (molar extinction coefficient of cytosine = $6 \times 10^{-3} \text{ mol}^{-1} \text{ cm}^{-1}$) on an ultraviolet (UV) spectrometer. The accurate extent of reaction was

determined by ^1H NMR spectroscopy when the reaction was about 90% complete. The NMR sample was made by collecting and lyophilizing 5 mL reaction solution, and dissolving the powder in 1 mL D_2O . The extent of reaction (F) was calculated according to Equation 2-1:

$$F = \frac{INT_{\text{cyt}}}{INT_{\text{cyt}} + INT_{\text{ura}}} \quad (2-1)$$

where INT_{cyt} and INT_{ura} are areas under the H5 peak of cytosine and uracil on the spectrum, respectively.

Purification of Residual Cytosine. After the reaction was quenched, the pH of the solution was adjusted back to 7.0 and the solution was dried by lyophilization. The powder was a mixture of uracil, residual substrate cytosine and phosphate salt. The powder was dissolved in 100 mL methanol to remove the large amount of phosphate. Methanol was removed by centrifuge under vacuum using a Savant SpeedVac system. A solution was made by dissolving the powder in 20 mM phosphate buffer, pH 2.5, and was loaded onto a Vydac 218TP 101522 (22 mm \times 250 mm) C18 peptide and protein column in reverse phase HPLC. 20 mM phosphate, pH 2.5 with 1% methanol was used as elution buffer and the elution was monitored by UV absorption at 256 nm. The retention times for cytosine and uracil were 3.6 and 5.8 min, respectively. The cytosine fraction was collected and adjusted to pH 7.0 before being dried by lyophilization. The powder was dissolved in methanol again to remove the large amount of phosphate. Methanol was removed by centrifuge under vacuum using the Savant SpeedVac system. The final product was the pure residual cytosine, which was not consumed during the reaction. The

residual cytosine and original cytosine from the same package were subjected to ^{13}C NMR analysis described in the next section.

NMR Experiments. NMR samples were prepared using both the original cytosine and residual cytosine. Cytosine was dissolved in 0.7 mL DMSO- D_6 and sealed in 5 mm NMR tube. The concentration of cytosine was ~ 200 mM. $\text{Cr}(\text{acac})_3$ was used as a relaxation reagent to reduce the collection time. The concentration of $\text{Cr}(\text{acac})_3$ was 20 mM. ^{13}C spectra were acquired on a Varian 600 MHz spectrometer at 25 °C.

The spectra were acquired with a spectral width of 32,000 Hz and acquisition time of 0.5 s. ^1H inverse-gated decoupling was applied only during acquisition to suppress the Nuclear Overhauser effect (NOE). During acquisition, the sample tubes were not spun. The 90° pulses were calibrated and spin-lattice relaxation time (T1) was measured by the inversion-recovery method to make sure that the relaxation delay (d1) is at least 5 times T1. The longest T1 for carbon atoms of cytosine was reduced to 0.8 s in the presence of 20 mM $\text{Cr}(\text{acac})_3$. Hence d1 was set at 5.0 s. 48,000 scans were recorded per free induction decay (FID) and an additional 96,000 points were added by zero-filling the FID. Peaks were assigned according to the deposited ^{13}C NMR spectrum of cytosine in DMSO- D_6 from the spectral database for organic compounds (SDBS) (40). ^{13}C chemical shifts were referenced to the solvent peak of DMSO- D_6 ($\delta = 39.52$ ppm) (41). The chemical shifts were 156.910, 166.674, 92.544 and 142.630 ppm for the C2, C4, C5 and C6 atom of cytosine, respectively.

Peak intensity was obtained by integrating the area under a given carbon peak. 10 Hz line broadening was used to improve the signal-to-noise ratio. Baseline correction was achieved using the Whittaker smoother algorithm (42) embedded in the NMR processing and analysis package Mnova (43). Peak intensities of C2, C4 and C5 were normalized relative to that of C6, which was set with a value of 100.

Experimental KIEs. Peak intensities of the carbon atoms in the original cytosine and residual cytosine were obtained by the procedure described above. The enrichment of ^{13}C at a given position was measured by the ratio of the intensity of the peak in the residual cytosine spectrum (R_i) to that in the original cytosine spectrum (R_0). KIE for each carbon atom was calculated by Equation 2-2 (29).

$$\text{KIE} = \frac{\ln(1-F)}{\ln\left[(1-F) \times R_i / R_0\right]} \quad (2-2)$$

where F is the extent of reaction calculated by Equation 2-1. R_i and R_0 for C6 in both original and residual cytosine were the same because C6 was used as the internal standard. Four individual reactions were performed and the KIEs were averaged for each carbon atom.

ONIOM Calculations. A two-layer ONIOM(QM:MM) method (44-46) as implemented in the Gaussian program (Gaussian 03) (47) was employed for the calculations on the yCD-transition state complex. The model system is based on the high-resolution crystal structure of yCD in complex with DHP (PDB code:1p6o) (12).

The system was divided into two layers: inner layer and outer layer (Figure 2-1, top). The whole protein is termed as the real system. The inner layer is termed as the

model system. The overall ONIOM energy E^{ONIOM} is obtained by combining three independent calculations (44-46):

$$E^{\text{ONIOM}} = E^{\text{real,MM}} + E^{\text{model,QM}} - E^{\text{model,MM}} \quad (2-3)$$

where, $E^{\text{real,MM}}$ is the energy for the entire system at the MM level, $E^{\text{model,QM}}$ is the energy for the inner layer at the QM level and $E^{\text{model,MM}}$ is the energy for the inner layer at the MM level. Hence, the ONIOM energy is defined as an extrapolation scheme.

The inner layer is composed of residues Glu64, His62, Cys91, Cys94, Asn51, Asp155, zinc, and the zinc-bound transition state (Figure 2-1, bottom). The inner layer was treated at a high energy level and fully optimized. The rest of the residues are in the outer layer. The outer layer was treated at a low energy level. In the previous study (14), the entire outer layer was restrained during optimization. Freezing of the residues in the outer layer, especially those that have direct interaction with the substrate, may result in unrealistic configurations and/or energy profiles. In this study, instead of restraining all residues in the outer layer, only residues outside of a radius of 10 Å of the substrate were restrained. Residues within a radius of 10 Å of the substrate include residues 18, 21-22, 25, 30-36, 49-50, 52-55, 58-61, 63, 65-68, 84-90, 92, 93, 95-98, 106-111, 113-119, 122, 137, 141, 144-145, 152-158, 223, 251, 254-255 and 257-259, and water molecules 326, 333, 338-339, 341, 353, 359, 376, 378, 380-381, 416, 448, 450, 454, 505, 600, 603, 629, 636, 645, 669, 684, 690-691, 726. Thus, both active site residues and surrounding residues are expected to adopt proper conformation during optimization. Hydrogen atoms were used as link atoms to saturate the dangling bonds for the interface between the inner layer and the outer layer.

The optimization procedure includes three steps. First, the protein was dissolved in a 2-Å water shell and the structure was optimized using the Sander program implemented in the Amber 10 package (48) to remove unfavorable interactions within the protein. The optimized protein was used as the starting structure for the ONIOM calculation. The whole system included 316 amino acids from the protein, 2 zinc atoms, the transition state and 286 additional water molecules. The total number of atoms is 5734. Second, the system was optimized by an ONIOM(B3LYP/6-31G*:Amber) - mechanical embedding (ME) scheme. Third, the electric embedding (EE) formalism (49) was applied in the ONIOM scheme to calculate the interaction between QM and MM. With ME, the electrostatic interaction between the inner layer and the outer layer is treated at the MM level. With EE electronic embedding, the electrostatic interaction is incorporated in the quantum mechanical Hamiltonian (49). Thus, the interaction is better described than in the ME formalism.

B3LYP/6-31+G** is a better basis set to describe a system containing anions by introduction of diffusion functions. However, the optimization is very time-consuming. Optimization was not performed at this energy level; instead, a single-point ONIOM energy was calculated for the optimized structure, using the ONIOM(B3LYP/6-31+G**:Amber) - EE scheme. The results were found to be similar to those presented here.

Potential transition state structures were obtained by systematically increasing the C4-N4 bonding distance at an interval of 0.05 Å. The complete reaction pathway by this ONIOM scheme can be found in chapter 3.

Transition State Structure. Using the experimental KIEs as restraints, the transition state structure was obtained by quantum mechanics calculations in the

Gaussian 03 package (47). According to previous ONIOM calculations (14), the deamination of cytosine is likely to proceed via a tetrahedral intermediate. To determine the transition state structure, a series of potential transition state structures was built in the processes of tetrahedral intermediate formation and decomposition. First, a cluster of structures was built to search for the potential transition state during the tetrahedral intermediate formation by varying the bond distance between O4 and C4. Similarly, a cluster of structures was built to search for the potential transition state during the decomposition of the tetrahedral intermediate by varying the bond distance between C4 and N4. All structures as well as the substrate cytosine were subjected to optimization using the density functional method B3LYP with the basis set 6-31G**. Geometries of the potential transition state optimized by ONIOM calculations were exacted and re-optimized at the same energy level mentioned above, but with restraints on the dihedral angles $\angle\text{N4-C4-N3-C2}$ and $\angle\text{N4-C4-C5-C6}$. Another potential transition state was calculated using the conformations after H3 transfers to O4, as in the tautomer of uracil. Vibrational frequencies were calculated for the optimized structures. All 3N-6 vibrational modes, except the trivial ones, were used to calculate KIEs using the ISOEFF98 package (50). The KIEs of C2, C4, and C5 were normalized relative to that of C6, as applied for the experimental KIEs.

RESULTS AND DISCUSSION

Deamination of Cytosine. yCD stock solutions with different concentrations were tested for the enzyme catalyzed reaction. Four reactions were set up with yCD in a final concentration of 6 nM. Under this condition, the reactions were 90% complete after ca. 16 h. The exact extent of reaction was measured by ^1H NMR spectroscopy. The peaks

for H5 of cytosine ($\delta = -5.83$ ppm) and H5 of uracil ($\delta = -5.66$ ppm) were well separated on the ^1H NMR spectrum. The reaction extents for the four independent reactions were in the range of 86% to 90% (Table 2-1). The residual cytosine was successfully purified from uracil by reverse phase HPLC. The ^1H NMR spectrum confirmed that there was no uracil contamination in the final product.

^{13}C NMR Spectra. The NMR Spectra were acquired for three samples of the original cytosine. Four independent deamination reactions on cytosine from the same bottle were set up and spectra of the residual cytosine were acquired. Peak intensity for each carbon atom was calculated by integrating the area under the peak and normalized relative to that of the C6 atom (Table 2-2 and 2-3). The natural abundance of ^{13}C is around 1.1%, but they may differ from batch to batch. Hence, it is important to compare spectra of residual and starting material from the same batch.

Experimental KIEs. ^{13}C enrichment at a given position was evaluated by comparing the peak intensity of the atom in the residual and original cytosine ^{13}C NMR spectra. KIEs were calculated for C2, C4 and C5 relative to C6 according to Equation 2-2. Hence, by definition, KIE of C6 is unity and KIEs of other carbons are relative. Because C6 is two bonds away from C4, which is involved in the bond formation and breakage during the reaction, KIEs relative to C6 are expected to be close to true KIEs. Four individual reactions were performed and KIEs were averaged for each carbon atom. KIEs were calculated to be 1.002(9), 1.023(4), 1.000(3) for C2, C4, C5, respectively (Table 2-4 and Figure 2-2). The numbers in the parentheses are the third digit of the standard derivation. A relatively large value of C4 KIE (1.023) indicates that the

tetrahedral intermediate formation or decomposition is the first irreversible step (Figure 2-3).

***In vacuo* Transition State Structure.** Crystal structures have been determined at high resolution for yCD with the inhibitor Py (12, 13), which was found to be the hydrated form DHP. The authors proposed that the deamination reaction proceeds through formation of a tetrahedral intermediate. Similarly, crystal structures have been solved for enzymes in the CDA family of purine/pyrimidine deaminase, such as *E. coli* cytidine deaminase (51-53), murine adenosine deaminase (54), *E. coli* tRNA deaminase (55) and *Bacillus subtilis* guanine deaminase (56). All of these structures have transition state based analogs bound in the active site. KIEs (25-28) were measured for these enzymes, based on which, the transition states were accessed. The results showed that the transition state is either in the formation or the decomposition of a tetrahedral intermediate.

These studies suggest that the deamination catalyzed by the enzymes using the same catalytic apparatus is a stepwise reaction via a tetrahedral intermediate, instead of being concerted. The transition state of yCD could be located in either the formation or decomposition of the tetrahedral intermediate (Figure 2-3). Thus, to determine the transition state structure, a series of conformations of potential transition state structures was constructed starting from the tetrahedral intermediate. These conformations represent different states during C4-O4 bond formation or C4-N4 bond cleavage by fixing the C4-O4 bond or the C4-N4 bond at different distances, respectively. Vibrational frequencies were calculated for the optimized structures as well as the substrate cytosine. There was no imaginary frequency for cytosine, suggesting that the optimized structure is a true

ground state minimum. For the transition state, it had one imaginary frequency corresponding to the reaction coordinate.

The KIEs were calculated using all 3N-6 vibrational modes, except the trivial ones. The KIEs of C2, C4, and C5 were normalized relative to that of C6. The primary ^{13}C KIE of C4 was found to match the experimental KIE when the C4-N4 bond distance is 1.88 Å. The optimized transition state (**TS1**) is presented in Figure 2-4, A. A proton atom is located in the middle between O4 and N4, with a distance of 1.21 and 1.31 Å, respectively. It represents the transition state for a concerted reaction including the proton transfer from O4 to N4 and the C4-N4 bond cleavage.

The primary ^{13}C KIE of C4 reflects the change of hybridization of C4 between the substrate and the transition state. The protonation state of O4 and N4 cannot be determined solely based on the ^{13}C KIE of C4. Additional deuterium KIEs of the proton and ^{15}N KIE of N4 are desirable to correctly decipher the transition state structure, which was not measured in this study. Instead, we tried to refine the transition state in the context of the enzyme environment by ONIOM calculations as presented in the section below.

ONIOM Calculations. The complete reaction path for the deamination of cytosine catalyzed by yCD has been proposed by the previous ONIOM calculation (14) and discussed in detail in later chapters (see Chapter 3). The reaction pathway mainly consisted of three events, formation of a tetrahedral intermediate, decomposition of the tetrahedral intermediate and Zn-uracil bond breakage.

The calculated ^{13}C KIE of C4 at the transition state (**TS1**) matched well with the experimental KIE. However, within the enzyme context, the proton may not be directly

transferred to N4 from O4. One study showed that the energy barrier is extremely high for a similar transition state in H₂O-catalyzed deamination of cytosine (57).

A conserved glutamate residue Glu64 was found to be critical to catalysis. Substituting Glu64 with Ala dramatically reduced the catalytic efficiency by 10⁸ times (see Chapter 5). During decomposition of the tetrahedral intermediate, Glu64 abstracts a proton from O4 and transfers to N4. Both the previous and new ONIOM calculations have proposed that the transition state was located in the process of the C4-N4 bond cleavage. To incorporate the input of the enzyme on the transition state, a series of structures was constructed starting from the tetrahedral intermediate, with different distances between C4 and N4. ONIOM optimizations were performed on yCD in complex with these structures. With the increase of the C4-N4 distance, C4 gradually switches from *sp*³ to *sp*² hybridization and the proton is transferred to N4 from the protonated carboxyl group of Glu64. Geometries of these potential transition states optimized by ONIOM calculations (Figure 2-5) were extracted and re-optimized at the same energy level, but with constraints on the dihedral angles ∠N4-C4-N3-C2 and ∠N4-C4-C5-C6. The purpose of imposing the dihedral angle constraints is to reproduce the configuration of the leaving ammonia. Similarly KIEs were obtained using the vibrational frequencies as input.

By comparing the calculated ¹³C KIE of C4 with the experimental KIE, the transition state (**TS2**) structure was found to be in the process of C4-N4 bond cleavage, with a C4-N4 bond distance at 1.93 Å (Figure 2-4, B). In **TS2**, the proton is already transferred to N4 with a distance of 1.02 Å. However, in the enzyme environment, the N4-H bond distance is 1.06 Å. Similarly, the N3-H3 bond distance also decreased, from

1.06 to 1.02 Å. Comparing with **TS1**, the C4-N4 bond increased slightly to 1.93 Å. The bond distance between C4 and O4 decreased greatly, from 1.33 to 1.24 Å. N3 adopts more towards *sp3* from *sp2*, indicated by the increased bond distance between N3 and C4.

Surprisingly, the calculated ^{13}C KIE of C2 was reversed and needs to be explained (Table 2-5). From the QM/MM MD simulation on yCD described in Chapter 4, the newly generated uracil loses a proton in H3 to Glu64 after the C4-N4 bond cleavage. Another potential transition state was calculated using the conformations after H3 transfers to O4, as in the tautomer of uracil. The new transition state (**TS3**) (Figure 2-4, C) was successfully constructed by matching the calculated KIEs with experimental KIEs (Table 2-5). The ^{13}C KIE of C2 was close to the experimental KIE too. The bond distance between C4 and N4 increased to 1.99 Å. Although it matches with the experimental KIEs better, one should be cautious to conclude that **TS3** is the true transition state, because the calculated secondary KIEs may not reflect true KIEs (18) in an enzyme environment. Furthermore, the Glu64 sidechain would need to undergo a large motion to reach the conformation in the transition state (**TS3**).

Commitment to Catalysis. It is important that the measured KIEs really reflect the process of chemical transformation, instead of substrate binding or conformational change. When an enzyme is so efficient that all substrates bound are consumed rapidly to product, instead of being in equilibrium with the free substrate in solution, substrate binding becomes the first irreversible step. Forward commitment can be measured by isotope trapping experiments (58), in which the equilibrium between isotope labeled substrate and enzyme is disrupted by addition of an excessive amount of unlabeled substrate. For example, commitment factors for ADA (27) were measured. However, the

commitment is very small, because the product turnover rate is much slower than the dissociation rate of the substrate. In the case of yCD, the K_m for cytosine is in the range of millimolar and the turnover rate is slow. The enzyme is not a very efficient enzyme with a k_{cat}/K_m of $8.2 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$. The most efficient enzymes, such as catalase, can turn over the substrate so quick that the association of substrate becomes the rating limiting step (18, 30).

The transient kinetics parameters were determined for the better substrate 5FC (39). The dissociation rate constant (k_{-1}) of the Michaelis complex (ES) is only 3 times lower than the forward rate constant k_2 . Most likely, the forward commitment is very small. Reverse commitment is not likely either, because the product ammonia can rapidly diffuse into solution, making the reaction irreversible. Even for the deamination of 5FC, although the rate-limiting step is the product release of 5FU (39), the first irreversible step is prior to it because ammonia has already diffused into solution before the release of 5FU. Hence, the measured KIEs should represent the intrinsic KIEs on k_{cat}/K_m .

Based on the transition state analysis, a reaction mechanism was proposed as shown in Figure 2-6. The substrate can go over a lower energy barrier, if any, and form the tetrahedral intermediate. The first irreversible step is located in the decomposition of the tetrahedral intermediate.

Molecular Electrostatic Potential Surface. The molecular electrostatic potential surfaces were plotted for cytosine, the tetrahedral intermediate, the transition states and uracil (Figure 2-7). The coordinates for cytosine, the tetrahedral intermediate, the transition states are listed in Figure 2-8 to 2-12, respectively. Prior to the O4-C4 bond formation, N3 is protonated and passes more positive charge to the C4 atom, facilitating

the nucleophilic attack on C4. The N4 atom in the tetrahedral intermediate gains some negative charge compared with that of cytosine and is likely to accept a proton. In the first transition state (**TS1**), the proton is trapped between O4 and N4, which may dramatically increase the energy level needed for the departure of ammonia. In the second transition state (**TS2**), the negative charge on N4 is neutralized and NH₃ is ready to be released. Similarly for the third transition state (**TS3**), the C4-N4 bond is ready to be cleaved. In the product uracil, C4 returns to *sp*² hybridization. The primary ¹³C KIE of C4 arises from the change of hybridization state. If the first irreversible step is the substrate binding or the product release, in which there is no hybridization change, the KIE should be close to unity.

Molecular electrostatic potential surfaces were also plotted for the inhibitor 2Py and its hydrated form DHP (Figure 2-7). 2Py is like a substrate mimic and hence cannot effectively inhibit yCD. DHP is more like a tetrahedral intermediate mimic because, like the tetrahedral intermediate, H3 of DHP and HO of DHP form hydrogen bonds with the carboxyl group of Glu64. According to the transition state theory, the protein environment is used to lower the energy of the transition state. If the interactions involved in the transition state stabilization can be immobilized by an inhibitor, the inhibitor is more effective. Although DHP is already a very good inhibitor, it is still possible to design a transition state analog with higher inhibition efficiency.

CONCLUSIONS

Measurement of ¹³C KIEs for the deamination of cytosine catalyzed by yCD at natural abundance was achieved by NMR spectroscopy. The transition state structure was

characterized by quantum chemical calculations. To get a more accurate transition state in the context of enzyme, ONIOM calculations were performed on yCD in complex with transition state structures. The transition state structure was found to be in the process of the C4-N4 bond cleavage, with a C4-N4 bond distance at 1.93 Å. Previous ONIOM calculations on the deamination of cytosine by yCD proposed a reaction pathway in which a similar transition state was obtained. The results indicate that the reaction proceeds by a stepwise A_N+D_N mechanism, with a transition state during the C4-N4 bond cleavage. The late transition state established in this study is similar to that reported for *E. coli* cytidine deaminase and TadA RNA deaminase, but different from the early transition states for adenosine deaminases from human, bovine and *Plasmodium falciparum* and AMP deaminase from *Saccharomyces cerevisiae*.

TABLES AND FIGURES

Table 2-1. Reaction extents for yCD catalyzed reactions.

	Reaction 1	Reaction 2	Reaction 3	Reaction 4
Reaction Extent	87.74%	87.94%	86.49%	89.39%

Table 2-2. Carbon peak intensities for the original cytosine.

	Original 1	Original 2	Original 3	Average	Standard Deviation
C2	101.37	101.32	98.67	100.45	1.54
C4	102.48	101.39	101.25	101.71	0.67
C5	100.36	101.98	98.77	100.37	1.61
C6	100 ^a	100	100	100	0

a: Intensity of C6 was set at 100 and other carbon intensities were relative to C6.

Table 2-3. Carbon peak intensities for the residual cytosine.

Carbon	Reaction 1	Reaction 2	Reaction 3	Reaction 4
C2	101.46	98.16	102.19	101.37
C4	105.75	106.03	107.25	107.5
C5	99.42	100.11	100.76	100.97
C6	100 ^a	100	100	100

a: Intensity of C6 was set at 100 and other carbon intensities were relative to C6.

Table 2-4. Experimental ^{13}C KIEs.

	Reaction 1	Reaction 2	Reaction 3	Reaction 4	Average	Standard Deviation
C2	1.005	0.989	1.009	1.004	1.002	0.009
C4	1.019	1.020	1.027	1.025	1.023	0.004
C5	0.995	0.999	1.002	1.003	1.000	0.003
C6	1 ^a	1	1	1	1	0

a: KIEs were relative to C6.

Table 2-5. Calculated ^{13}C KIEs.

Carbon	KIE _{exp}	TS1 KIE _{calc}	TS2 KIE _{calc}	TS3 KIE _{calc}
C2	1.002(9)	0.987	0.987	0.998
C4	1.023(4)	1.023	1.023	1.022
C5	1.000(3)	0.995	0.998	0.998
C6	1 ^a	1	1	1

a: KIEs were relative to C6.

Table 2-6. Bond distances for cytosine and the transition states.

	Cytosine	Transition State 1	Transition State 2	Transition State 3
N1-C2	1.425	1.392	1.392	1.429
C2-N3	1.372	1.381	1.382	1.368
C2-O1	1.220	1.219	1.222	1.223
N3-C4	1.332	1.415	1.452	1.330
N3-H3	NA	1.011	1.013	NA
C4-O4	NA	1.329	1.236	1.366
C4-N4	1.349	1.880	1.93	1.99
O4-H4	NA	1.206	NA	0.972
N4-H4	NA	1.314	1.019	1.019
C4-C5	1.442	1.474	1.492	1.457
C5-C6	1.358	1.344	1.342	1.351
C6-N1	1.356	1.385	1.387	1.366

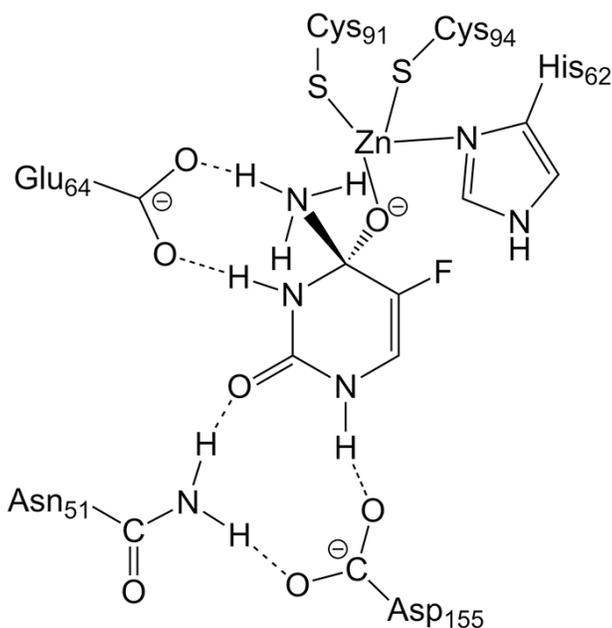
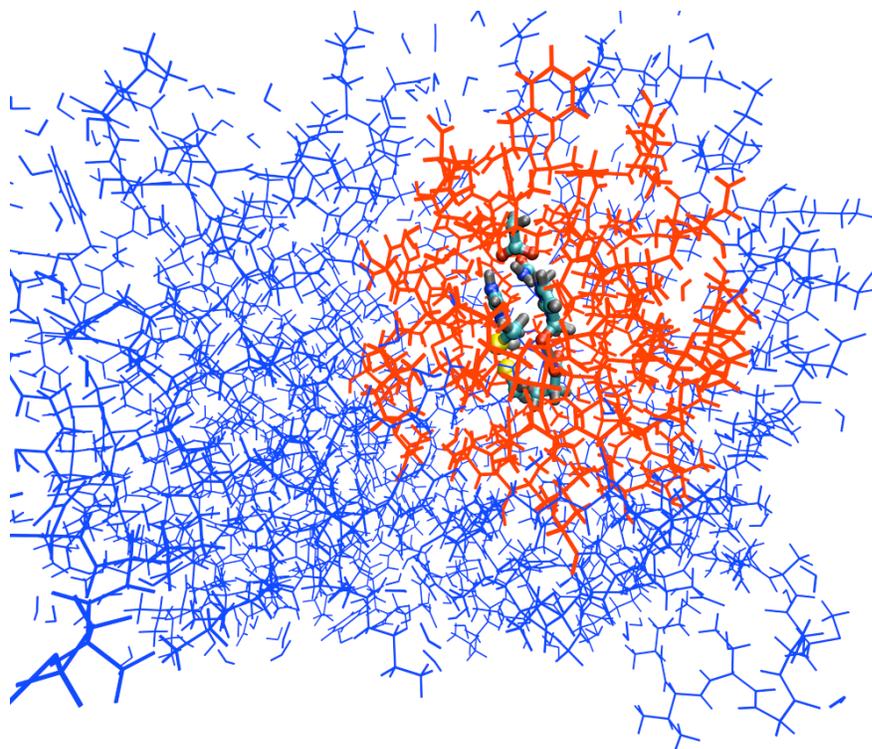


Figure 2-1. Model system for ONIOM calculations. (Top) Partition of the inner and outer layers for the ONIOM calculations. The inner layer is composed of residues Glu64, His62, Cys91, Cys94, Asn51, Asp155, zinc and the TS (displayed by ball and stick in the middle of the active site). The outer layer is composed of the rest of the components. Residues in the outer layer but within 10 Å of the ligand are allowed to move freely (lines in red) and the rest are frozen (lines in blue) during optimization. (Bottom) Schematic drawing of the inner layer.

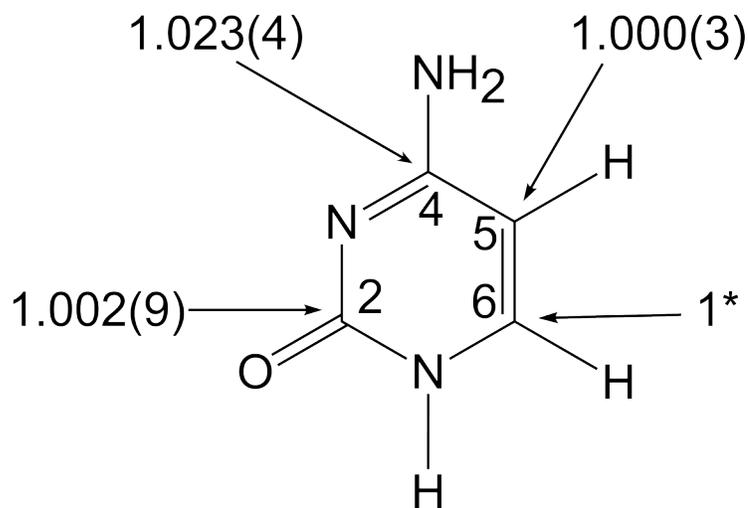


Figure 2-2. Experimental KIEs mapped on the cytosine structure. Number in the parentheses is the third digit of the standard deviation. C6 KIE is used as internal standard.

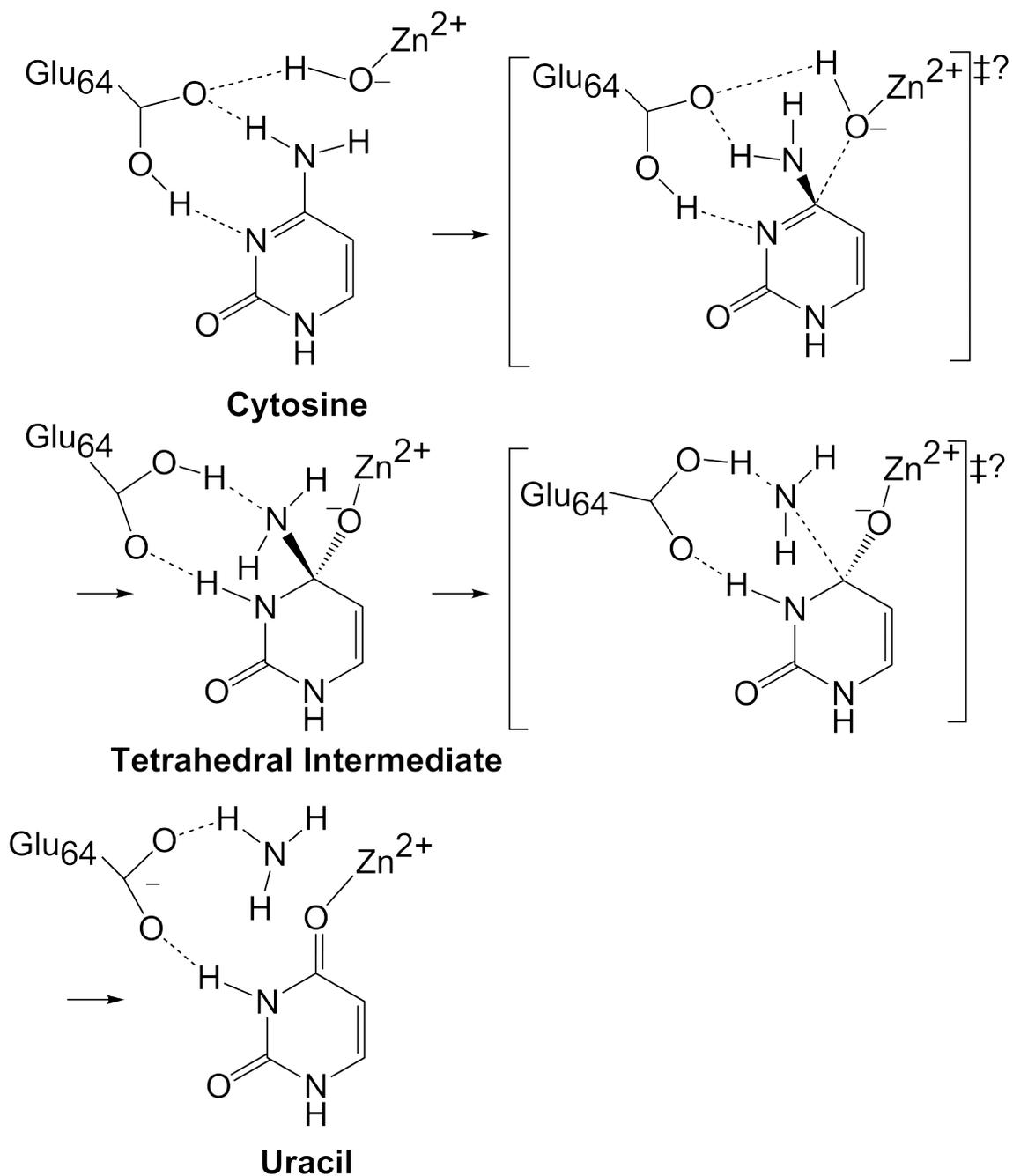


Figure 2-3. Proposed reaction pathway for yCD-catalyzed deamination of cytosine to uracil. As a general base, Glu64 abstracts a proton from the zinc-coordinated water. The hydroxide attacks C4 of cytosine to form a tetrahedral intermediate. Two possible transition state structures were proposed, one in the formation of the tetrahedral intermediate and the other in the decomposition of the tetrahedral intermediate.

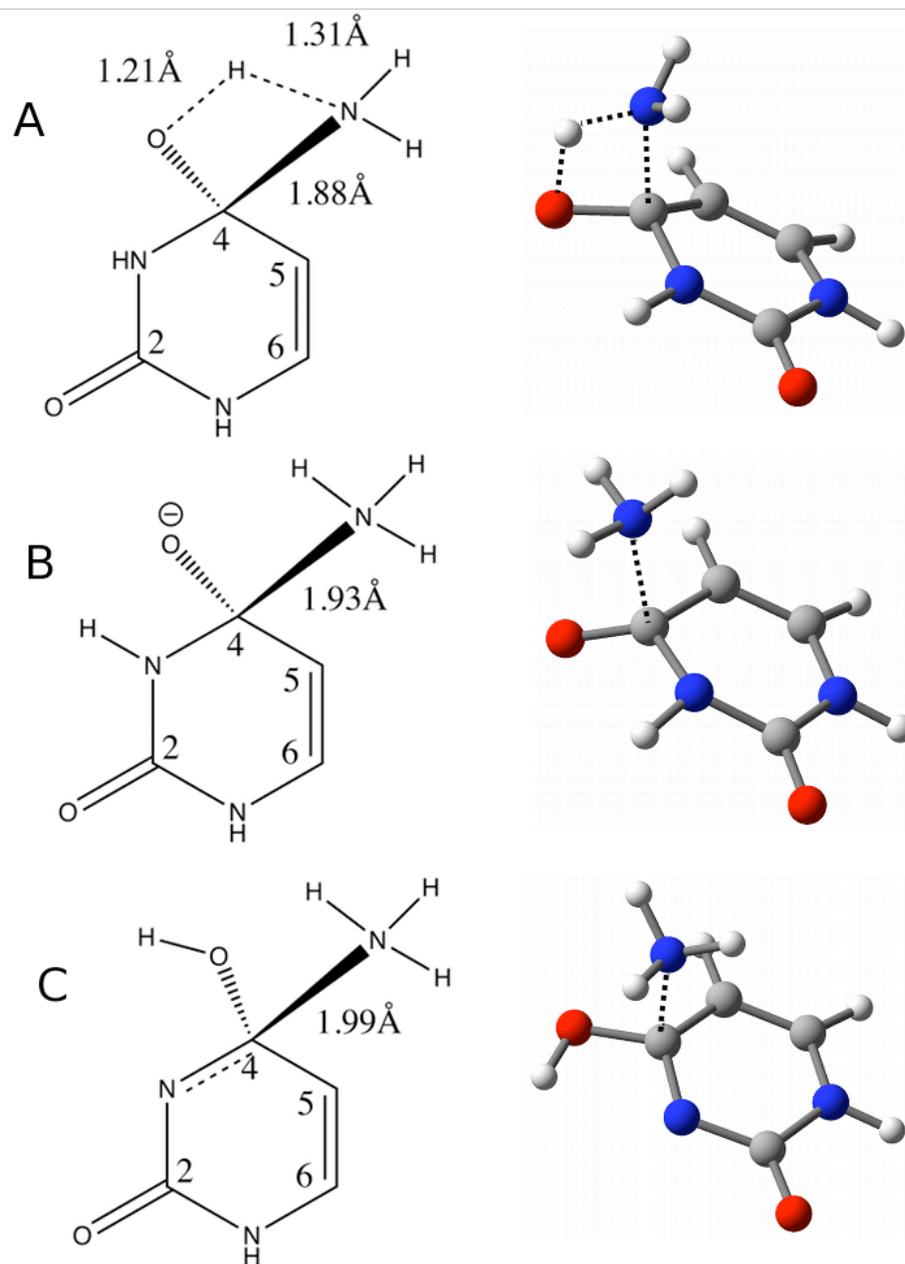


Figure 2-4. Possible transition state structures. A: Transition state for the proton transfer from O4 to N4 with C4-N4 distance of 1.88 Å. B: Transition state with constraints on the N4-C4-N3-C2 and N4-C4-C5-C6 dihedral angles, the values of which were adapted from the ONIOM optimized structure. C: Transition state after H3 transfers from N3 to O4.

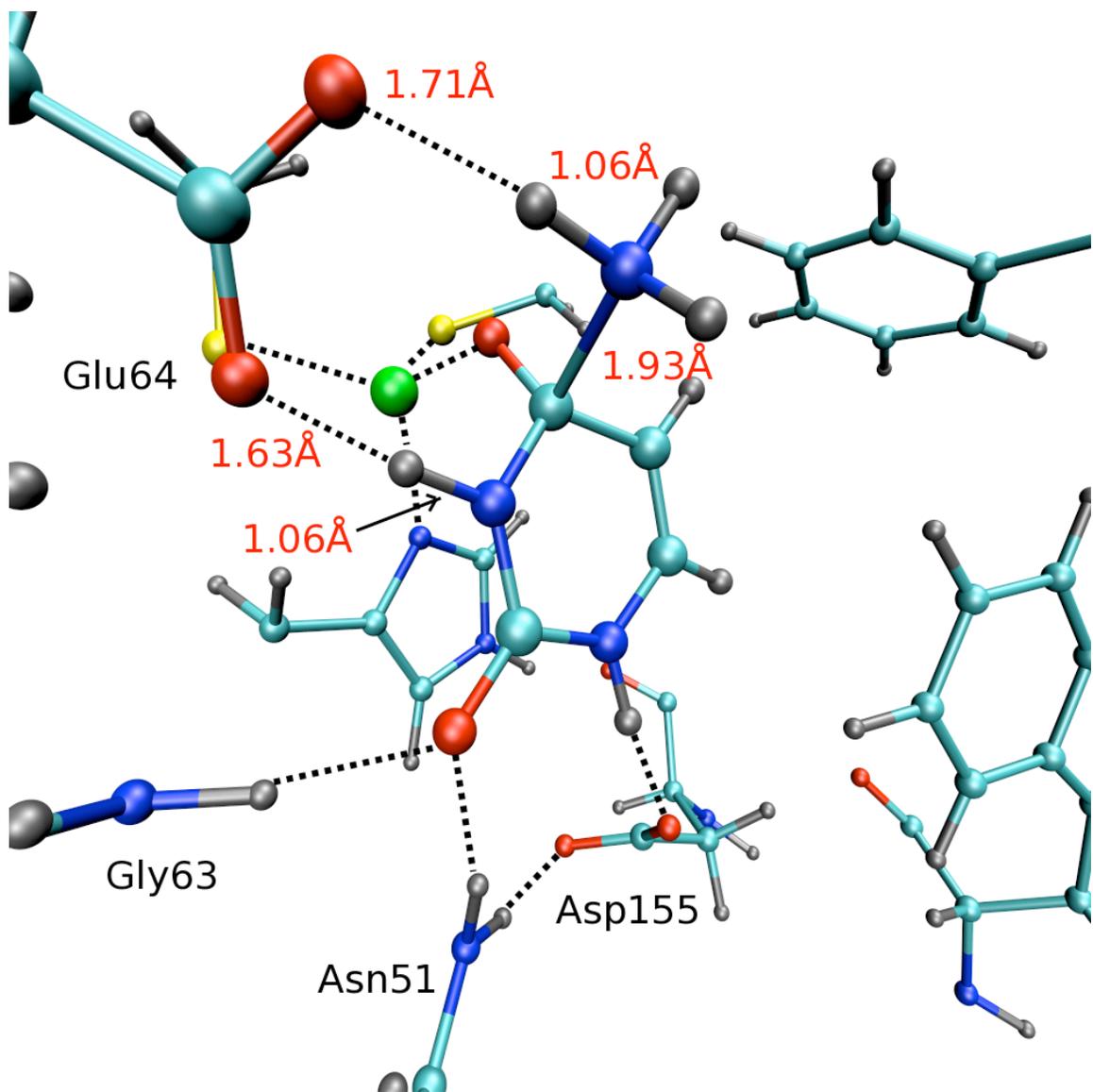


Figure 2-5. ONIOM optimized transition state structure in the active site of yCD.

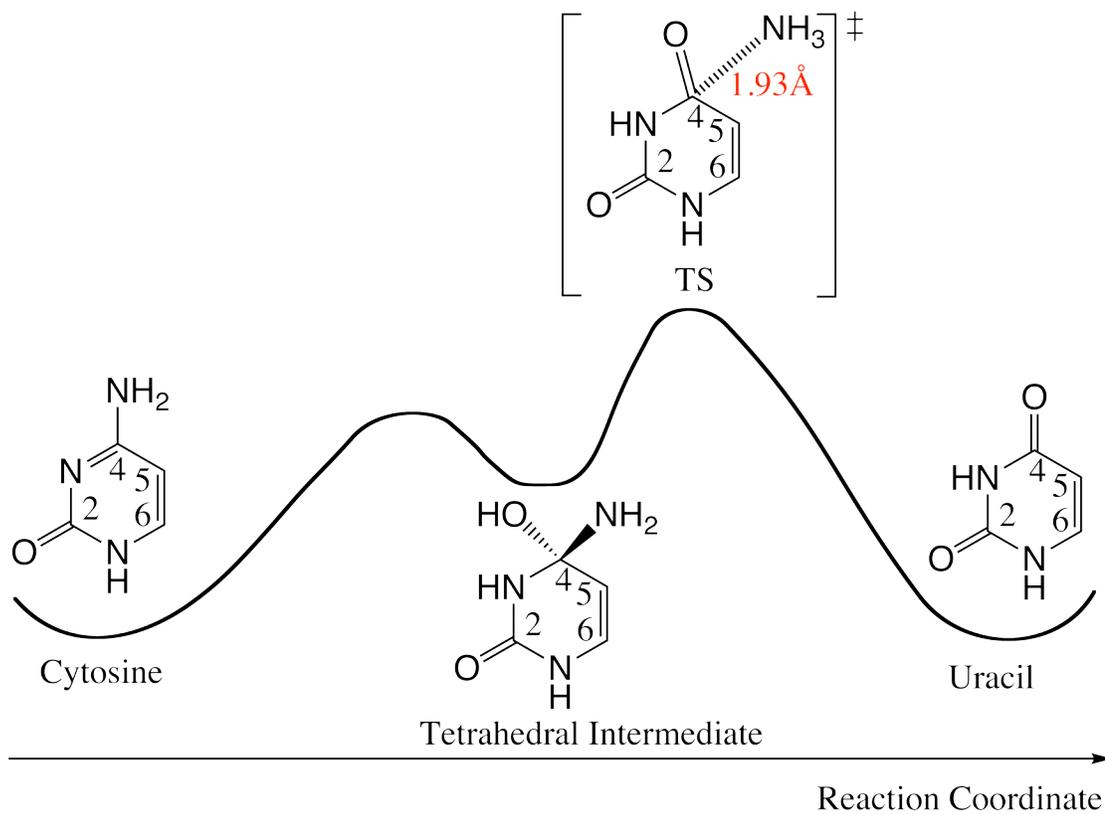


Figure 2-6. Proposed reaction mechanism based on KIEs.

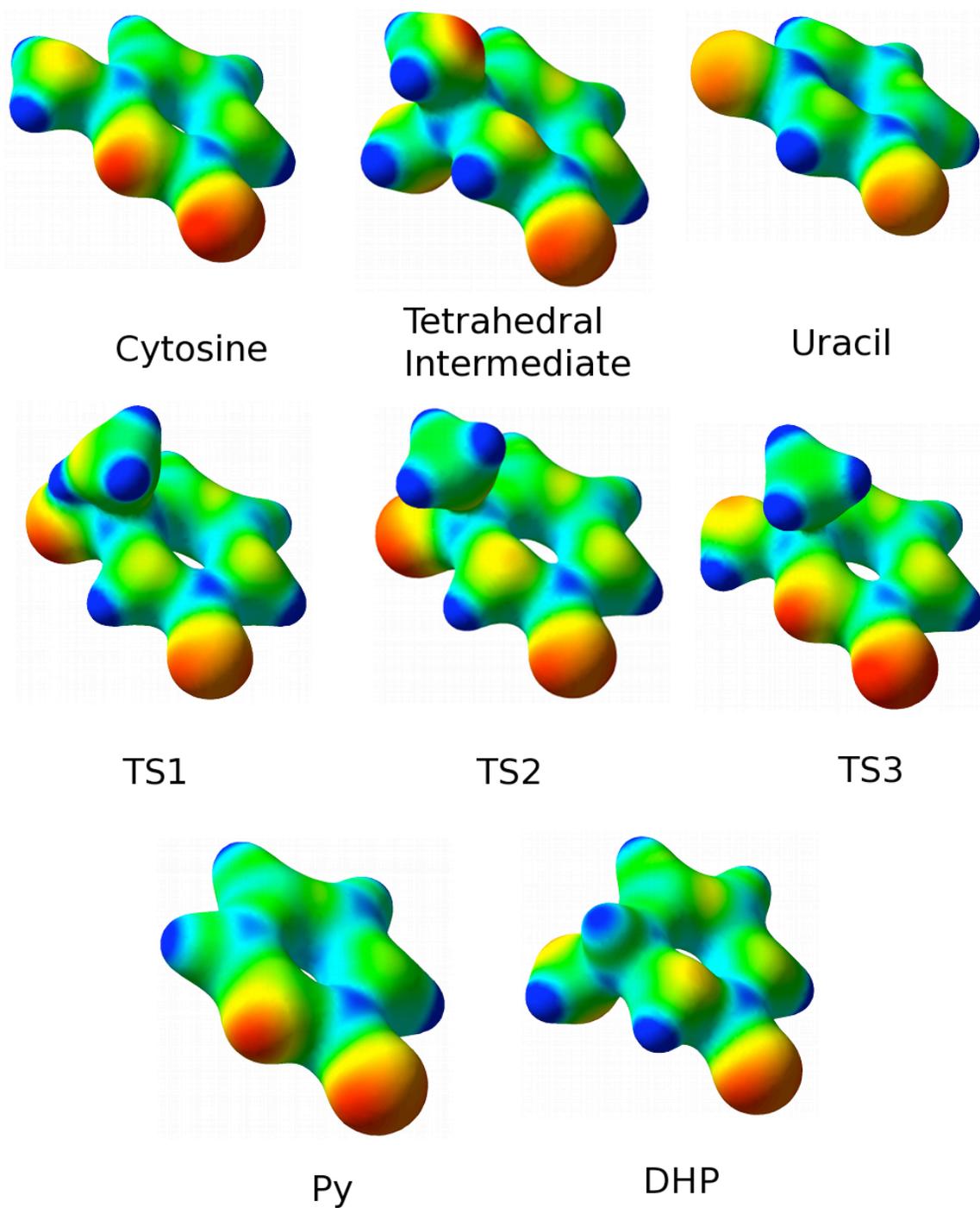


Figure 2-7. Molecular electrostatic potential (MEP) surfaces for cytosine, tetrahedral intermediate, uracil, Py, DHP, and three transition states.

N	-1.279358	0.896720	0.002476
C	-1.187607	-0.529956	-0.000089
O	-2.223508	-1.173897	0.001017
N	0.081818	-1.053172	0.002868
C	1.131126	-0.253206	-0.003479
C	1.052295	1.185972	-0.002115
C	-0.199374	1.715173	0.001656
N	2.358192	-0.845124	-0.041488
H	-2.221649	1.261800	0.005276
H	3.186568	-0.319468	0.183021
H	2.377873	-1.845160	0.091603
H	1.933523	1.814016	-0.013667
H	-0.391450	2.783130	0.002793

Figure 2-8. Coordinates of optimized cytosine.

N	-1.599229	0.881741	0.022234
C	-1.445535	-0.497695	-0.021928
O	-2.396864	-1.264444	0.021043
N	-0.139585	-0.905555	-0.174018
C	1.073877	-0.121820	0.018930
O	1.651511	-0.299420	1.303213
C	0.734951	1.339690	-0.104803
C	-0.534206	1.763635	-0.073670
N	2.042968	-0.649870	-0.941939
H	-2.552252	1.204832	0.075162
H	-0.025887	-1.910527	-0.184581
H	2.910914	-0.125155	-0.859440
H	1.692812	-0.528363	-1.888981
H	2.078347	-1.171024	1.286337
H	1.558224	2.040990	-0.153054
H	-0.802946	2.813089	-0.124607

Figure 2-9. Coordinates of optimized tetrahedral intermediate.

N	1.603991	0.852013	0.224228
C	1.460325	-0.516794	0.001499
O	2.408808	-1.285229	0.018672
N	0.160231	-0.894918	-0.251952
C	-1.023077	-0.079281	-0.250487
O	-2.017052	-0.532690	-1.011433
C	-0.671874	1.367665	-0.258483
C	0.587908	1.764135	-0.022908
N	-1.911794	-0.436782	1.191983
H	2.560422	1.156598	0.324449
H	0.032033	-1.849285	-0.561520
H	-2.247455	0.374598	1.709908
H	-1.479083	-1.099839	1.834785
H	-2.513146	-0.774011	0.165942
H	-1.462138	2.072626	-0.480089
H	0.888629	2.806120	-0.018929

Figure 2-10. Coordinates of optimized transition state 1.

H	-2.321350	-1.431980	1.014592
N	1.616375	0.870920	0.171906
H	2.570046	1.190265	0.242591
C	1.472422	-0.507881	0.050070
O	2.430534	-1.265663	0.072293
N	0.162149	-0.922234	-0.092934
H	0.082255	-1.872979	-0.433825
C	-0.991128	-0.095243	-0.399089
O	-1.787719	-0.487004	-1.259097
C	-0.699134	1.345207	-0.143437
H	-1.503138	2.051249	-0.307932
C	0.564973	1.762689	0.023476
H	0.837689	2.810878	0.082558
N	-2.083146	-0.451247	1.151930
H	-1.827778	-0.316275	2.131648
H	-2.930681	0.079474	0.962364

Figure 2-11. Coordinates of optimized transition state 2.

H	-1.643690	-1.461771	1.459898
N	1.568048	0.832763	0.175106
H	2.526941	1.106920	0.332236
C	1.379391	-0.562229	-0.069774
O	2.351185	-1.299316	0.025193
N	0.112153	-0.947459	-0.411065
H	-1.832686	-1.352316	-1.334939
C	-0.889512	-0.072817	-0.370520
O	-2.020667	-0.445590	-1.038944
C	-0.682683	1.365937	-0.269190
H	-1.497580	2.060466	-0.418920
C	0.579712	1.761061	0.006277
H	0.868235	2.802947	0.101626
N	-1.645961	-0.443588	1.432375
H	-1.034963	-0.120265	2.180497
H	-2.591528	-0.120460	1.633947

Figure 2-12. Coordinates of optimized transition state 3.

REFERENCES

REFERENCES

1. Aghi, M., Hochberg, F., and Breakefield, X. (2000) Prodrug activation enzymes in cancer gene therapy, *J Gene Med* 2, 148-164.
2. Greco, O., and Dachs, G. U. (2001) Gene directed enzyme/prodrug therapy of cancer: historical appraisal and future prospectives, *J Cell Physiol* 187, 22-36.
3. Hamstra, D., Rice, D., Fahmy, S., Ross, B., and Rehemtulla, A. (1999) Enzyme/prodrug therapy for head and neck cancer using a catalytically superior cytosine deaminase, *Hum Gene Ther* 10, 1993-2003.
4. Kievit, E., Bershad, E., Ng, E., Sethna, P., Dev, I., Lawrence, T., and Rehemtulla, A. (1999) Superiority of yeast over bacterial cytosine deaminase for enzyme/prodrug gene therapy in colon cancer xenografts, *Cancer Res* 59, 1417-1421.
5. Erbs, P., Regulier, E., Kintz, J., Leroy, P., Poitevin, Y., Exinger, F., Jund, R., and Mehtali, M. (2000) *In vivo* cancer gene therapy by adenovirus-mediated transfer of a bifunctional yeast cytosine deaminase/uracil phosphoribosyltransferase fusion gene, *Cancer Res* 60, 3813-3822.
6. Kievit, E., Nyati, M., Ng, E., Stegman, L., Parsels, J., Ross, B., Rehemtulla, A., and Lawrence, T. (2000) Yeast cytosine deaminase improves radiosensitization and bystander effect by 5-fluorocytosine of human colorectal cancer xenografts, *Cancer Res* 60, 6649-6655.
7. Nyati, M., Symon, Z., Kievit, E., Dornfeld, K., Rynkiewicz, S., Ross, B., Rehemtulla, A., and Lawrence, T. (2002) The potential of 5-fluorocytosine/cytosine deaminase enzyme prodrug gene therapy in an intrahepatic colon cancer model, *Gene Ther* 9, 844-849.
8. Nyati, M., Sreekumar, A., Li, S., Zhang, M., Rynkiewicz, S., Chinnaiyan, A., Rehemtulla, A., and Lawrence, T. (2002) High and selective expression of yeast cytosine deaminase under a carcinoembryonic antigen promoter-enhancer, *Cancer Res* 62, 2337-2342.
9. Warmann, S., Armeanu, S., Heigoldt, H., Ruck, P., Vonthein, R., Heitmann, H., Seitz, G., Lemken, M., Bitzer, M., Fuchs, J., and Lauer, U. (2009) Adenovirus-mediated cytosine deaminase/5-fluorocytosine suicide gene therapy of human hepatoblastoma *in vitro*, *Pediatr Blood Cancer* 53, 145-151.
10. Dachs, G., Hunt, M., Syddall, S., Singleton, D., and Patterson, A. (2009) Bystander or no bystander for gene directed enzyme prodrug therapy, *Molecules* 14, 4517-4545.
11. Xing, L., Deng, X., Gade, T., Ling, C., Koutcher, J., and Li, G. (2007) Combination of cytosine deaminase with uracil phosphoribosyl transferase with 5-fluorocytosine radiosensitizes rat prostate cancer cells, *Int J Radiat Oncol* 69, S617-S618.

12. Ireton, G., Black, M., and Stoddard, B. (2003) The 1.14 angstrom crystal structure of yeast cytosine deaminase: Evolution of nucleotide salvage enzymes and implications for genetic chemotherapy, *Structure* 11, 961-972.
13. Ko, T., Lin, J., Hu, C., Hsu, Y., Wang, A., and Liaw, S. (2003) Crystal structure of yeast cytosine deaminase - Insights into enzyme mechanism and evolution, *J Biol Chem* 278, 19111-19117.
14. Sklenak, S., Yao, L., Cukier, R., and Yan, H. (2004) Catalytic mechanism of yeast cytosine deaminase: An ONIOM computational study, *J Am Chem Soc* 126, 14879-14889.
15. Lee, J. K., Bain, A. D., and Berti, P. J. (2004) Probing the transition states of four glucoside hydrolyses with ^{13}C kinetic isotope effects measured at natural abundance by NMR spectroscopy, *J Am Chem Soc* 126, 3769-3776.
16. Berti, P. J., and Tanaka, K. S. E. (2002) Transition state analysis using multiple kinetic isotope effects: Mechanisms of enzymatic and non-enzymatic glycoside hydrolysis and transfer, *Adv Phys Org Chem* 37, 239-314.
17. Northrop, D. B. (1981) The expression of isotope effects on enzyme-catalyzed reactions, *Annu Rev Biochem* 50, 103-131.
18. Berti, P. J. (1999) Determining transition states from kinetic isotope effects, in *Enzyme kinetics and mechanism Part E: Energetics of Enzyme Catalysis*, 355-397. Academic Press.
19. Schimerlik, M. I., Rife, J. E., and Cleland, W. W. (1975) Equilibrium perturbation by isotope substitution, *Biochemistry* 14, 5347-5354.
20. Cook, P. F. (2007) Enzyme kinetics and mechanism. Garland Science, London.
21. Schramm, V. L. (2007) Enzymatic transition state theory and transition state analogue design, *J Bio Chem* 282, 28297-28300.
22. Cleland, W. W. (2007) Use of isotope effects to determine enzyme mechanisms, *J Label Compd Radiopharm* 50, 1006-1015.
23. Schramm, V. L. (2007) Binding isotope effects: Boon and bane, *Curr Opin Chem Biol* 11, 529-536.
24. Cook, P. F., Oppenheimer, N. J., and Cleland, W. W. (1981) Secondary deuterium and nitrogen-15 isotope effects in enzyme-catalyzed reactions. Chemical mechanism of liver alcohol dehydrogenase, *Biochemistry* 20, 1817-1825.
25. Snider, M. J., Reinhardt, L., Wolfenden, R., and Cleland, W. W. (2002) ^{15}N kinetic isotope effects on uncatalyzed and enzymatic deamination of cytidine, *Biochemistry* 41, 415-421.

26. Luo, M., and Schramm, V. L. (2008) Transition state structure of *E. coli* tRNA-specific adenosine deaminase, *J Am Chem Soc* 130, 2649-2655.
27. Luo, M., Singh, V., Taylor, E. A., and Schramm, V. L. (2007) Transition-state variation in human, bovine, and *Plasmodium falciparum* adenosine deaminases, *J Am Chem Soc* 129, 8008-8017.
28. Merkler, D. J., Kline, P. C., Weiss, P., and Schramm, V. L. (1993) Transition-state analysis of AMP deaminase, *Biochemistry* 32, 12993-13001.
29. Singleton, D. A., and Thomas, A. A. (1995) High-precision simultaneous determination of multiple small kinetic isotope effects at natural abundance, *J Am Chem Soc* 117, 9357-9358.
30. Fersht, A. (1999) Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding. W. H. Freeman, New York.
31. DelMonte, A. J., Haller, J., Houk, K. N., Sharpless, K. B., Singleton, D. A., Strassner, T., and Thomas, A. A. (1997) Experimental and theoretical kinetic isotope effects for asymmetric dihydroxylation. Evidence supporting a rate-limiting "(3+2)" cycloaddition, *J Am Chem Soc* 119, 9907-9908.
32. Singleton, D. A., Hang, C., Szymanski, M. J., Meyer, M. P., Leach, A. G., Kuwata, K. T., Chen, J. S., Greer, A., Foote, C. S., and Houk, K. N. (2003) Mechanism of ene reactions of singlet oxygen. A two-step no-intermediate mechanism, *J Am Chem Soc* 125, 1319-1328.
33. Meyer, M. P., DelMonte, A. J., and Singleton, D. A. (1999) Reinvestigation of the isotope effects for the claisen and aromatic claisen rearrangements: The nature of the claisen transition states, *J Am Chem Soc* 121, 10865-10874.
34. Nowlan, D. T., Gregg, T. M., Davies, H. M. L., and Singleton, D. A. (2003) Isotope effects and the nature of selectivity in rhodium-catalyzed cyclopropanations, *J Am Chem Soc* 125, 15902-15911.
35. Singleton, D. A., Merrigan, S. R., Kim, B. J., Beak, P., Phillips, L. M., and Lee, J. K. (2000) ¹³C kinetic isotope effects and the mechanism of the uncatalyzed decarboxylation of orotic acid, *J Am Chem Soc* 122, 3296-3300.
36. Keating, A. E., Merrigan, S. R., Singleton, D. A., and Houk, K. N. (1999) Experimental proof of the non-least-motion cycloadditions of dichlorocarbene to alkenes: Kinetic isotope effects and quantum mechanical transition states, *J Am Chem Soc* 121, 3933-3938.
37. Frantz, D. E., Singleton, D. A., and Snyder, J. P. (1997) ¹³C kinetic isotope effects for the addition of lithium dibutylcuprate to cyclohexenone. Reductive elimination is rate-determining, *J Am Chem Soc* 119, 3383-3384.
38. Brecker, L., Kögl, M. F., Tyl, C. E., Kratzer, R., and Nidetzky, B. (2006) NMR study

- of ^{13}C kinetic isotope effects at ^{13}C natural abundance to characterize oxidations and an enzyme-catalyzed reduction, *Tetrahedron Lett* 47, 4045-4049.
39. Yao, L., Li, Y., Wu, Y., Liu, A., and Yan, H. (2005) Product release is rate-limiting in the activation of the prodrug 5-fluorocytosine by yeast cytosine deaminase, *Biochemistry* 44, 5940-5947.
 40. Spectral Database for Organic Compounds, SDBS, http://riodb01.ibase.aist.go.jp/sdbs/cgi-bin/cre_index.cgi
 41. Gottlieb, H. E., Kotlyar, V., and Nudelman, A. (1997) NMR chemical shifts of common laboratory solvents as trace impurities, *J Org Chem* 62, 7512-7515.
 42. Carlos Cobas, J., Bernstein, M. A., Martín-Pastor, M., and Tahoces, P. G. (2006) A new general-purpose fully automatic baseline-correction procedure for 1D and 2D NMR data, *J Magn Reson* 183, 145-151.
 43. Claridge, T. (2009) Software review of MNOVA: NMR data processing, analysis, and prediction software. *J Chem Inf Model* 49, 1136-1137.
 44. Humbel, S., Sieber, S., and Morokuma, K. (1996) The IMOMO method: Integration of different levels of molecular orbital approximations for geometry optimization of large systems: Test for *n*-butane conformation and S_N2 reaction: $\text{RCl} + \text{Cl}^-$, *J Chem Phys* 105, 1959-1967.
 45. Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., and Morokuma, K. (1996) ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels–Alder reactions and $\text{Pt}(\text{P}(t\text{-Bu})_3)_2 + \text{H}_2$ oxidative addition, *J Phys Chem* 100, 19357-19363.
 46. Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K., and Frisch, M. J. (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, *J Mol Struct: THEOCHEM* 461-462, 1-21.
 47. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W.,

- Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2004) Gaussian 03, Revision D.01. Gaussian, Inc., Wallingford CT.
48. Case, D. A., Darden, T. A., Cheatham, I., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossvary, I., Wong, K. F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., and Kollman, P. A. (2008) AMBER 10. University of California, San Francisco.
 49. Vreven, T., Byun, K. S., Komáromi, I., Dapprich, S., Montgomery, Morokuma, K., and Frisch, M. J. (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM, *J Chem Theory Comput* 2, 815-826.
 50. Anisimov, V., and Paneth, P. (1999) ISOEFF98. A program for studies of isotope effects using Hessian modifications, *J Math Chem* 26, 75-86.
 51. Frick, L., Yang, C., Marquez, V. E., and Wolfenden, R. (1989) Binding of pyrimidin-2-one ribonucleoside by cytidine deaminase as the transition-state analogue 3,4-dihydrouridine and the contribution of the 4-hydroxyl group to its binding affinity, *Biochemistry* 28, 9423-9430.
 52. Betts, L., Xiang, S., Short, S. A., Wolfenden, R., and Carter, C. W. (1994) Cytidine deaminase. The 2.3 Å crystal structure of an enzyme: Transition-state analog complex, *J Mol Biol* 235, 635-656.
 53. Xiang, S., Short, S. A., Wolfenden, R., and Carter, C. W. (1995) Transition-state selectivity for a single hydroxyl group during catalysis by cytidine deaminase, *Biochemistry* 34, 4516-4523.
 54. Wilson, D. K., Rudolph, F. B., and Quioco, F. A. (1991) Atomic structure of adenosine deaminase complexed with a transition-state analog: understanding catalysis and immunodeficiency mutations, *Science* 252, 1278-1284.
 55. Kim, J., Malashkevich, V., Roday, S., Lisbin, M., Schramm, V. L., and Almo, S. C. (2006) Structural and kinetic characterization of *Escherichia coli* TadA, the Wobble-specific tRNA deaminase, *Biochemistry* 45, 6407-6416.
 56. Liaw, S., Chang, Y., Lai, C., Chang, H., and Chang, G. (2004) Crystal structure of *Bacillus subtilis* guanine deaminase: the first domain-swapped structure in the cytidine deaminase superfamily, *J Biol Chem* 279, 35479-35485.
 57. Almatarneh, M. H., Flinn, C. G., and Poirier, R. A. (2008) Mechanisms for the deamination reaction of cytosine with $\text{H}_2\text{O}/\text{OH}^-$ and $2\text{H}_2\text{O}/\text{OH}^-$: A computational study, *J Chem Inf Model* 48, 831-843.
 58. Rose, I. A. (1980) The isotope trapping method: Desorption rates of productive ES complexes, *Meth Enzymol* 64, 47-59.

CHAPTER 3

REACTION PATHWAY OF DEAMINATION OF 5FC: ONIOM CALCULATION REVISITED

INTRODUCTION

yCD is of great biomedical interest because of its potential application in gene-directed enzyme prodrug therapy (GDEPT). yCD catalyzes the conversion of 5FC to 5FU, which is widely used in the treatment of breast, rectal, stomach, colon, and pancreatic cancers (1, 2). In GDEPT, cancer cells gain cytosine deaminase activity after a yCD gene is introduced. When the nontoxic prodrug 5FC is administered, it is converted to the active and lethal form 5FU. This strategy takes advantage of the fact that there is no cytosine deaminase expressed in mammalian cells (3). Thus, in principle, only the cancer cells are exposed to the toxic 5FU and are killed, and the normal cells are not affected.

Therapeutic efficacy of GDEPT is affected by many factors, such as efficiency of gene delivery, and stability and half-life of the enzyme, among which the kinetic activity of enzyme is particularly important. The kinetic properties of yCD have been studied by steady-state kinetic experiments for both cytosine and 5FC as substrates (4). 5FC is a better substrate than cytosine regarding binding affinity. The K_m value for 5FC is 7-fold

lower than that for cytosine. However, the turnover rate of 5FC is 5-fold slower than that for cytosine and the catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$) of 5FC is only slightly higher than that for cytosine. Hence, enzyme variants with improved catalytic efficiency for 5FC and/or substrate preference towards 5FC are desirable candidates in GDEPT.

Efforts have been attempted to address substrate specificity towards 5FC by random mutagenesis, directed evolution, and protein engineering. In one such study(5), a yCD variant (D92E) was successfully designed with relative specificity for 5FC over cytosine, although the variant is not better than the wild-type enzyme in cell sensitivity assay. More promising variants have been found for *Escherichia coli* cytosine deaminase (eCD), the counterpart of yCD used in combination with 5FC in GDEPT (6-8). For example, the relative specificity of substrate preference towards 5FC by the eCD/D314A increases 19-fold compared with that by the wild-type eCD.

To facilitate design of potential candidates in GDEPT, the reaction mechanism for the deamination of 5FC and cytosine have been investigated. In Chapter 2, we have shown that the deamination of cytosine by yCD proceeds by a stepwise $A_{\text{N}}+D_{\text{N}}$ mechanism, with the transition state involving C4-N4 bond breakage. The detailed reaction pathway was proposed by previous two-layered ONIOM(B3LYP:PM3) calculations (9). The results suggest that the C4-N4 bond breakage step is the rate-limiting step in generating the zinc-coordinated uracil. The Zn-O4 of uracil bond breakage was further investigated by a two-stepped ONIOM MD calculation (10), which shows that the activation energy for the bond breakage is only 2.9 kcal/mol, indicating that product release is not the rate-limiting step for the overall reaction. The catalytic mechanism of the deamination of 5FC by yCD is different as revealed by transient kinetic

and NMR studies (4). The results show that product release is indeed the rate-limiting step. The rate constants for the forward reaction involving the chemical step are at least 8 times that of the product release step. NMR spectra for yCD in the free form and in the complex form with 5FU were acquired and analyzed. The results show that the K_d of 5FU with yCD is estimated to be around 20 mM, indicating a low affinity of 5FU with yCD. NMR saturation transfer experiments show that 5FU is in slow exchange with its yCD complex on the NMR time scale. The association rate constant and dissociation rate constant are $0.6 \text{ mM}^{-1} \text{ s}^{-1}$ and 14 s^{-1} , respectively.

In this study, an ONIOM electronic embedding (EE) scheme (11) was employed to investigate the reaction mechanism of the deamination of 5FC. The same strategy was applied on cytosine as well to enable direct comparison between the reaction mechanisms. By doing so, we expect to understand the difference between the catalytic mechanisms for 5FC and cytosine as substrate and hopefully to provide a guidance for rational protein design of effective enzyme variants specific towards 5FC.

COMPUTATIONAL METHODS

The reaction pathways for conversion of 5FC to 5FU were calculated by the ONIOM method (12-14) as implemented in the Gaussian program (Gaussian 03) (15). The initial structure is from the high-resolution crystal structure of yCD in complex with DHP (PDB code:1p6o) (16). Instead of only twenty-two amino acids being included, as in the previous ONIOM calculation (9), the whole protein was used to construct the molecular system to incorporate the protein environmental effects (17). The detailed methodology has been described in Chapter 2. In the following paragraphs, I will only

describe briefly how to the calculations were set up.

Two-layered ONIOM (B3LYP/6-31G*:Amber) - mechanical embedding (ME) calculations were performed for the yCD inhibitor complexes, yCD substrate complexes, yCD product complexes and a series of intermediate complexes. The structures were further optimized by an ONIOM (B3LYP/6-31G*:Amber) - electronic embedding (EE) scheme and the ONIOM energies were calculated. We used scans to find approximate geometries of the transition states rather than optimization. Consequently, the calculated values of the barriers are their upper bounds.

RESULTS AND DISCUSSION

yCD-DHP Complex. ONIOM calculations were first performed on the yCD-DHP complex. Key coordinates in the ONIOM optimized structure for the yCD-DHP complex were summarized and compared with those of the crystal structures in Table 3-1. The results show that both ME and EE schemes are sufficient to describe the geometry of the molecular system. Since the zinc atom and its coordination were completely included in the inner layer, they were described at the B3LYP energy level. The bond distances and angles between Zn and its coordination are consistent with those in the crystal structures and those calculated in the previous calculations. Similarly, the conformation of DHP was well described too. DHP adopts a tetrahedral geometry in which the C4 atom shows obvious *sp*³ characteristic. The dihedral angles \angle O4-C4-N3-C2 and \angle O4-C4-C5-C6 are 111° and -107°, respectively, in close agreement with those in the crystal structure (Table 3-1). O4-C4 is clearly a single bond with a distance of 1.45 Å.

DHP is stabilized by a number of hydrogen bonds. O4 of DHP forms a strong hydrogen bond with O^{ε1} of Glu64 with a distance of 2.49 Å. The hydrogen bond is believed to be a low barrier hydrogen bond (LBHB). Indeed, a peak at 16.2 ppm was observed on a 1D proton NMR spectrum of the yCD-DHP complex (unpublished data). Normally, a hydrogen bond showing a proton peak in the range between 16 and 22 ppm is considered a LBHB. Additional experiments are being constructed to further investigate the nature of this hydrogen bond. The bond distances of other hydrogen bonds between DHP and the surrounding residues are in close agreement with those experimentally determined.

Superposition of the optimized structure onto the crystal structure (Figure 3-1) shows that most residues around the active site are well aligned. Different from the previous ONIOM calculation, in which all residues in the outer layer were fixed during optimization, the residues in the outer layer and within a radius of 10 Å of DHP were not restrained. The results show that removal of the restraints did not lead to unreasonable conformation. On the contrary, these residues might adopt a more realistic conformation. For example, the three residues, Trp152, I156 and F114, which cover the active site, move slightly away from their positions in the crystal structure. These residues sit in either the C-terminal helix or the F114 loop, both of which are very flexible fragments, suggested by NMR relaxation experiments (Yao, unpublished data). Binding of DHP may rigidify these fragments through hydrogen bonding, electrostatic and Van der Waals interactions.

Deamination of 5FC. yCD catalyzes the deamination of 5FC to 5FU, as well as the deamination of cytosine to uracil. The latter has been analyzed by previous ONIOM

calculations and a reasonable pathway has been proposed. In this work, we studied the deamination pathway for 5FC by yCD using a different ONIOM scheme. The deamination for cytosine was investigated as well in order to compare with the reaction mechanism of the deamination of 5FC.

ONIOM calculations were performed on a series of yCD substrate, intermediate and product complexes, including (1) water/5FC complex, (2) hydroxide/Glu64H 5FC complex, (3) hydroxide/Glu64 5FC-H complex, (4) protonated tetrahedral intermediate complex, (5) unprotonated tetrahedral intermediate complex, (6) 5FU/ammonia complex, (7) 5FU/water complex and (8) water/5FU complex. Similarly, the calculations were performed for the deamination of cytosine on the same complexes with cytosine and its derivatives. The reaction schemes along the reaction pathways are proposed in Figure 3-2 and the optimized structures for each species are given in Figure 3-3 and 3-4. Same as the previous ONIOM calculation, we used scans to find approximate geometries of the transition states rather than optimization. Consequently, the calculated values of the barriers are their upper bounds.

The reaction pathway starts from the yCD water/5FC complex (1). In 1, 5FC is stabilized by a number of hydrogen bonds. At one end of the pyrimidine ring, N1 forms a hydrogen bond with O^{δ1} of D155 and O2 forms a hydrogen bond with N^{δ2} of Asn51. At the other end, N4 forms a hydrogen bond with O^{ε1} of Glu64. It is also stabilized by hydrophobic interaction between, C5 and C6 of 5FC, and the surrounding nonpolar residues. The zinc-coordinated water forms a hydrogen bond with O^{ε2} of Glu64. The Glu64 side chain can actually adapt two conformations with different $\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\varepsilon 2}$

dihedral angles. In the first conformation (**1**), the Glu64 side chain dihedral angle ($\angle C^\beta-C^\gamma-C^\delta-O^{\epsilon 2}$) is 3.5° . To adapt to the other conformation (**1A**), this angle needs to rotate clockwise by 83° . The energy barrier for the rotation is trivial.

The substrate 5FC is activated by protonating its N3 atom via a two-step proton transfer. First, a proton of the zinc-coordinated water is transferred to $O^{\epsilon 2}$ of Glu64 to form the hydroxide/Glu64H 5FC complex (**2**). In **2**, the Glu64 side chain rotates 36° to form a stable hydrogen bond with N3 of 5FC. **2** is less stable than **1** by 3.9 kcal/mol at ONIOM (Figure 3-5). The $E^{\text{model,QM}}$ energy for the model system (the inner layer plus the link atoms) was calculated to be 5.5 kcal/mol higher than **1**. The transition state (**TS12**) is found in the stage where the hydrogen bond between $O^{\epsilon 2}$ and O^{Zn} is breaking and a new hydrogen bond between $O^{\epsilon 2}$ and N3 is forming (Figure 3-2). The energy barrier was calculated to be 7.7 kcal/mol, suggesting the proton transfer is not a trivial event. The difference of $E^{\text{model,QM}}$ between **1** and **TS12** is 11.0 kcal/mol. The energy difference (3.4 kcal/mol) indicates that the model system is stabilized by the protein environment (the outer layer).

In the next step, the same proton is transferred from $O^{\epsilon 2}$ to N3 to form the complex **3**. **3** is relatively more stable than **2** by 0.3 kcal/mol at ONIOM. $E^{\text{model,QM}}$ of **3** is 1.8 kcal/mol lower than **2**. The transition state (**TS23**) is found when the proton is in the middle of the two heavy atoms. The energy barrier is 2.9 kcal/mol at ONIOM and 0.8

kcal/mol on the $E^{\text{model,QM}}$ energy surface. The results indicate that the environment does not contribute too much for this step.

The distance between O^{Zn} (hydroxide) and C4 is reduced from 2.94 Å, to 2.66 Å, to 2.53 Å from **1**, to **2**, to **3**. In **3**, the zinc-coordinated hydroxide is well positioned for the nucleophilic attack on C4. The transition state (**TS34**) is found when the distance between O4 and C4 is reduced to 1.88 Å. The corresponding energy barrier is calculated to be 2.9 kcal/mol at ONIOM and 6.7 kcal/mol on the $E^{\text{model,QM}}$ energy surface. The energy difference of 3.8 kcal/mol manifests the contribution from the protein environment for the nucleophilic attack process. Complex **4** (Figure 3-3) is formed after nucleophilic attack. In **3**, $O^{\epsilon 1}$ of Glu64 is involved in the hydrogen bond with N4 as a hydrogen acceptor. In **4**, the hydrogen bond is disrupted and $O^{\epsilon 1}$ forms a new hydrogen bond with O4. **4** is found to be more stable than **3** by 5.6 kcal/mol at ONIOM and by 4.5 kcal/mol on the $E^{\text{model,QM}}$ energy surface. Again, the protein stabilizes the model system by 1.1 kcal/mol.

Complex **5** is formed after $O^{\epsilon 1}$ of Glu64 abstracts the proton from O4 and forms a new hydrogen bond with N4 as hydrogen donor (Figure 3-3). The proton is already transferred to $O^{\epsilon 1}$ at the transition state. The energy barrier is mainly contributed to the hydrogen bond breakage between $O^{\epsilon 1}$ and O4 and the rotation of the H41-N4 bond. The energy barrier is calculated to be 3.9 kcal/mol at ONIOM and 5.2 kcal/mol on the

$E^{\text{model,QM}}$ energy surface. The difference between **4** and **5** is the protonation state of the intermediate. In **4**, the intermediate is protonated at O4, and a hydrogen bond is presented between $O^{\epsilon 1}$ of Glu64 and O4. In **5**, the intermediate is unprotonated and a hydrogen bond is present between $O^{\epsilon 1}$ and N4. **5** is more stable than **4** by 5.2 kcal/mol at ONIOM. However, the model system in **5** is less stable than that in **4** by 5.3 kcal/mol. The difference of 10.5 kcal/mol is due to the stabilization effect of the protein environment.

The bond cleavage between N4 and C4 is initialized by the proton transfer from $O^{\epsilon 1}$ to N4. After the proton transfer, N4 keeps moving apart from C4. The transition state (**TS56**) is found when the bond distance is about 2.05 Å. **6** is formed after the N4-C4 bond is completely cleaved. The product ammonia is hydrogen bonded with $O^{\epsilon 1}$ of Glu64. The other product 5FU is covalently bonded with the zinc atom. The energy barrier for this step was calculated to be 5.1 kcal/mol at ONIOM. The energy barrier on the $E^{\text{model,QM}}$ energy surface is 3.0 kcal/mol. However, it is not for the same transition state. The highest point is found on the energy surface during the proton transfer from $O^{\epsilon 1}$ to N4.

Dynamics of the Active Site. Throughout the deamination of 5FC, one end of 5FC ring (the N1 and O2 atoms) is anchored by Asn51 and Asp155 via two hydrogen bonds with N1 and O2, respectively. The other end (the C4 atom) rotates towards the zinc-coordinated water molecule and eventually forms a covalent bond (Figure 3-6). The ring rotates about 39° from the initial position. Obviously, the anchoring residues can

stabilize the enzyme-substrate complex. The transition state structure is stabilized as well, as suggested by mutagenesis studies which show that substituting Asn51 and Asp155 results in great reduction of turnover rate (Chapter 6). Gln156 in eCD functions in a similar way, which forms hydrogen bonds with N1 and O2 by the single side-chain amide group (18). The function of these two residues is further investigated and presented in Chapter 6.

Dynamic effect of the surrounding residues was investigated. Since only energy optimizations were performed in the ONIOM calculations, the dynamic effects cannot be directly observed and evaluated. Instead, it was implied by the movements of the surrounding residues along the reaction pathway. Three fragments, including residues 32 to 34, 62 to 65 and 88 to 91, in the vicinity of the substrate, oscillate gently to accommodate the conformational changes of the reaction species (Figure 3-7). These motions may represent the minimum dynamic efforts required for the reaction, since these fragments are the most rigid regions according to ^{15}N NMR relaxation experiments (Yao, unpublished data) and molecular dynamics simulations. There are close contacts between these fragments and the substrate. The main chain amide group of Gly63 forms a hydrogen bond with O2. The main chain C=O group of Ser89 forms a hydrogen bond with N3. The main chain amide of Cys91 forms a hydrogen bond with the zinc-coordinated water. I33 sandwiches the 5FC ring with the side chain of His62 on the other side. With the progress of the reaction, the position and conformation of these residues change accordingly.

Motions of side chains of F114 and W152 were also observed. These residues sit in the C-terminal helix and the F114 loop, respectively. The conformational change of

these two residues did not prolong to the C-terminal helix and the F114 loop. These two regions are quite flexible according to the ^{15}N NMR relaxation experiments and MD simulations. The movements of these regions are required for the substrate binding and product release (10). The dynamics may not be equally important for the chemical steps in the catalysis. However, their presence is definitely important for maintaining the hydrophobic environment near the C5-C6 side of the pyrimidine ring. Substituting either Phe114 or Trp152 results in reduced turnover rate (unpublished data).

Comparison the Deamination of 5FC and Cytosine. The first transition state (TS12) from complex **1** and **2** for the deamination of 5FC is surprisingly high, with a value of 7.7 kcal/mol at ONIOM. The approximate geometry of TS12 is found by scanning the distance between O^{Zn} and HW (which is the proton transferred to $\text{O}^{\epsilon 1}$ of Glu64). To examine whether it is only for the deamination of 5FC or not, we also performed calculations on the deamination of cytosine using the same scheme. The results show that the energy barrier is almost the same (7.6 kcal/mol). And the $E^{\text{model,QM}}$ energy is quite similar too (11.0 kcal/mol for 5FC and 10.1 kcal/mol for cytosine). It basically means the energy barrier for the first proton transfer is not determined by the substrate. However, the presence of the substrate may prevent the other O^{Zn} -HW bond from rotating and forming a hydrogen bond with $\text{O}^{\epsilon 2}$ of Glu64 and, hence, results in an unrealistically high energy barrier.

The energy barrier for the proton transfer from **2** to **3** is very similar for these two reactions too. This is reasonable since the only difference between 5FC and cytosine is

the fluorine substitution on C5, which is two bonds away from N3. Although fluorine has strong electronegativity, it has little effect on the protonation state of N3.

On the other hand, more pronounced effect of fluorine substitution on the tetrahedral intermediate formation in the next step is evidenced by the energy profile. The energy barrier (TS34) from **3** to **4** for the deamination of 5FC is only 2.9 kcal/mol, 2.4 kcal/mol lower than that for cytosine. Fluorine pulls electrons towards itself, leading the C4 atom more liable to nucleophilic attack by the zinc-coordinated hydroxide. Under the same logic, **4** is more stable than **3** by 5.6 kcal/mol for 5FC. While for cytosine, **4** is only more stable than **3** by 1 kcal/mol.

For the C4-N4 bond cleavage, a positively charged leaving group ($-\text{NH}_3^+$) is more likely to depart from 5FC, because C4 is more positive due to the presence of fluorine on C5. The energy barrier is 5.1 kcal/mol for this step for the deamination of 5FC, while it is 6.3 kcal/mol for cytosine.

Steady-state kinetic constants have been measured for both 5FC and cytosine (4). The K_m for cytosine (1.1 mM) is about 7 times higher than that for 5FC (0.16 mM). For a simple Michaelis-Menten reaction, K_m is highly related to the binding affinity between the enzyme and substrate. While for a reaction involving multiple intermediates, like the deamination of 5FC by yCD, K_m is not just that of the enzyme-substrate complex, but rather reflects the overall binding affinity with all intermediate complexes (19). In the deamination of 5FC, relatively stable intermediate complexes **4** and **5** are formed during the reaction (Figure 3-5), which results in a smaller K_m for 5FC. This finding is interesting for protein engineering of effective yCD variants specific towards 5FC. If an

enzyme variant can lead to strong interaction with these intermediates, it would probably be a better candidate for GDEPT.

For the deamination of 5FC, the energy barriers for each step are smaller when compared with the same step for the deamination of cytosine. If the C4-N4 cleavage step is the rate-limiting step for both reactions, the overall turnover rate k_{cat} of 5FC would be higher than that of cytosine. However, k_{cat} of 5FC is 17 s^{-1} , 5 times lower than that of cytosine (91 s^{-1}). This means this step cannot be rate-limiting for the deamination of 5FC. Indeed, transient kinetic experiments and NMR experiments have shown that product-release is the rate-limiting step for the deamination of 5FC. KIE studies have suggested that the C4-N4 cleavage step is the rate-limiting step for the deamination of cytosine. A two-step ONIOM MD simulation has suggested product release is not the rate-limiting step for the deamination of cytosine. The energy barrier is calculated to be 2.9 kcal/mol. However, the energy barrier for the C4-N4 cleavage step is not compared using the same protocol. In the next section, the Zn-5FU bond breakage is evaluated using the same scheme as shown above.

Zn-5FU Bond Breakage. Ammonia can rapidly diffuse into the solution once it is generated (20). However, direct cleavage of the bond between 5FU and Zn is surprisingly hard according to the previous ONIOM calculation. An alternative mechanism involving oxygen exchange was proposed and is energetically better than direct cleavage. However, oxygen exchange experiments show the exchange rate with 5FU is too slow to be the main route for the 5FU release. A further study combining ONIOM and MD simulation shows rearrangement of the active site residues is needed for the Zn-O4 bond

cleavage (10), which cannot be achieved in the previous ONIOM calculation due to the restraints on the outer layer atoms. In that study, the author proposed that 5FU is released with the help of the carboxyl group of Glu64, which replaces O4 of 5FU to coordinate with Zn.

In this study, we tried to investigate if the proper transition state can be located by the ONIOM scheme alone if the active site residues are set free. Moreover, we want to see if direct replacement of 5FU with a water molecule is a possible pathway (Figure 3-8).

Complex **7** was constructed by replacing the ammonia with a water molecule. The reaction pathway was constructed by a series of calculations with different distances between Zn and OW of water. The optimized structures are presented in Figure 3-9 and 3-10. The energy profile is plotted in Figure 3-11. When the water is approaching Zn, the ONIOM energy increases accordingly. The energy reaches a plateau when the Zn-OW distance decreases to ~ 4.0 Å, with a value of more than 20 kcal/mol. The energy of the system stays at a high level until the Zn-OW distance reduces to 2.6 Å, when the Zn-OW bond is about to form and the Zn-O4 bond is about to break. Complex **8** is formed when the water completely replaces 5FU in coordination with Zn. **8** is less stable than **7** by 2.0 kcal/mol.

At first glance, the high energy barrier is reasonable because product release is supposed to be the rate-limiting step for the overall reaction of deamination of 5FC as revealed by transient kinetics and NMR experiments. However, the energy profile is too high to be true. Moreover, a similar energy barrier was obtained for the bond cleavage of Zn-uracil, which is not the rate-limiting step for the deamination of cytosine as suggested

by our KIE studies and the two-stepped ONIOM MD simulations (Figure 3-11).

The $E^{\text{model,QM}}$ energy is much lower than the ONIOM energy after the Zn-OW distance reduces to 4.2 Å, suggesting that the protein environment does not help, but impedes the reaction. In other words, although the surroundings are set free during the optimization, proper arrangement is not achieved. Figure 3-12 shows that the water accesses Zn through the space between Glu64 and Pro90. Glu64 and Pro90 are pushed away from their original positions. Other regions should move accordingly to reach the conformation of the global minimum along the reaction pathway. However, the conformation is settled in a nearby local minimum, which results in an unrealistically high energy at ONIOM. This suggests that for a large rearrangement, a simple ONIOM optimization cannot properly sample the right reaction path.

Hence, proper sampling technique is needed to assess the Zn-5FU bond breakage pathway. The two-stepped ONIOM MD approach is a good method and has been successfully applied to uracil release with the assist of Glu64 carboxyl group. Such an approach can also be applied on the 5FU release.

CONCLUSIONS

The deamination of 5FC catalyzed by yeast cytosine deaminase was investigated using the ONIOM (B3LYP/6-31G*:Amber) - EE scheme. In the molecular system, the whole protein was included in order to properly estimate contributions from the protein environments.

yCD catalyzes the deamination of 5FC via a sequential mechanism, including substrate activation, tetrahedral intermediate formation and C4-N4 bond cleavage. 5FC is activated by protonation at N3 via $O^{\epsilon 2}$ of Glu64. The zinc-coordinated water is converted

to a hydroxide after losing a hydrogen atom and is well positioned for nucleophilic attack on C4. The tetrahedral intermediate is formed after nucleophilic attack and the deprotonated tetrahedral intermediate is formed after $O^{\epsilon 1}$ of Glu64 abstracts the hydrogen on O4 and forms hydrogen bond with N4. In the next step, the C4-N4 bond is cleaved, in which ammonia and zinc-coordinated 5FU are generated.

Throughout the deamination of 5FC, one end of the 5FC ring (the N1 and O2 atoms) is anchored by Asn51 and Asp155 through two hydrogen bonds. The other end (the C4 atom) moves towards the zinc-coordinated water molecule and eventually forms a covalent bond. From the initial 5FC to the zinc-coordinated 5FU, the ring rotates about 38° . Dynamics of the three fragments, including residues 32 to 34, 62 to 65 and 88 to 91, in the vicinity of the substrate, may facilitate the progress of the reaction. Interestingly, these fragments are the most rigid regions in the protein according to NMR relaxation experiments and molecular dynamics simulations. These motions may represent the minimum dynamic effects required for the reaction.

The energy profile shows that the tetrahedral intermediate complexes are more stable than the initial complex during the deamination of 5FC. However, the stabilization of these intermediates during the deamination of cytosine is relatively weak. It could be at least part of the reason why K_m for cytosine is higher than that of 5FC.

For the Zn-5FU bond breakage, we proposed that a water molecule attacks the zinc atom as a nucleophile and replaces 5FU in coordination with Zn. The access of the zinc atom by the water is through the space between Glu64 and Pro90. The close contacts disturb the conformation of these residues and thus results in unrealistic high energies at

ONIOM, though the $E^{\text{model,QM}}$ energy surface is reasonable. The results suggest that a sole energy optimization starting from a static structure cannot sample the reaction pathway if large rearrangements are involved.

TABLES AND FIGURES

Table 3-1. Comparison of key coordinates of structures from ONIOM calculations with crystal structures ^a

Internal Coordinate	X-ray Structure (1uaq)	X-ray Structure (1p6o)	ONIOM ME	ONIOM EE
His62-N ^{δ1} ... Zn	2.07, 2.06 ^b	2.01, 2.00	2.03	2.01
Cys91-S ^γ ... Zn	2.35, 2.31	2.29, 2.30	2.32	2.36
Cys94-S ^γ ... Zn	2.27, 2.25	2.27, 2.27	2.30	2.29
DHP -O4 ... Zn	2.06, 2.08	2.05, 2.06	2.05	2.03
Cys91-N ... O4-DHP	3.02, 3.10	3.02, 3.03	3.11	3.13
Glu64-N ... O ^{ε1} -Glu64	2.69, 2.70	2.69, 2.72	2.73	2.73
DHP-O4 ... O ^{ε2} -Glu64	2.60, 2.45	2.49, 2.52	2.48	2.49
DHP-N3 ... O ^{ε1} -Glu64	2.82, 2.81	2.80, 2.78	2.94	2.93
DHP-O2 ... N-Gly63	2.84, 2.79	2.89, 2.87	2.89	2.84
DHP-O2 ... N ^{δ2} -Asn51	2.96, 2.95	2.92, 2.90	3.07	3.03
DHP-N1 ... O ^{δ1} -Asp155	2.67, 2.70	2.69, 2.68	2.73	2.70
Asn51-N ^{δ2} ... O ^{δ2} -Asp155	2.95, 2.98	2.92, 2.90	2.89	2.91
DHP-O4 ... C4-DHP	1.44, 1.44	1.48, 1.48	1.44	1.45
DHP-N3 ... C4-DHP	1.42, 1.41	1.46, 1.45	1.45	1.45
DHP-C5 ... C4-DHP	1.45, 1.45	1.47, 1.46	1.49	1.49
DHP-N3 ... C2-DHP	1.37, 1.37	1.31, 1.31	1.37	1.36
DHP-N1 ... C2-DHP	1.39, 1.39	1.32, 1.33	1.38	1.37
DHP-N1 ... C6-DHP	1.38, 1.38	1.43, 1.42	1.38	1.38
DHP-C5 ... C6-DHP	1.34, 1.35	1.31, 1.33	1.34	1.34
DHP-O2 ... C2-DHP	1.21, 1.21	1.30, 1.30	1.24	1.25
∠O4-C4-N3	110, 109	107, 108	111	111
∠O4-C4-C5	110, 109	112, 109	111	111
∠O4-C4-N3-C2	113, 119	113, 114	105	107
∠O4-C4-C5-C6	-125, -116	-113, -114	-110	-111
DHP-HO4 ... O4-DHP			1.07	1.06
DHP-HO4 ... O ^{ε2} -Glu64			1.40	1.43

a: Distances between the paired atoms are in Å. Angles and dihedral angles are in degree.

b: Distances in subunits A and B, respectively.

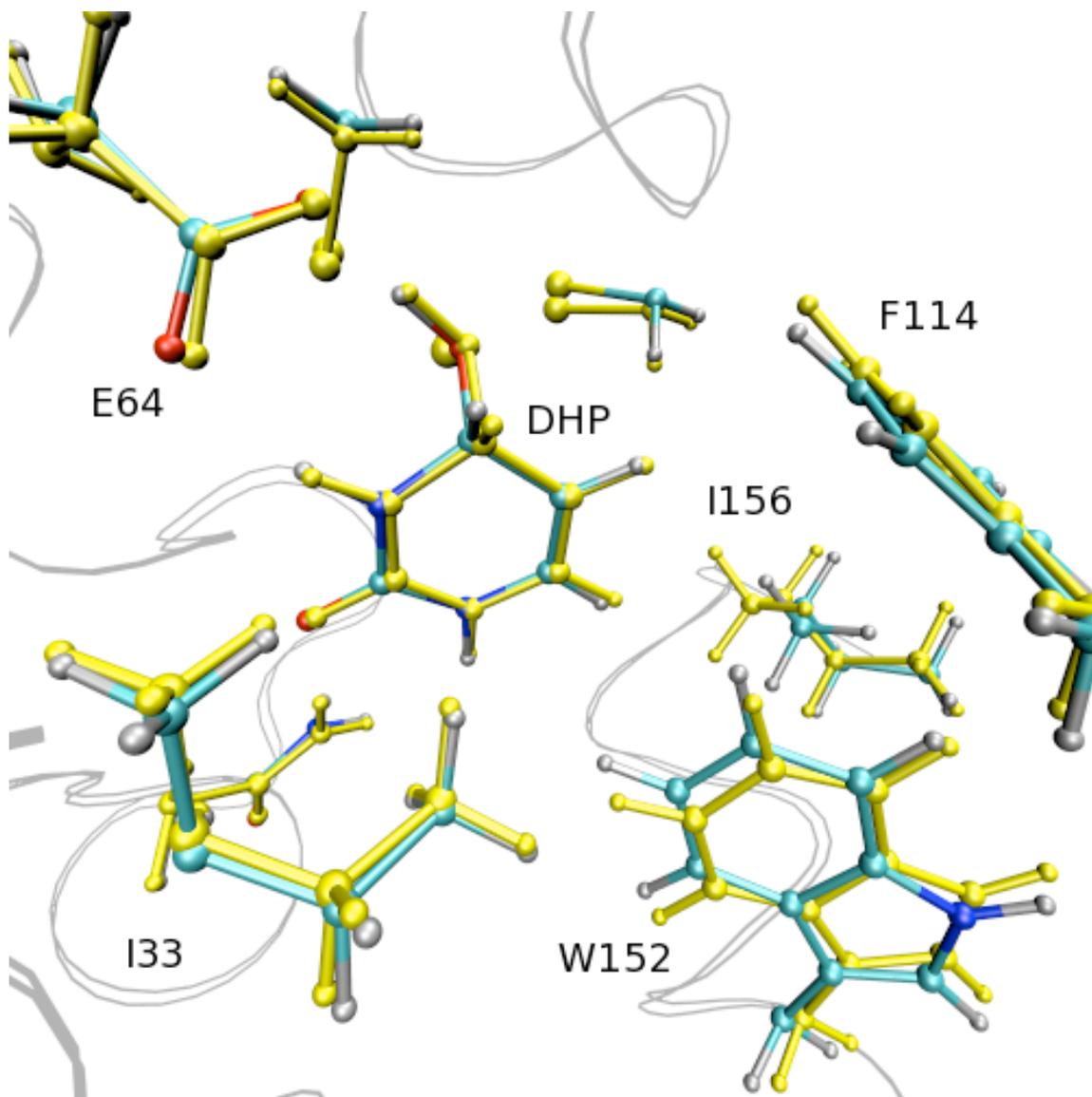


Figure 3-1. Superposition of the ONIOM optimized yCD-DHP complex structure with the crystal structure (PDB id: 1p6o). The backbone is displayed as a gray tube. DHP and several surrounding residues are displayed as ball and stick. The residues are colored yellow for the crystal structure.

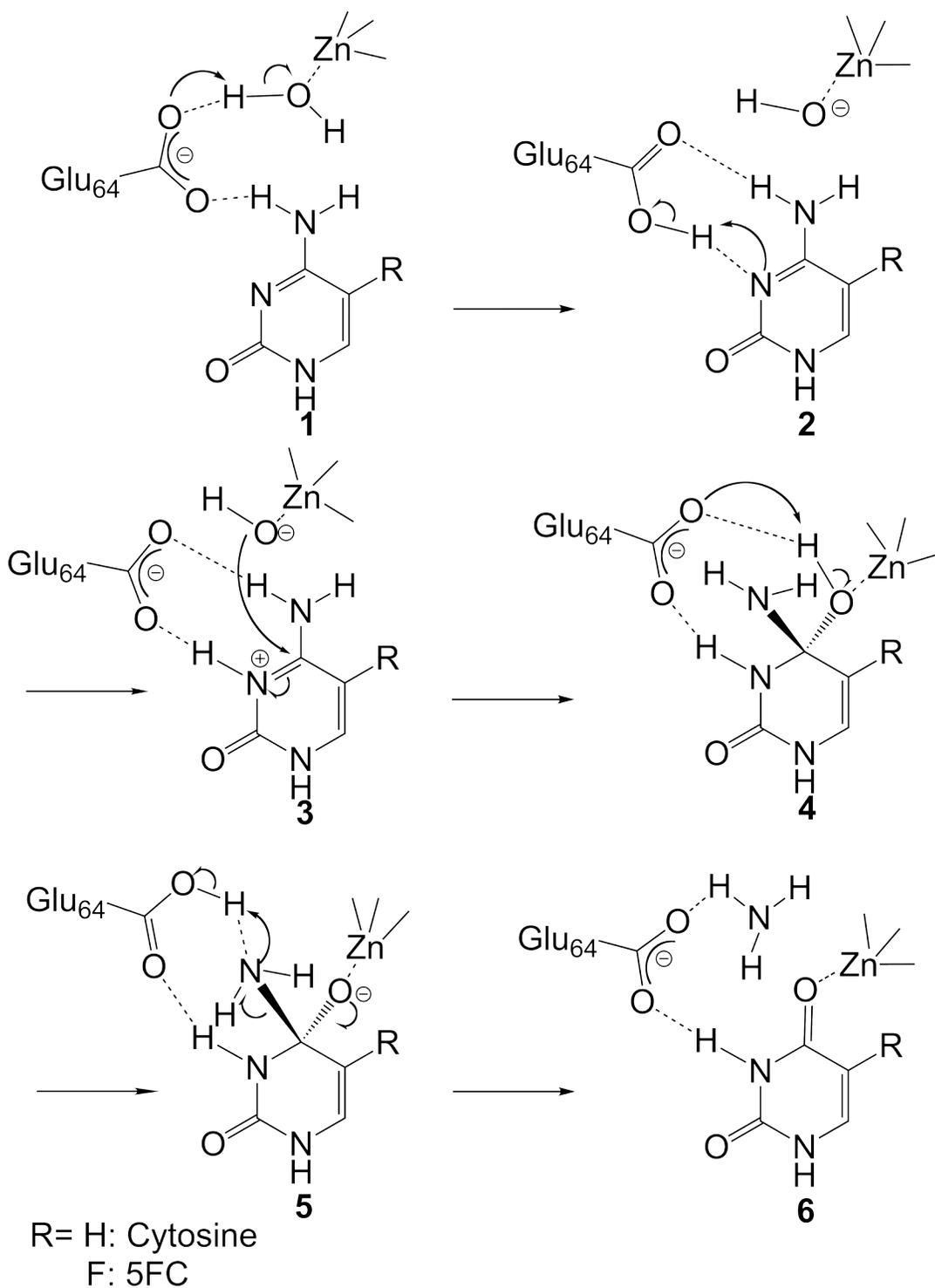


Figure 3-2. Proposed catalytic pathway for the conversion of cytosine (5FC) to Zn-coordinated uracil (5FC). (1) water/cytosine complex, (2) hydroxide/Glu64H cytosine complex, (3) hydroxide/Glu64 cytosine-H complex, (4) protonated tetrahedral intermediate complex, (5) unprotonated intermediate complex, (6) uracil/ammonia complex.

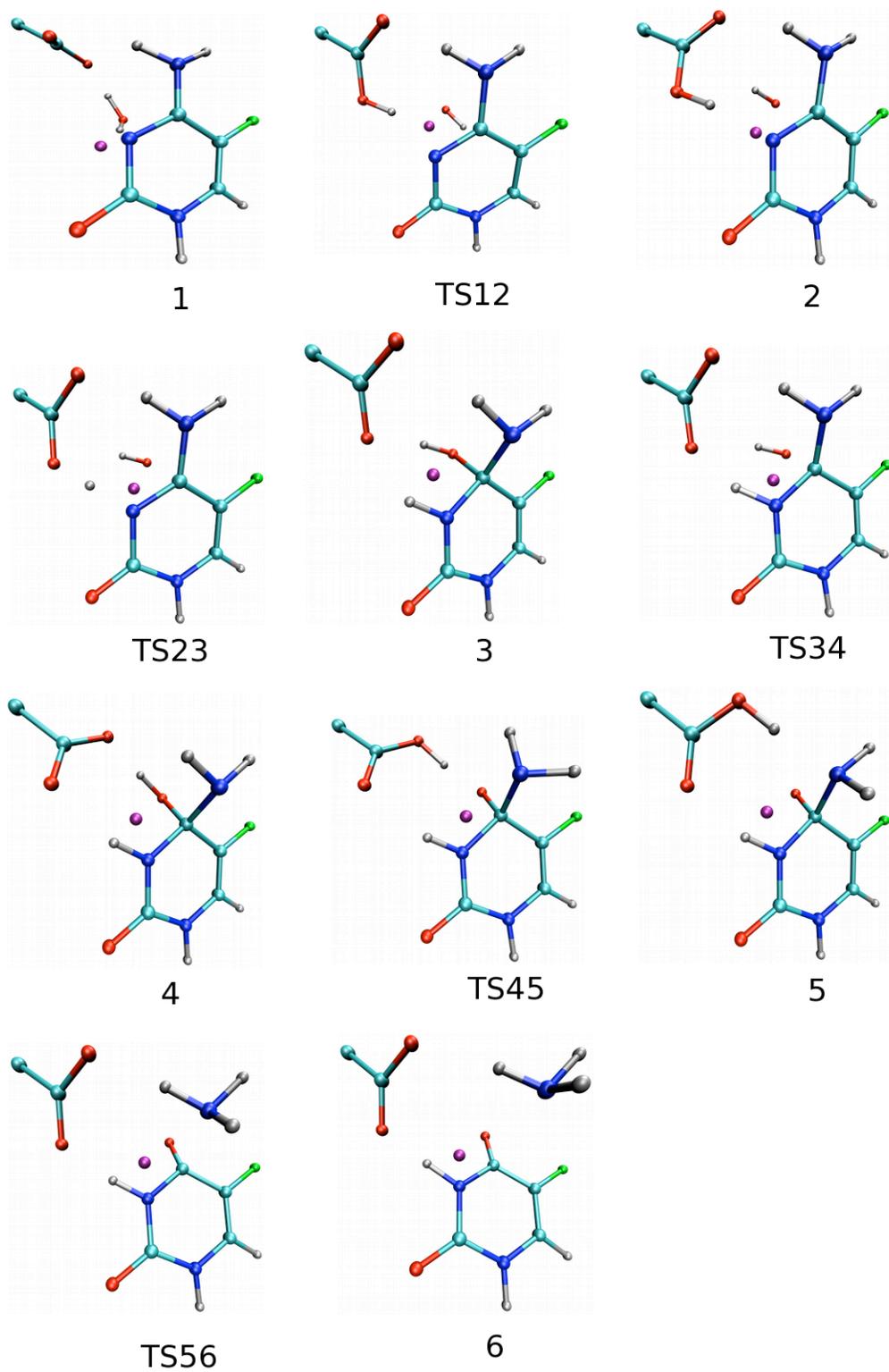


Figure 3-3. ONIOM optimized structures for complex 1-6 and the transition state structures during the deamination of 5FC.

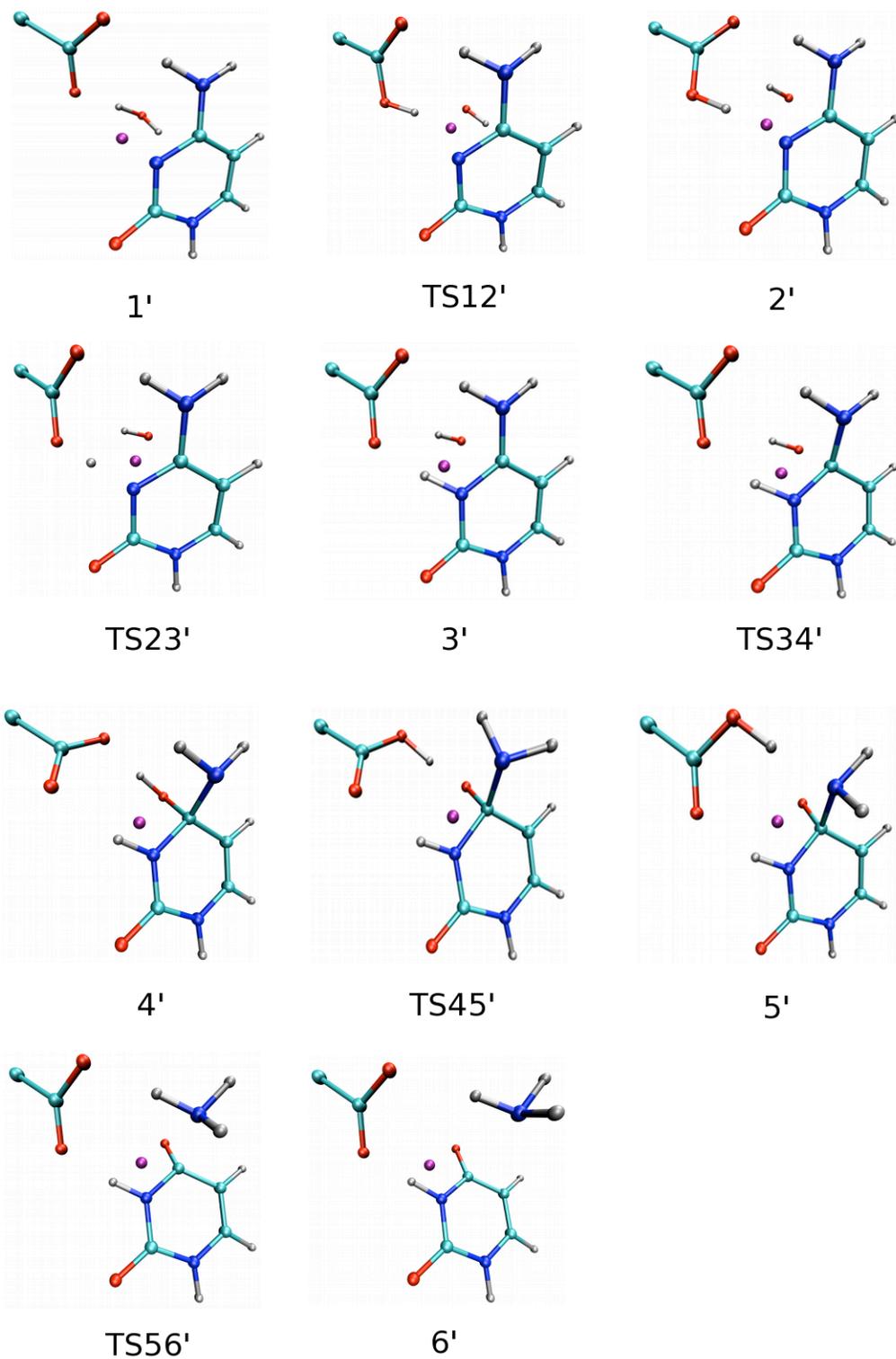


Figure 3-4. ONIOM optimized structures for complex 1'-6' and the transition state structures during the deamination of cytosine.

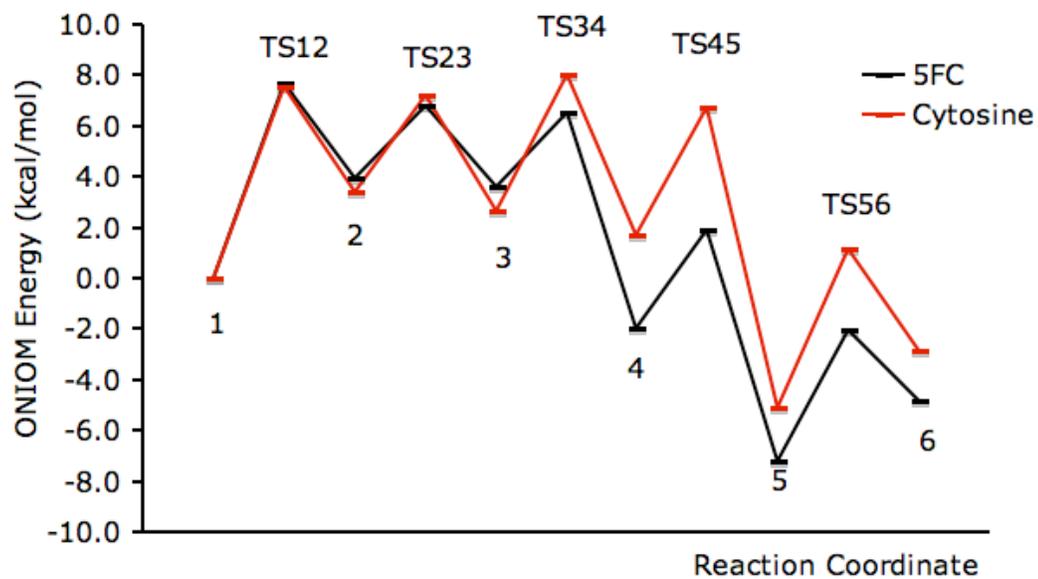


Figure 3-5. Energy profiles for the conversion of cytosine to the Zn-coordinated uracil and the conversion of 5FC to the Zn-coordinated 5FC.

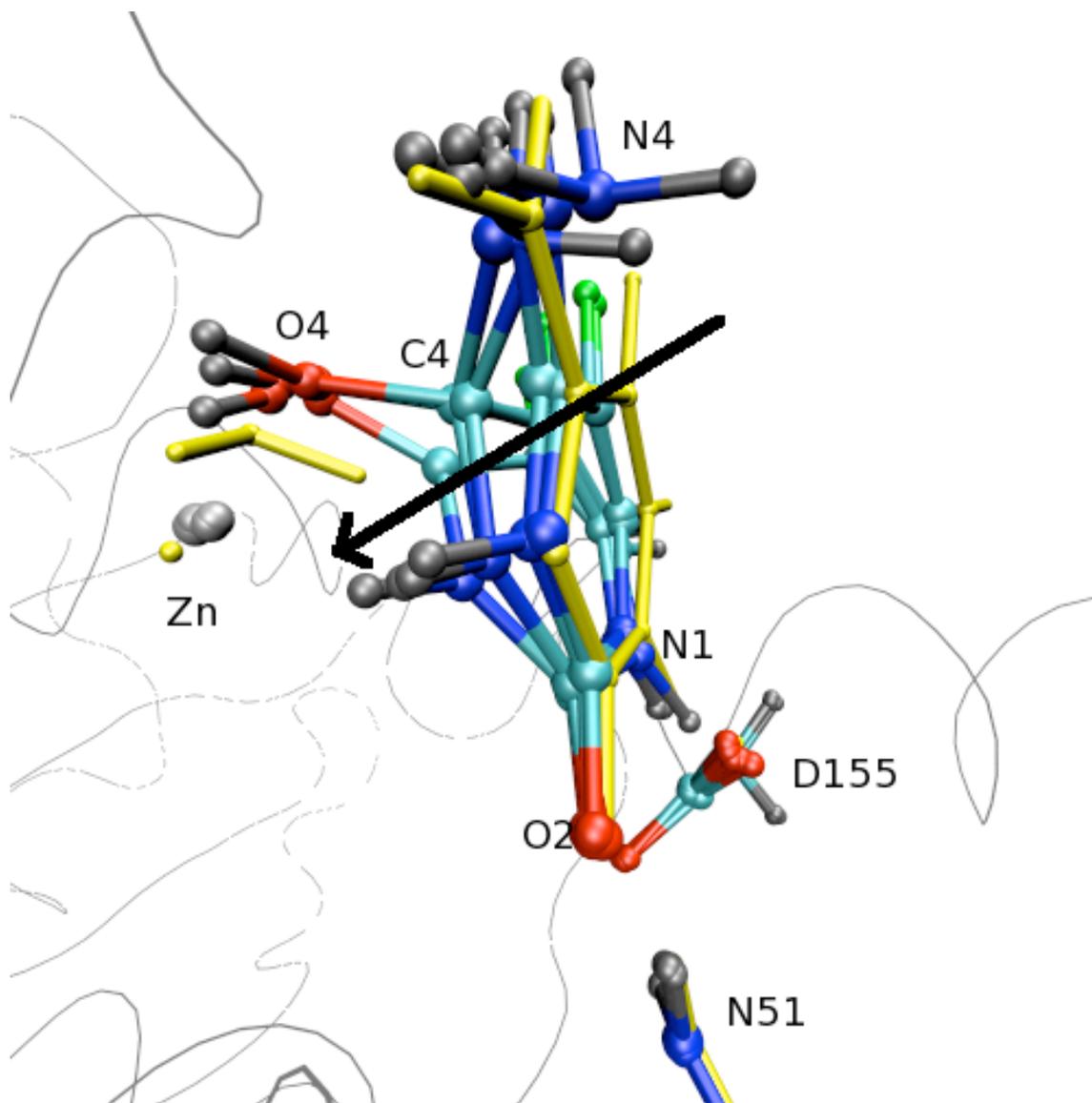


Figure 3-6. Rotation of the 5FC ring during the deamination of 5FC. The initial substrate is colored yellow. N1 and O2 are anchored by N51 and D155. The other end (C4) rotates towards the zinc-coordinated water and forms a covalent bond with O4. The arrow shows the direction how the ring rotates along the reaction coordinate.

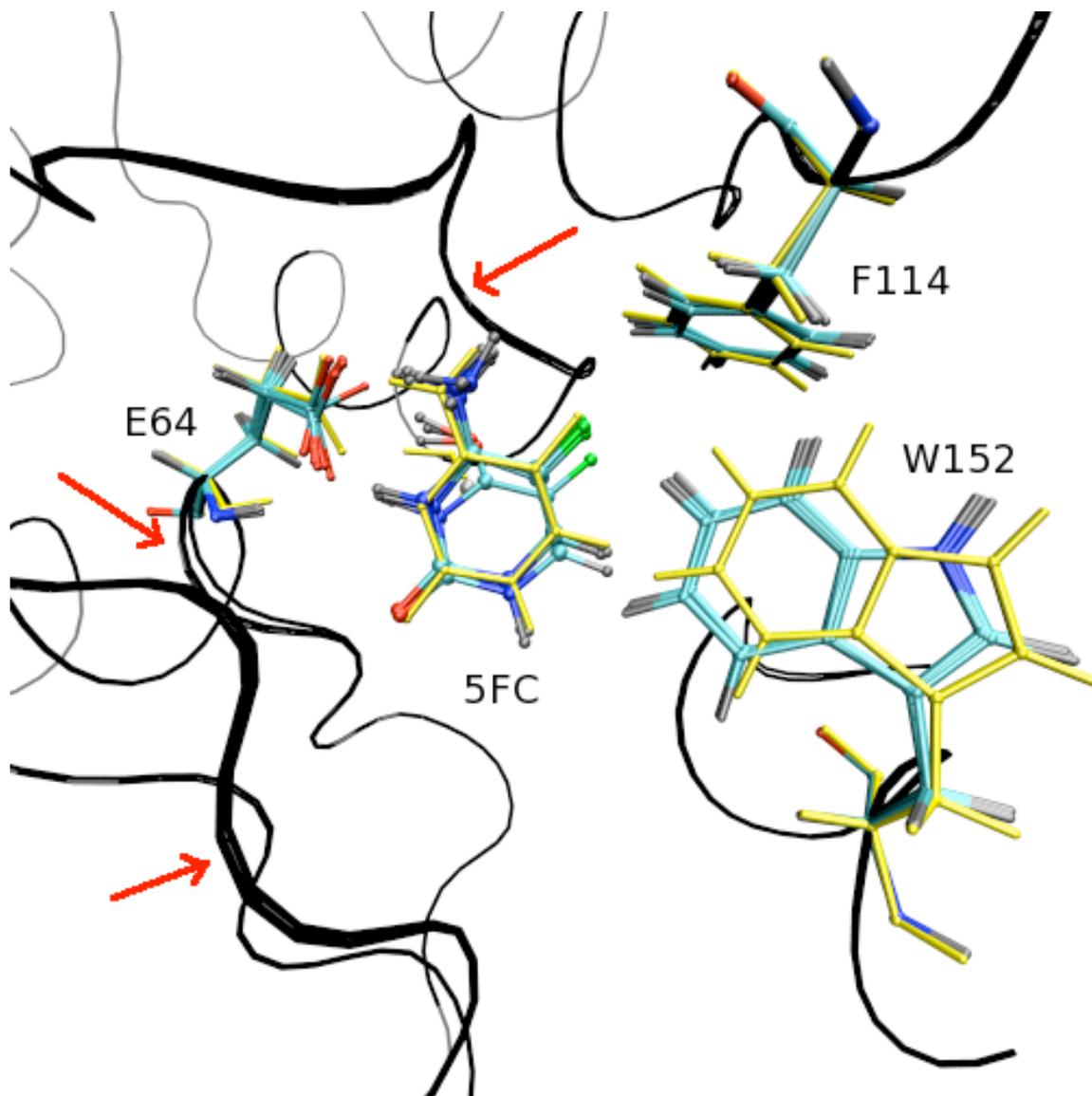
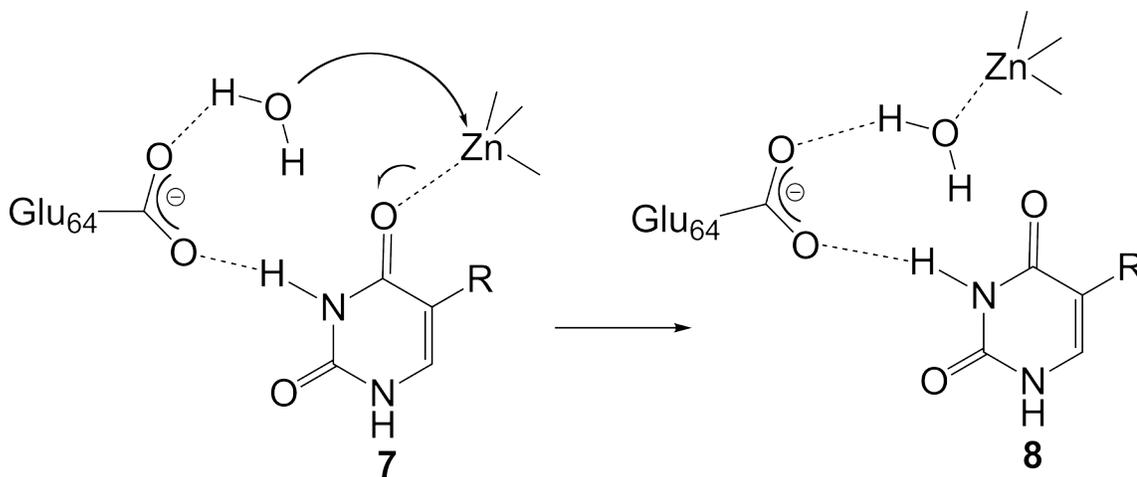


Figure 3-7. Motion of the active site residues during the deamination of 5FC. Several residues need to move around to accommodate the change of the intermediate (pointed by red arrow). These residues are within the most rigid regions. Motion of W152 and F114 is also observed. Residues in the initial complex are colored yellow.



R= H: Cytosine
 F: 5FC

Figure 3-8. Proposed catalytic pathways for the replacement of the Zn-coordinated uracil or 5FU with water.

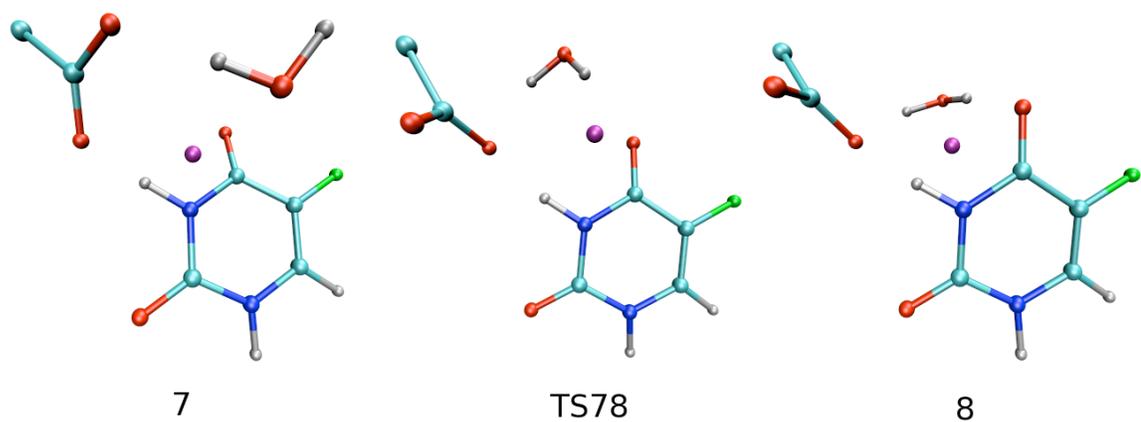


Figure 3-9. ONIOM optimized structures for Complex **7**, **8** and the transition state structure for the deamination of 5FC.

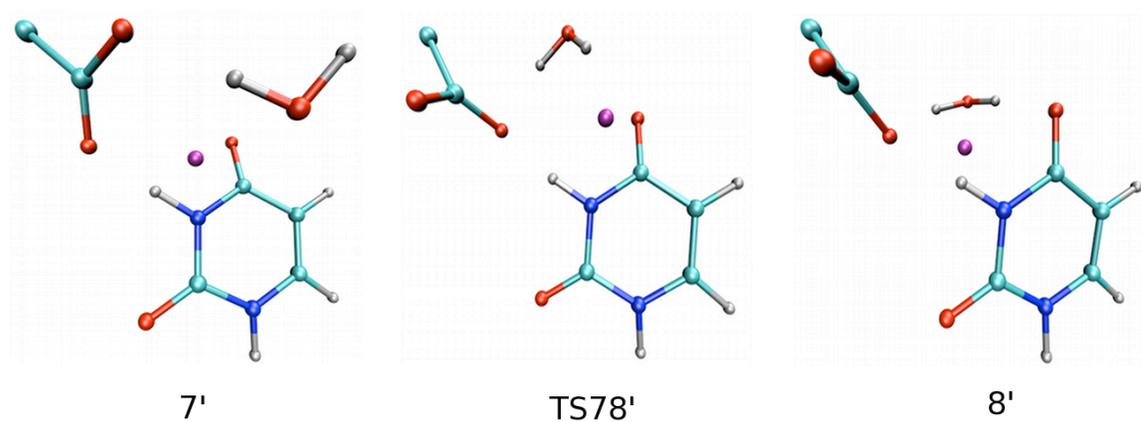


Figure 3-10. ONIOM optimized structures for Complex **7'**, **8'** and the transition state structure for the deamination of cytosine.

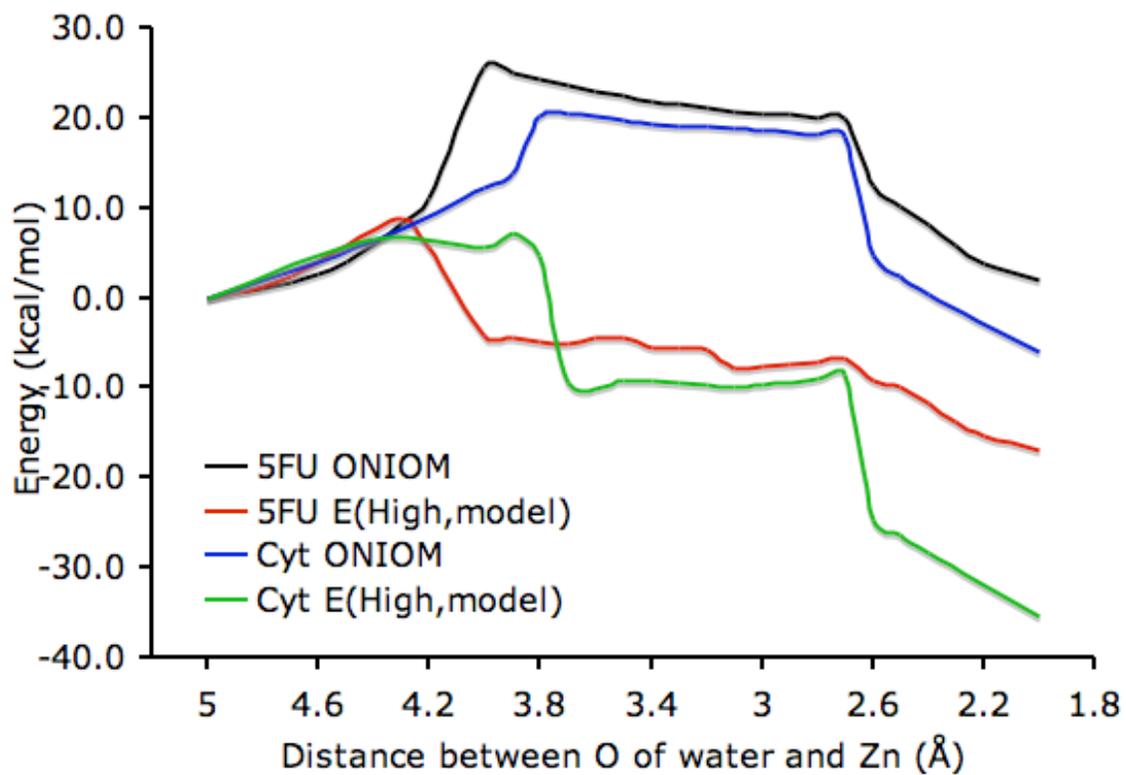


Figure 3-11. Energy profiles for replacement of the Zn-coordinated uracil or 5FU with water. ONIOM energies for 5FU and uracil are represented by black and blue curves, respectively. $E^{\text{model,QM}}$ energies for 5FU and uracil are represented by red and green curves, respectively.

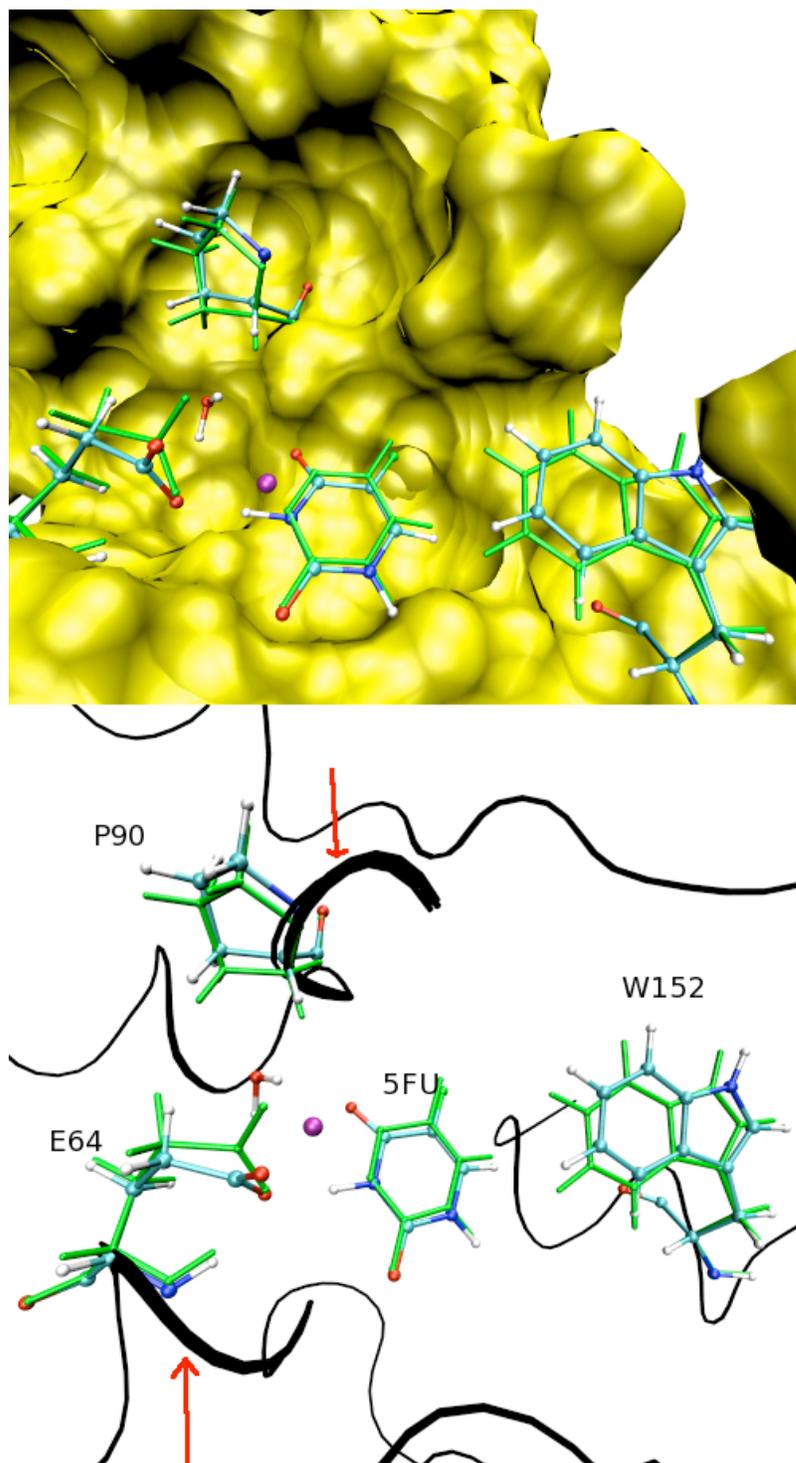


Figure 3-12. Conformational change caused by the access of the water towards the zinc atom. (top) the water molecule enters the space between Pro90 and Glu64 to access the zinc atom. (bottom) Motion of fragments to allow the access of the water.

REFERENCES

REFERENCES

1. Greco, O., and Dachs, G. U. (2001) Gene directed enzyme/prodrug therapy of cancer: Historical appraisal and future perspectives, *J Cell Physiol* 187, 22-36.
2. Aghi, M., Hochberg, F., and Breakefield, X. (2000) Prodrug activation enzymes in cancer gene therapy, *J Gene Med* 2, 148-164.
3. Ko, T., Lin, J., Hu, C., Hsu, Y., Wang, A., and Liaw, S. (2003) Crystal structure of yeast cytosine deaminase - Insights into enzyme mechanism and evolution, *J Biol Chem* 278, 19111-19117.
4. Yao, L., Li, Y., Wu, Y., Liu, A., and Yan, H. (2005) Product release is rate-limiting in the activation of the prodrug 5-fluorocytosine by yeast cytosine deaminase, *Biochemistry* 44, 5940-5947.
5. Stolworthy, T., Korkegian, A., Willmon, C., Ardiani, A., Cundiff, J., Stoddard, B., and Black, M. (2008) Yeast cytosine deaminase mutants with increased thermostability impart sensitivity to 5-fluorocytosine, *J Mol Biol* 377, 854-869.
6. Mahan, S., Ireton, G., Knoeber, C., Stoddard, B., and Black, M. (2004) Random mutagenesis and selection of *Escherichia coli* cytosine deaminase for cancer gene therapy, *Protein Eng Des Sel* 17, 625-633.
7. Mahan, S. D., Ireton, G. C., Stoddard, B. L., and Black, M. E. (2004) Alanine-scanning mutagenesis reveals a cytosine deaminase mutant with altered substrate preference, *Biochemistry* 43, 8957-8964.
8. Fuchita, M., Ardiani, A., Zhao, L., Serve, K., Stoddard, B., and Black, M. (2009) Bacterial cytosine deaminase mutants created by molecular engineering show improved 5-fluorocytosine-mediated cell killing *in vitro* and *in vivo*, *Cancer Res* 69, 4791-4799.
9. Sklenak, S., Yao, L., Cukier, R., and Yan, H. (2004) Catalytic mechanism of yeast cytosine deaminase: An ONIOM computational study, *J Am Chem Soc* 126, 14879-14889.
10. Yao, L., Yan, H., and Cukier, R. (2006) A combined ONIOM quantum chemical-molecular dynamics study of zinc-uracil bond breaking in yeast cytosine deaminase, *J Phys Chem B* 110, 26320-26326.
11. Vreven, T., Byun, K. S., Komáromi, I., Dapprich, S., Montgomery, Morokuma, K., and Frisch, M. J. (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM, *J Chem Theory Comput* 2, 815-826.
12. Humbel, S., Sieber, S., and Morokuma, K. (1996) The IMOMO method: Integration of different levels of molecular orbital approximations for geometry optimization of large systems: Test for *n*-butane conformation and S_N2 reaction: $\text{RCl}+\text{Cl}^-$, *J Chem*

Phys 105, 1959-1967.

13. Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., and Morokuma, K. (1996) ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and $\text{Pt}(\text{P}(t\text{-Bu})_3)_2 + \text{H}_2$ oxidative addition, *J Phys Chem* *100*, 19357-19363.
14. Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K., and Frisch, M. J. (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, *J Mol Struct: THEOCHEM* *461-462*, 1-21.
15. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2004) Gaussian 03, Revision D.01. Gaussian, Inc., Wallingford CT.
16. Ireton, G., Black, M., and Stoddard, B. (2003) The 1.14 angstrom crystal structure of yeast cytosine deaminase: Evolution of nucleotide salvage enzymes and implications for genetic chemotherapy, *Structure* *11*, 961-972.
17. Torrent, M., Vreven, T., Musaev, D. G., Morokuma, K., Farkas, O., and Schlegel, H. B. (2002) Effects of the protein environment on the structure and energetics of active sites of metalloenzymes. ONIOM study of methane monooxygenase and ribonucleotide reductase, *J Am Chem Soc* *124*, 192-193.
18. Ireton, G. C., McDermott, G., Black, M. E., and Stoddard, B. L. (2002) The structure of *Escherichia coli* cytosine deaminase, *J Mol Biol* *315*, 687-697.
19. Fersht, A. (1999) Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding. W. H. Freeman and Company, New York.
20. Yao, L., Sklenak, S., Yan, H., and Cukier, R. (2005) A molecular dynamics exploration of the catalytic mechanism of yeast cytosine deaminase, *J Phys Chem B* *109*, 7500-7510.

CHAPTER 4

QM/MM MD SIMULATIONS SHED NEW INSIGHTS ON THE CATALYTIC MECHANISM

INTRODUCTION

Yeast cytosine deaminase (yCD) is of significant biomedical interest because it catalyzes the conversion of the prodrug 5FC to 5FU which is widely used in the treatment of cancers (1, 2). Combination of yCD and 5FC is a promising strategy in gene directed enzyme prodrug therapy (GDEPT), through which the administered nontoxic 5FC is converted to the toxic and lethal 5FU in cancer cells.

Crystal structures of yCD determined at high resolution revealed that yCD is a homodimeric metalloenzyme (3, 4). The complete reaction path for the deamination of cytosine catalyzed by yCD has been proposed by the previous ONIOM calculations (5) and in Chapter 3. First, the substrate cytosine is activated by protonation at N3 with the help of the carboxyl group of Glu64. After transferring a proton to N3, the zinc-coordinated hydroxide is well positioned for nucleophilic attack on C4 to form a more stable tetrahedral intermediate, which is further stabilized by Glu64 and other surrounding residues via a set of hydrogen bonds and hydrophobic interactions. Glu64 is also involved in proton transfer during the C4-N4 bond cleavage. It abstracts the proton from O^{Zn} and then transfers it to N4 to form ammonia and the zinc-coordinated uracil. The N4-C4 bond cleavage is suggested to be the rate-limiting step up to the formation of

the zinc-coordinated uracil. The O^{Zn} -Zn bond breakage was studied by the same ONIOM calculation. The results show that the activation energy is quite high and the O^{Zn} -Zn bond breakage is suggested to be the overall rate-limiting step.

A revised ONIOM scheme was employed in Chapter 3 to compare the reaction mechanism of the deamination of 5FC by yCD with that of cytosine. In the new system, the whole enzyme was included to incorporate the protein environmental effects (6). Residues within a radius of 10 Å of the substrate were not restrained in order to accommodate the rearrangement of the active site along the reaction coordinate. Both molecular systems were well described before the formation of the zinc-coordinated uracil/5FU. However, the energy barrier of the O^{Zn} -Zn bond cleavage was still too high to be true. The $E^{\text{model,QM}}$ energy was low enough, suggesting that the protein environments did not help much in sampling the proper configurations along the proposed pathway.

Hence, configurational sampling needs to be incorporated while the chemical transformations are investigated to properly describe a process in which large conformational changes happen. A number of QM/MM protocols (7-14) have been developed either at the cost of accuracy to permit more configurational sampling or the opposite to precisely describe the chemistry.

A two-stepped ONIOM-MD calculation developed by Yao *et al* (15) was also applied on the O^{Zn} -Zn bond breakage. The results show that the activation energy for the bond breakage is only 2.9 kcal/mol, indicating that product release is not the rate-limiting step for the overall reaction. Fixing atoms in the outer layer in the first ONIOM calculation causes unrealistic effects, which were minimized by incorporating

configurational sampling in the MD scheme (15). Matsubara and co-workers developed a method (16) to directly couple the ONIOM method with the molecular dynamics (MD) method and applied this method to the process of zinc-uracil bond breakage (17). The results show that the O^{Zn} -Zn bond abruptly breaks at 11 ps during the simulation. The authors suggest that the thermal motion of Ile33, which sandwiches uracil with His62, perturbs the dynamics of uracil through steric contact and contributes to the bond breakage. A QM/MM MD method implemented in the CHARMM MD package (18) was performed on the yCD-DHP complex (19) to investigate the origin of the stabilization of the tight binding of DHP. The authors suggest that the inhibitor is converted to the alkoxide-like transition-state analogue and is stabilized.

In this chapter, the reaction mechanism of the deamination of cytosine by yCD was investigated by QM/MM MD simulations, using the fast semiempirical density-functional approach with self-consistent charge density functional tight-binding (SCC-DFTB) (8, 20) for QM and the Amber PARM99SB force field (21) for MM. The reaction mechanism was proposed starting from the substrate binding, the tetrahedral intermediate formation, and the C4-N4 bond cleavage to the O^{Zn} -Zn bond cleavage. Events were explored in details and analyzed, and the activation energy for each step was calculated and incorporated into the energy profile.

EXPERIMENTAL PROCEDURES

Molecular Dynamics Simulation. MD simulations were carried out on the free form yCD, the yCD-DHP complex and the enzyme complexes with the reaction species, including (1) water/cytosine complex, (2) hydroxide/Glu64H cytosine complex, (3) hydroxide/Glu64 cytosine-H complex, (4) protonated tetrahedral intermediate complex,

(5) unprotonated intermediate complex, (6) uracil/ammonia complex, (7) uracil/water complex and (8) water/uracil complex. The starting structures were obtained from the 1.14-Å resolution crystal structure of γ CD in complex with the inhibitor DHP (PDB ID: 1p6o) (4). The free form γ CD was built by replacing DHP with a zinc-coordinated water molecule. The water/cytosine complex (1) was built by replacing DHP with a zinc-coordinated water and a cytosine. The hydroxide/Glu64H cytosine complex (2) was built by replacing DHP with a zinc-coordinated hydroxide and a cytosine and changing Glu64 to the protonated form Glu64H. The hydroxide/Glu64 cytosine-H complex (3) was built by replacing DHP with a zinc-coordinated hydroxide and a protonated cytosine (cytosine-H). The protonated tetrahedral intermediate complex (4) was built by replacing DHP with the zinc-coordinated protonated tetrahedral intermediate. The unprotonated tetrahedral intermediate complex (5) was built by replacing DHP with the zinc-coordinated unprotonated tetrahedral intermediate. The uracil/ammonia complex (6) was built by replacing DHP with a zinc-coordinated uracil and an ammonia molecule. The uracil/water complex (7) was built by replacing DHP with a zinc-coordinated uracil and a water molecule. The water/uracil complex (8) was built by replacing DHP with a zinc-coordinated water and a uracil molecule. Ten molecular systems were constructed, in each of which, only the subunit A was modified. Subunit B remained the same as in the crystal structure.

All crystal water molecules were removed. All ionizable residues were assigned to their normal protonation state at pH 7.0 except for His50, which was set as doubly protonated because a carboxyl group presents at each side of the imidazole ring. The catalytic zinc ion is tetrahedrally coordinated with His62, Cys91, Cys94, and a water molecule. The sulfur atoms of Cys91 and Cys94 were set as deprotonated and His62 was

singly protonated at the epsilon nitrogen. Explicit bonds between the zinc atom and its coordination were used. The force constants for bonds and bond angles and the charges of Zn and its ligands have been described previously (22). The protein was placed in a periodic box of 12.5 Å TIP3P water molecules. Four Na⁺ ions were placed to counteract the negative charge of the molecular system. The parallel version Sander in Amber 10 (23) with the PARM99SB (21) force field was employed for the MD simulations.

The energy minimization of the molecular system is divided into three steps. First, the solvent was minimized with restraints on the protein for 15,000 steps, using the steepest descent method for the first 7,500 steps followed by the conjugate gradient method for another 7,500 steps. Second, the protein was minimized with restraints on the solvent for 5,000 steps, using the steepest descent method for the first 2,500 steps followed by the conjugate gradient method for another 2,500 steps. Third, the whole system was minimized without restraints, using the steepest descent method for the first 15,000 steps followed by the conjugate gradient method for another 15,000 steps. During the minimization, nonbonded interactions were calculated using the nonbonded pair list cutoff at 12.0 Å, and the nonbonded pair list was updated every 25 steps.

The molecular system was heated up gradually from 0 to 300 K at 50 K intervals for a total of 240,000 steps using a 1 fs time step. Temperature was regulated by Langevin dynamics (24) with $\gamma = 1.0 \text{ ps}^{-1}$. Another 40,000 steps were run at 300 K before the simulation was switched to an NPT (atom number (N), pressure (P) and temperature (T)) ensemble. During the heating process, nonbonded interactions were calculated using the nonbonded pair list cutoff at 12.0 Å, and the nonbonded pair list was updated every 25 steps.

Equilibration simulation was performed using 1 fs as the time step. The SHAKE algorithm (25) was turned on for all bonds involving hydrogen. The pressure was set at 1 atm and the temperature was controlled at 300 K by Langevin dynamics (24) with $\gamma = 1.0 \text{ ps}^{-1}$. Nonbonded interactions were calculated using the nonbonded pair list cutoff at 12.0 Å, and the nonbonded pair list was updated every 25 steps. Production simulations were performed under the same condition as the equilibration simulation. All of the MD results were analyzed with the Ptraj module of Amber 10 (23).

QM/MM MD Simulations. The self-consistent charge density functional tight-binding (SCC-DFTB) method is a semi-empirical method developed by Elstner *et al* (20, 26), which has been applied on a number of enzymatic systems (7, 8, 27-29). In this study, QM/MM MD simulations were performed for yCD using the SCC-DFTB method implemented in the amber molecular dynamics package (8). The molecular systems for the yCD-DHP complex, the yCD free form complex and the yCD water/cytosine complex (**1**) were built as described in the previous section. The active site of subunit A was modeled in the QM region (Figure 4-1), including the side chains of His62, Glu64, Cys91, and Cys94 and the catalytic zinc atom, and the ligands for the respective complexes. The rest of the protein was modeled in the MM region. SCC-DFTB and PARM99SB force field were used for the QM and MM regions, respectively. The SHAKE algorithm was turned on for all bonds except those in the QM region to allow proton transfer.

The molecular systems were minimized and heated up to 300 K using the same procedure as in the conventional MD simulation, except that the residues in the QM region were calculated by the SCC-DFTB method. The structures after being heated up to 300 K were used as the starting structures for the equilibration and production MD

simulations. As an alternative, representative structures from the conventional production MD simulations were tested as the starting structures. The results were similar.

During the equilibration and production MD simulations, temperature was regulated by Langevin dynamics with $\gamma = 1.0 \text{ ps}^{-1}$. Tight convergence on density matrix in the SCF was used. The electrostatic cutoff was set at 12.0 Å for QM/MM electrostatic interaction. Long-range QM-QM and QM-MM electrostatic interactions were calculated by Ewald sum (28, 30) and a QM compatible PME approach (7). The time step is 1 fs. MD simulations were carried out using the parallel version Sander module in Amber 10. All of the MD results were analyzed with the Ptraj module of Amber 10.

The free energy profile along a reaction coordinate was obtained from the potential of mean force (PMF) by adopting an umbrella sampling technique (29). Proper amount of windows were used to sample enough space along the reaction coordinate. The force constants were set in the range of 100 to 600 kcal/mol for simulations during a bond formation and the spring constants were set at $0.02 \text{ kcal/mol}\cdot\text{deg}^2$ for simulations during the rotation of the side chain of Glu64. The data from the simulations were analyzed by the weighted histogram analysis method (WHAM) (31, 32) to compute the PMF. Data from a 200 ps trajectory were collected after 200 ps of equilibration for each window.

RESULTS AND DISCUSSION

Two crystal structures of the yCD-DHP complex have been solved (3, 4) at high resolution. These structures not only serve as great starting conformations for the MD simulation, but also can be used to estimate how well the QM/MM method performs when compared with the MD simulations.

QM/MM MD Simulation on the yCD-DHP Complex. In the conventional MD simulation on yCD, the residues coordinated with Zn were constrained by explicit bonds as described elsewhere (22). The zinc coordination stayed stable and matched well with the geometry of the crystal structure. However, the chemical change cannot be captured due to the inherited nature of this methodology. A QM/MM MD simulation on the yCD-DHP complex was carried out instead. Since the residues (His62 N^{δ1}, Cys91 S^γ, Cys94 S^γ and DHP-O4) coordinated with zinc were included in the QM core, it is not necessary to constrain the bond distances by explicit forces. Moreover, chemical changes, such as proton transfer, bond formation and bond cleavage, can be studied directly.

A 1.3 ns simulation trajectory was collected and analyzed. Figure 4-2 shows the time evolution of the C α atom root mean squared deviation (RMSD) of the snapshots along the trajectory relative to the initial crystal structure. The average RMSDs for subunit A and subunit B are 0.54 and 0.58 Å, respectively, indicating that the structure is very rigid during the simulation course. The first 0.5 ns was regarded as the equilibrium stage and discarded and key internal coordinates are summarized in Table 4-1 based on the last 0.8 ns trajectory of the simulation. The structural similarity is shown in Figure 4-3 by superimposing the average structure of the snapshots in the MD trajectory with the crystal structure.

The results show the zinc coordination was described very well. The average bond distance between Zn and S^γ of Cys91 is 2.37 Å; the average bond distance between Zn and S^γ of Cys94 is 2.32 Å; the average bond distance between Zn and N^{δ1} of His62 is 2.00 Å; the average bond distance between Zn and O4 of DHP is 2.12 Å. The bond

lengths are comparable with those in the crystal structure and previous ONIOM calculations in Chapter 3 (Table 4-1).

The majority of the inhibitor in solution is in the unhydrated form 2Py. Only less than 0.05% of the inhibitor is in the covalent hydrate form DHP (33). The crystal structure suggests that the inhibitor in the active site of γ CD is DHP. The simulation captures the tetrahedral feature of the inhibitor in which the C4 atom shows obvious sp^3 hybridization characteristics. The dihedral angles $\angle O4-C4-N3-C2$ and $\angle O4-C4-C5-C6$ are 120° and -117° , respectively. The bond distances are also in close agreement with those in the crystal structure (Table 4-1). The O4-C4 bond is a single bond with a bond distance at 1.47 Å and the N3-C4 bond is a single bond too, with a bond distance at 1.45 Å. When the inhibitor is in its unhydrated form, the N3-C4 bond is clearly a double bond with a bond distance at 1.32 Å and C4 adopts a sp^2 hybridization.

DHP is stabilized by a number of hydrogen bonds. N1 is hydrogen bonded with the carboxyl group of Asp155 with a distance of 2.80 Å. O2 is hydrogen bonded with the side chain amide group of Asn51 and the backbone amide group of Gly63 with distances of 2.91 and 2.86 Å, respectively. DHP forms two hydrogen bonds with the carboxyl group of Glu64, through O4 and N3, with a distance of 2.48 and 2.88 Å, respectively. O4 is also hydrogen bonded with the backbone amide of Cys91 with distances of 3.14 Å. Other than the hydrogen bonds, hydrophobic interactions contribute to the stabilization too. DHP is surrounded with a cluster of nonpolar residues, such as Ile33, Leu88, Phe114, Trp152, and Ile156. The imidazole ring of His62 interacts with the pyrimidine ring of DHP through a π -stacking interaction. The docking surface is well defined for DHP and is rigidified by hydrogen bonds and hydrophobic packing. For example,

hydrogen bonds are formed between the amide group of Asn51 the carboxyl group of Asp155 and between N^{ε2} of His62 and the backbone oxygen of Asp155.

Characteristic of the O4...O^{ε1} Hydrogen Bond. Hydrogen bonds between O4 of DHP and O^{ε1} of Glu64, and between N3 of DHP and O^{ε2} of Glu64 may be particularly important for the stabilization of DHP because these two hydrogen bonds are not present in the yCD-2Py complex. The former is a hydrogen bond with a very short bond distance at 2.48 Å. The hydrogen does not exclusively form a covalent bond with either heteroatom. It is located in the middle of the two heteroatoms, with an average distance between HO4 and O4 at 1.31 Å and between HO4 and O^{ε1} at 1.17 Å in the 0.8 ns trajectory (Table 4-1 and Figure 4-4). The hydrogen bond may have only one potential energy well or two wells separated by a low energy barrier.

A 2D umbrella sampling technique was employed to illustrate the potential of mean force for the proton transfer between the two heavy atoms. The reaction coordinate is based on the bond distance between HO4 and O4 of DHP and between HO4 and O^{ε1} of Glu64 (Figure 4-5). The results show that the energy profile indeed has a very flat well where the hydrogen atom tends to sit in the middle between the two heteroatoms. There is no obvious energy barrier. This calculation suggests that the hydrogen bond is very likely a low barrier hydrogen bond (LBHB) (34).

In comparison, a simulation on the yCD-2Py complex was performed. O^{ε1} of Glu64 forms a hydrogen bond with the zinc-coordinated water with a distance of 2.60 Å. The distance is rather short too when compared with the average O...O type hydrogen

bond. However, the hydrogen in the hydrogen bond (H^{WAT}) is covalently bonded with the oxygen of the zinc-coordinated water (O^{WAT}) and the distance between H^{WAT} and $\text{O}^{\varepsilon 2}$ of Glu64 is 1.8 Å. The energy profile for the proton transfer from the zinc-coordinated water to $\text{O}^{\varepsilon 1}$ of Glu64 was obtained using the same protocol (Figure 4-5, bottom). The reaction coordinate is set based on the distance between H^{WAT} and O^{WAT} and between H^{WAT} and $\text{O}^{\varepsilon 1}$ of Glu64. The energy profile shows that the system is much more stable when H^{WAT} stays with O^{WAT} . The proton transfer is not a preferred process probably due to the $\text{p}K_{\text{a}}$ mismatch between $\text{O}^{\varepsilon 1}$ of Glu64 and the zinc-coordinated water. The $\text{p}K_{\text{a}}$ of the zinc-coordinated water is 8-9 (27), while the $\text{p}K_{\text{a}}$ of Glu64 is normally between 2 and 5.5 in protein (35). Preferably, the hydrogen will stay with the water instead of transferring to Glu64. However, the $\text{p}K_{\text{a}}$ of the water would probably be reduced to somewhere close to that of Glu64 when it is incorporated into the inhibitor. The strong hydrogen bond in the yCD DHP complex was inferred by a ^1H jump return NMR experiment in which a proton peak at 16.2 ppm was observed on the NMR spectrum (unpublished data).

In Chapter 3, ONIOM calculations also suggested that this $\text{O4}\dots\text{O}^{\varepsilon 1}$ hydrogen bond is an LBHB. In the ONIOM calculation, the most essential elements were treated at the semiempirical DFT method B3LYP with the 6-31G* basis set. B3LYP is a higher

energy level than DFTB used in this study. The ability to capture this important feature also indicates the validation of this simulation protocol.

The same $O4 \cdots O^{\epsilon 1}$ hydrogen bond in the γ CD-DHP complex was investigated recently by QM/MM MD simulations (19), using SCC-DFTB for QM and the CHARMM force field for MM. The authors suggest the proton is completely transferred to $O^{\epsilon 1}$ and this proton transfer may play an important role in stabilizing the inhibitor. In the calculation of the simple model in that study (19), the hydrogen is covalently bonded to $O^{\epsilon 2}$, but the bond length is 1.0 Å and the distance between $O^{\epsilon 2}$ and O4 is 2.7 Å, characteristic of a conventional H-bond. It is noted, however, that the zinc coordination was not reproduced well because the distances of three out of the four ligands to the zinc are significantly different from those measured by X-ray diffraction.

There have been debates (36-40) about whether the LBHB exists or not, and if it does, what is its function. By using high-resolution neutron crystallography, a LBHB was indeed discovered by Yamaguchi et al (41) in photoactive yellow protein (PYP) between the chromophore of PYP, *p*-coumaric acid (pCA), and the carboxyl oxygen of Glu46. The bond distance is 2.56 Å. The proton sits in the middle between the two heavy atoms, with distances of 1.37 Å to pCA and 1.21 Å to Glu46. It was suggested to stabilize the buried charge in the hydrophobic environments.

For the $O4 \cdots O^{\epsilon 1}$ hydrogen bond in the γ CD-DHP complex, the LBHB may play an important role in the stabilization of the inhibitor. Further experiments are needed to investigate the nature of this hydrogen bond and its function in catalysis.

Substrate Activation. To investigate the deamination of cytosine by yCD, a series of QM/MM MD simulations were performed starting from the yCD water/cytosine complex (**1**). **1** was constructed by replacing DHP with a cytosine and a zinc-coordinated water molecule in the active site of yCD (Figure 4-6 and 4-7). Cytosine is placed in the same orientation as DHP.

A 0.8 ns QM/MM MD simulation trajectory after equilibration (0.5 ns) was collected and analyzed. Figure 4-8 shows the time evolution of the C α atom RMSDs of the snapshots along the trajectory relative to the initial crystal structure. The average RMSDs for subunit A and subunit B are 0.65 and 0.54 Å, respectively, indicating that the structure is quite rigid during the simulation course.

Key internal coordinates are summarized in Table 4-1. Similarly, cytosine is stabilized by a set of hydrogen bonds and hydrophobic interactions (Table 4-1). Two additional hydrogen bonds were formed because of the presence of the amide group (N4). One is between N4 and the main chain carbonyl of Ser89 and the other is between N4 and O $^{\epsilon 2}$ of Glu64. O $^{\epsilon 1}$ of Glu64 forms a hydrogen bond with the zinc-coordinated water, with an average distance of 2.58 ± 0.10 Å.

Glu64 is believed to serve as a general base/acid during multiple steps. The side chain rotation is critical to interact with different hydrogen donor/acceptor species during different stages. The information on side chain rotation is extracted from the MD trajectory (Figure 4-9). Moreover, we performed a simulation on the free form yCD. The dihedral angle of the side chain (defined as $(\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\epsilon 2})$) was analyzed and compared with that in the yCD cytosine complex.

For the free form yCD, the 180° rotation of the side chain of Glu64 was observed in the 0.8 ns trajectory (Figure 4-9, C). The other rotation mode is between two states, between which the difference of the side chain dihedral angle ($\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\epsilon 2}$) is 50° (Figure 4-9, C). The hydrogen bond between the zinc-coordinated water and the carboxyl group of Glu64 can switch rapidly from one carboxyl oxygen of Glu64 to the other (Figure 4-9, A). Rotation of Glu64 side chain was investigated by the umbrella sampling technique. The energy profile is plotted along the reaction coordinate based on the dihedral angle ($\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\epsilon 2}$) (Figure 4-9, D). Two shallow wells were found at -20 and -70°, indicating the rotation is indeed facile. The 180° rotation costs about 2.0 kcal/mol, again suggesting that the rotation of the side chain of Glu64 is a common event for the free form of yCD on the MD time scale.

After the binding of cytosine, the side chain of Glu64 adopts a slightly different configuration, with the dihedral angle ($\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\epsilon 2}$) oscillating around -120° (Figure 4-9). Furthermore, the presence of cytosine obscures the 180° rotation of the side chain of Glu64, suggested by the PMF, in which the activation energy is increased to 4.5 kcal/mol. Two potential energy wells were found, corresponding to two stable states with ($\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\epsilon 2}$) at -100° and -150°, respectively. However, the rotation did not break the hydrogen bond between the carboxyl oxygen and the zinc-coordinated water (Figure 4-9, B).

The other carboxyl oxygen of Glu64 forms a hydrogen bond with N4 of cytosine and the hydrogen bond pulls cytosine closer to the zinc-coordinated water. The distance

between the water oxygen (O^{Zn}) and C4 of cytosine is 2.90 Å. In the complex with 2Py, the distance is 3.07 Å. However, the water is not likely to attack C4 directly. The water needs to be activated by losing a proton to the carboxyl group of Glu64, which, in turn, activates cytosine by transferring a proton to cytosine at N3.

This process was investigated by a 2D umbrella sampling calculation along two reaction coordinates. One is the distance between HO1 (hydrogen of the zinc-coordinate water) and N3 and the other is the distance between HW and $O^{\epsilon 2}$ of Glu64 (Figure 4-10). First, the proton (HO1) transfers to $O^{\epsilon 2}$ to form the yCD hydroxide/Glu64H cytosine complex (**2**). But there is no noticeable energy barrier. **2** is ~ 5 kcal/mol less stable than complex **1**. It suggests that although the water is coordinated with the zinc atom, its pK_a is still not low enough to match that of Glu64. The energy surface around **2** is very flat. When the hydrogen bond between $O^{\epsilon 2}$ and O^{Zn} breaks, a new hydrogen bond is formed between $O^{\epsilon 2}$ and N3, which compensates for the energy loss of the former hydrogen bond. **2** in this form is ~ 8 kcal/mol less stable than **1**. The system is stabilized after the proton (HO1) transfers to N3 to form the complex **3**. The transition state (TS2) is located when the proton is in the middle of $O^{\epsilon 2}$ and N3. The activation energy for the overall reaction from **1** to **3** is about 14.9 kcal/mol. **3** is less stable than **1** by 10.9 kcal/mol. In **3**, the energy level is elevated because two charged groups are generated, one is the zinc-coordinated hydroxide and the other is the positive charged cytosine (Cytosine-H). However, this special configuration and charge distribution facilitate the nucleophilic attack in the next step. The addition of the proton atom on N3 draws electrons from C4.

The distance between O^{Zn} and C4 decreases to 2.10 Å, from 2.45 Å in **2** and from 2.90 Å in **1**. With the shortened distance between O^{Zn} and C4, O^{Zn} is well positioned to attack C4.

Nucleophilic Attack. The zinc-coordinated hydroxide is well positioned for the nucleophilic attack on C4. The energy profile was obtained by umbrella sampling along the reaction coordinate by shortening the distance between O^{Zn} and C4 (Figure 4-11). The transition state was found when the O^{Zn} -C4 distance is around 1.82 Å. The energy barrier is only about 1.4 kcal/mol and the tetrahedral intermediate (**4**) is slightly more stable than the hydroxide/Glu64 Cytosine-H complex (**3**) by 1.0 kcal/mol.

With the decrease of the distance between O^{Zn} and C4, the side chain of Glu64 ($\angle C^\beta-C^\gamma-C^\delta-O^{\epsilon 2}$) rotates from 128° to 104°, leading to the formation of a hydrogen bond between $O^{\epsilon 1}$ of Glu64 and O^{Zn} . At the same time, the hydrogen bond between $O^{\epsilon 1}$ of Glu64 and N4 of cytosine is broken. The newly formed hydrogen bond between $O^{\epsilon 1}$ and O^{Zn} is quite strong, with a distance of 2.50 Å. It is comparable with that found in the yCD-DHP complex, which is believed to be important for the stabilization of the inhibitor (also see Chapter 3). Proton transfer within this hydrogen bond was investigated by umbrella sampling. The energy profile shows that the energy barrier is negligible.

The average structure of the snapshots of **4** in the MD simulation trajectory is superimposed onto that of the yCD-DHP complex, shown in Figure 4-12, from which, we can see that the structures align very well. All hydrogen bonds between DHP and yCD

are preserved for the intermediate and yCD. This may be the origin where the tight-binding of DHP comes from.

C4-N4 Bond Cleavage. From **3** to **4**, the hybridization state of C4 changes from *sp*² to *sp*³. The bond distance between N4 and C4 is increased to 1.44 Å, from 1.37 Å in **3**. **4** is stabilized by a network of hydrogen bonds and hydrophobic interactions. To investigate how the C4-N4 bond is cleaved, first, a 1D umbrella sampling calculation was carried out along the reaction coordinate based on the distance between C4 and N4. The results show that with the increase of the C4-N4 distance, the hydrogen of the hydroxyl group (HO4) transfers to O^{ε1} of Glu64. The transition state was found with a distance between C4 and N4 at 1.91 Å. The energy barrier is 16.9 kcal/mol. This is quite consistent with the transition state obtained from the KIE studies. However, whether the proton transfer to N4 and the breakage of C4-N4 bond occur in a concerted or stepwise manner is not clear from this energy profile.

Hence, a 2D umbrella sampling calculation was carried out along two reaction coordinates. One is the distance between HO4 and N4 and the other is the distance between C4 and N4. The energy profile is plotted in Figure 4-13. HO4 is first transferred to O^{ε1} of Glu64 by crossing over a low energy barrier. We have already seen this kind of hydrogen bond in the yCD-DHP complex. This strong hydrogen bond may be important for the stabilization of the tetrahedral intermediate and hence increase the catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$). With the increase of the distance of the C4-N4 bond, HO4 prefers to stay with O^{ε1}, instead of O^{Zn}, and forms a hydrogen bond with N4. It turns out that the proton needs to transfer to N4 before the C4-N4 bond completely breaks. The transition

state is found for the proton transfer when C4-N4 bond distance is increased to 1.65 Å and the energy barrier is 19.7 kcal/mol. It is also the overall activation energy for the C4-N4 bond cleavage. After the proton transfers to N4, the C4-N4 bond keeps increasing and the ammonia is formed. The remaining part of the intermediate forms the other product uracil, which is still covalently bonded with the zinc atom.

A surprising feature is that H3 is transferred to O^{ε2} of Glu64 when the C4-N4 bond distance is increased to 1.95 Å. The protonated side chain of Glu64 forms a hydrogen bond with O^{Zn} as a proton donor. The proton transfer was investigated by umbrella sampling calculations along the reaction coordinate, in which the N3-H3 bond distance is gradually increased. Ammonia is believed to diffuse out of the active site relatively quickly. Instead, a water molecule enters the active site and binds at the ammonia binding site. Hence, a new molecular system is set up, in which the ammonia is replaced by a water molecule, to investigate the proton transfer from N3 to O^{ε2} of Glu64. The results show the activation energy is quite low, with a value of 2.4 kcal/mol (Figure 4-14). Furthermore, the deprotonated uracil with its cognate enzyme is ~ 5.0 kcal/mol more stable than the protonated uracil complex.

The p*K*_a of N3 of uracil in solution is 9.5 (42). Binding with Zn at the O4 position can lower the p*K*_a of N3 to match that of the carboxyl group of Glu64. Hence the proton can transfer to O^{ε2}. Once the proton is transferred, the hydrogen bond between Glu64 and N3 is broken, which may facilitate the release of uracil.

O^{Zn}-Zn Bond Cleavage. Uracil is covalently bonded with the zinc atom after the C4-N4 bond cleavage. It cannot diffuse out unless the O^{Zn}-Zn bond is cleaved. The O^{Zn}-Zn bond cleavage has been proposed by previous ONIOM calculations. A two-stepped ONIOM-MD calculation developed by Yao *et al* (15) was also applied on the O^{Zn}-Zn bond breakage. The results show that the activation energy for the bond breakage is only 2.9 kcal/mol, indicating product release is not the rate-limiting step for the overall reaction. Fixing atoms in the outer layer in the first ONIOM calculation causes unrealistic effects, which were minimized by incorporating configurational sampling in the MD scheme (15). Matsubara and co-workers developed a method (16) to directly couple the ONIOM method with the MD method and applied this method on the process of the zinc-uracil bond breakage (17). The results show that the O^{Zn}-Zn bond abruptly breaks at 11 ps during the simulation. The authors suggest that the thermal motion of Ile33, which sandwiches uracil with His62, perturbs the dynamics of uracil through steric contact and contributes to the bond breakage.

In this study, a pathway was proposed, in which a solvent water attacks the zinc atom and releases uracil (Figure 4-15). It has been attempted by ONIOM calculations in Chapter 3, but the right path was not sampled. A series of MD simulations were carried out to push the water towards the zinc atom and replace uracil and the PMF was obtained (Figure 4-16). The transition state was located when the distance between water and zinc was at 2.9 Å. The energy barrier is 8.1 kcal/mol. The O^{Zn} atom is expected to be quickly released from Zn afterwards. The zinc atom forms a new coordination with the water. The uracil would now be ready to diffuse out of the active site.

Proton transfer from N3 to $O^{\epsilon 1}$ of Glu64 plays an important role in the $O^{\text{Zn}}\text{-Zn}$ bond cleavage. The hydrogen bond between N3 and $O^{\epsilon 1}$ is disrupted because the side chain of Glu64 rotates away to form a new hydrogen bond with O^{Zn} or the water molecule. Hence the zinc-coordinated uracil is not as rigidified as the intermediates occurring in the prior steps. When the water moves close to Zn, the pyrimidine ring can shift away a bit to decrease the steric hindrance. Once water moves into the cavity close to Zn, the $O^{\epsilon 1}\text{-H}$ group swings back to form a hydrogen bond with N3 and eventually transfers the proton back to N3. This also facilitates the release of O^{Zn} from Zn.

The overall energy profile for the enzyme reaction is plotted in Figure 4-17. Even with the help of Glu64, the proton transfer from the zinc-coordinated water to N3 of cytosine needs a large activation energy to cross over the transition state. Once the complex **3** is formed, it can be easily converted to the tetrahedral intermediate (**4**). **4** is stabilized by a hydrogen bond network and hydrophobic interactions. The hydrogen bond between O4 and $O^{\epsilon 1}$ of Glu64 may be particularly important. For the C4-N4 bond cleavage, the activation energy is very high. It is the overall rate-limiting step. The energy barrier is 19.7 kcal/mol. The C4-O4 bond is cleaved during the water attacks Zn. The energy barrier is only 8.1 kcal/mol, which is much lower than that in the C4-N4 bond cleavage step.

CONCLUDING REMARKS

In this study, a series of QM/MM MD simulations were performed to explore the

catalytic mechanism of the deamination of cytosine by yCD. Different from the ONIOM calculations in Chapter 3, dynamic effects were taken into consideration in the MD simulation. The essential elements, composed of the side chains of Glu64, His62, Cys91 and Cys94, the catalytic zinc and the ligands, were included in the QM part and simulated by the fast semiempirical density-functional approach SCC-DFTB. The remainder of the protein and the solvent were simulated by the PARM99SB force field. The method was applied on the yCD-DHP complex to validate the methodology on the yCD system. The key coordinates of the resulting structures in the equilibrium ensemble are in close agreement with those of the crystal structure, especially those for the catalytic zinc apparatus. The results also suggest that the hydrogen bond between the carboxyl group of Glu64 and O4 of DHP is possibly a low barrier hydrogen bond, which may play an important role in the stabilization of the tetrahedral intermediate.

The reaction pathway of the deamination of cytosine was proposed. First, the substrate cytosine is activated by abstracting one proton from the zinc-coordinated water via $O^{\epsilon 1}$ of Glu64. The energy is elevated due to charge redistribution and geometry change. However, the zinc-coordinated hydroxide is well positioned for the nucleophilic attack on C4. Second, the tetrahedral intermediate is formed after the nucleophilic attack. The energy barrier is quite low, with a value of 1.4 kcal/mol. The intermediate is stabilized by a hydrogen bond network and hydrophobic interactions. The hydrogen bond between $O^{\epsilon 1}$ of Glu64 and O^{Zn} is very strong, with a distance of 2.51 Å, which is quite similar to the hydrogen bond in the yCD-DHP complex. This hydrogen bond is likely to play an important role in stabilizing the intermediate. Third, the C4-N4 bond is cleaved. Several proton transfer events were observed during this process. HO4 first transfers

from O^{Zn} to $O^{\epsilon 1}$ of Glu64. With the increase of the distance between C4 and N4, HO4 transfers to N4 to form the product ammonia. H3 transfers to $O^{\epsilon 2}$ of Glu64 when the distance C4-N4 increases to 1.95 Å, probably because the pK_a of N3 is lowered due to the interaction between Zn and O^{Zn} . The transition state for the C4-N4 step is found when the C4-N4 distance is about 1.65 Å, where HO4 already transferred to N4 and H3 still bonds with N3. It is also the overall transition state, which is believed to be the rate-limiting step with an energy barrier of 19.7 kcal/mol. The product ammonia can easily diffuse into the solution. The other product uracil is covalently bonded with the zinc atom. We propose that a solvent water molecule enters the active site and replaces uracil in coordination with Zn. The energy barrier is about 8.1 kcal/mol, which is lower than that for the C4-N4 bond cleavage step.

TABLES AND FIGURES

Table 4-1. Comparison of changes of key coordinates during the conversion of cytosine to uracil with crystal structures ^a.

Internal Coordinate	X-ray Structure (1uaq) ^b	X-ray Structure (1p6o)	yCD-DHP	Complex 1	Complex 4
His62-N ^{δ1} ...Zn	2.07, 2.06	2.01, 2.00	2.00	2.00	2.00
Cys91-S ^γ ...Zn	2.35, 2.31	2.29, 2.30	2.37	2.35	2.38
Cys94-S ^γ ...Zn	2.27, 2.25	2.27, 2.27	2.32	2.30	2.32
O4 ^c ...Zn	2.06, 2.08	2.05, 2.06	2.12	2.14	2.14
Cys91-N ...O4	3.02, 3.10	3.02, 3.03	3.14	3.23	3.13
Glu64-N ...O ^{ε1} -Glu64	2.69, 2.70	2.69, 2.72	2.78	3.05	2.80
O4 ...O ^{ε2} -Glu64	2.60, 2.45	2.49, 2.52	2.48	2.58	2.51
N3 ...O ^{ε1} -Glu64	2.82, 2.81	2.80, 2.78	2.88	3.64	2.87
O2 ...N-Gly63	2.84, 2.79	2.89, 2.87	2.86	2.91	2.83
O2 ...N ^{δ2} -Asn51	2.96, 2.95	2.92, 2.90	2.91	2.82	2.93
N1 ...O ^{δ1} -Asp155	2.67, 2.70	2.69, 2.68	2.80	2.81	2.81
Asn51-N ^{δ2} ...O ^{δ2} -Asp155	2.95, 2.98	2.92, 2.90	3.02	2.90	3.04
His62-N ^{ε2} ...O-Asp155	2.74, 2.76	2.67, 2.70	2.89	2.86	2.89
N4 ...O-Ser89				3.48	3.41
N4 ...O ^{ε1} -Glu64				2.94	3.24
O4 ...C4	1.44, 1.44	1.48, 1.48	1.47	2.90	1.51
N3 ...C4	1.42, 1.41	1.46, 1.45	1.45	1.34	1.46
C5 ...C4	1.45, 1.45	1.47, 1.46	1.50	1.45	1.50
N3 ...C2	1.37, 1.37	1.31, 1.31	1.36	1.37	1.37
N1 ...C2	1.39, 1.39	1.32, 1.33	1.38	1.40	1.38
N1 ...C6	1.38, 1.38	1.43, 1.42	1.38	1.36	1.38
C5 ...C6	1.34, 1.35	1.31, 1.33	1.35	1.37	1.35
O2 ...C2	1.21, 1.21	1.30, 1.30	1.28	1.27	1.27
∠O4-C4-N3	110, 109	107, 108	108	91	106
∠O4-C4-C5	110, 109	112, 109	110	94	108
∠O4-C4-N3-C2	113, 119	113, 114	120	99	119
∠O4-C4-C5-C6	-125, -116	-113, -114	-117	-91	-115
∠C ^β -C ^γ -C ^δ -O ^{ε2} (Glu64)	-65, -68	-68, -68	-72	-120	-78

a: Distances between the paired atoms are in Å. Angles and dihedral angles are in degrees. b: Distances in subunits A and B, respectively. c: O4 is the oxygen atom in the water, or the hydroxide oxygen of DHP, cytosine or other intermediates. Atom names are presented in the same manner for C1, C2, C4, C5, C6, N1, N3, N4 and O2.

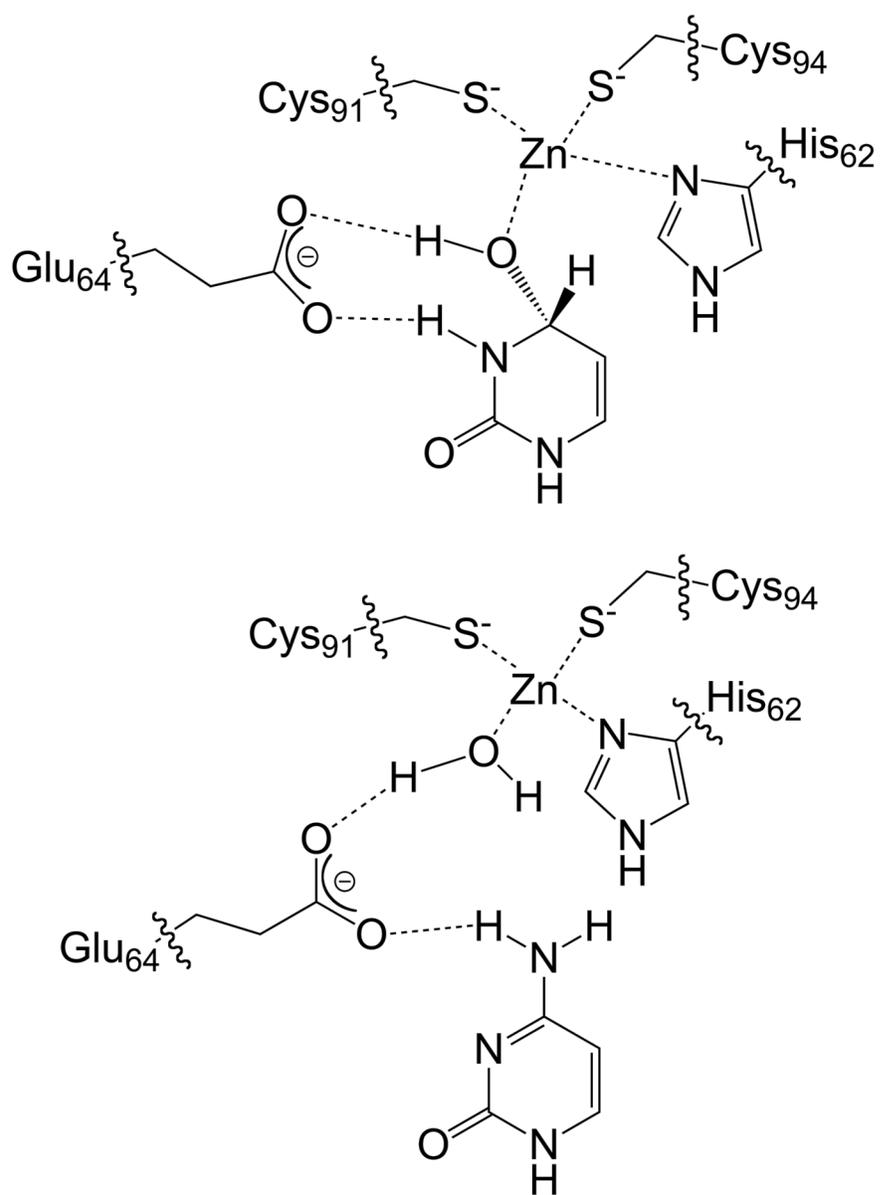


Figure 4-1. QM partition in the simulation systems for the yCD-DHP complex (top) and with the yCD water/cytosine complex (bottom).

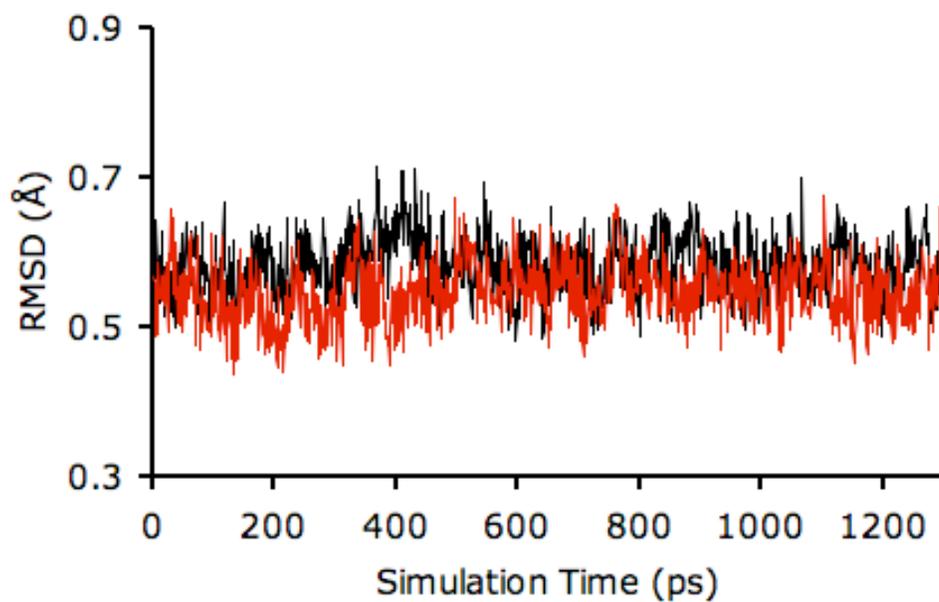


Figure 4-2. RMSDs of the backbone C α atoms relative to the crystal structure during the QM/MM MD simulation for the yCD-DHP complex. Black curve is for subunit A and red curve is for subunit B. The first 500 ps was regarded as the equilibrium stage. The data analysis was based on the last 800 ps trajectory.

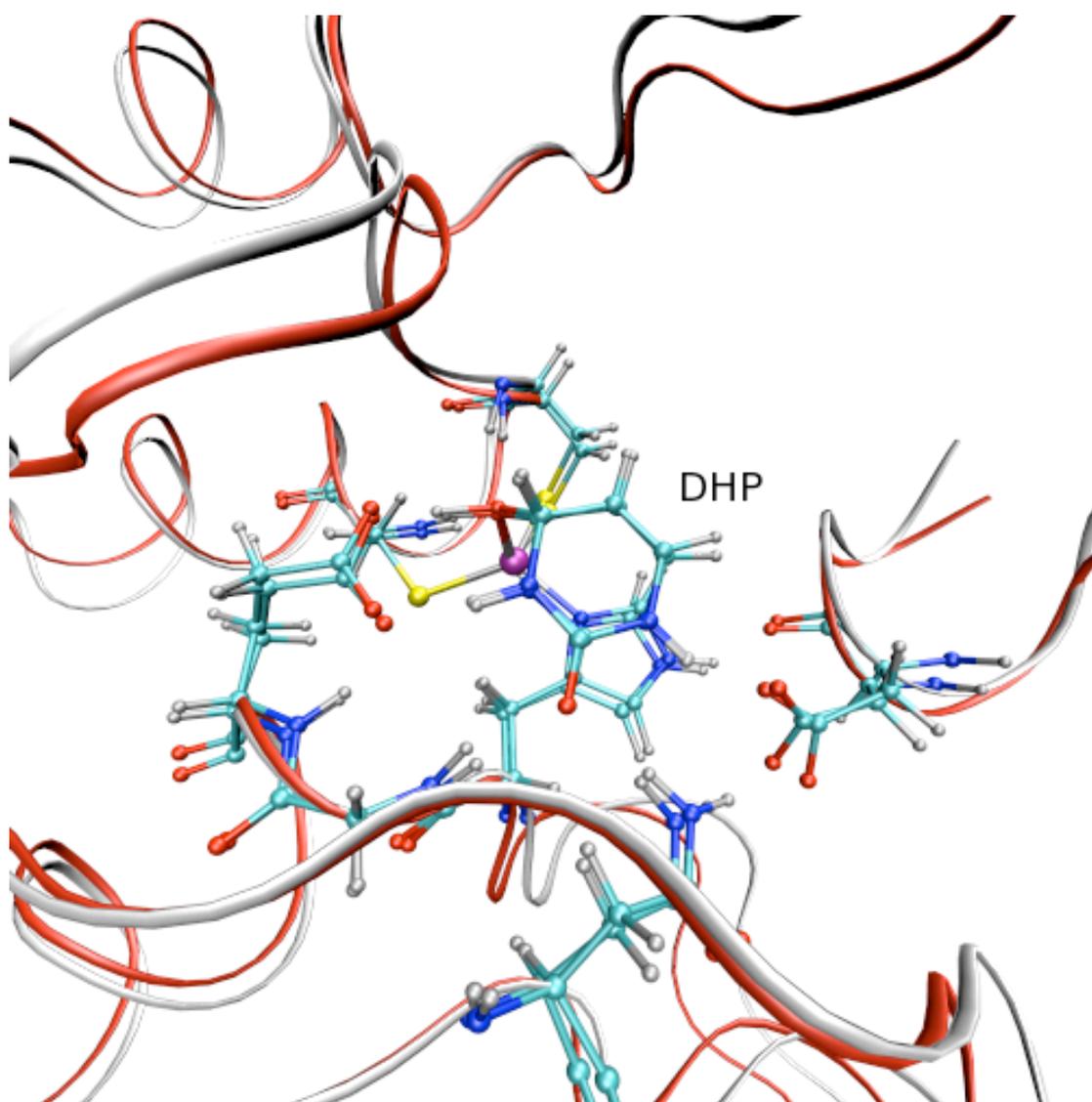


Figure 4-3. Superimposition of a representative snapshot during the yCD-DHP complex simulation with the crystal structure. The backbone was displayed as a red tube (crystal structure) and gray tube (MD). DHP and several surrounding residues were displayed as ball and stick type.

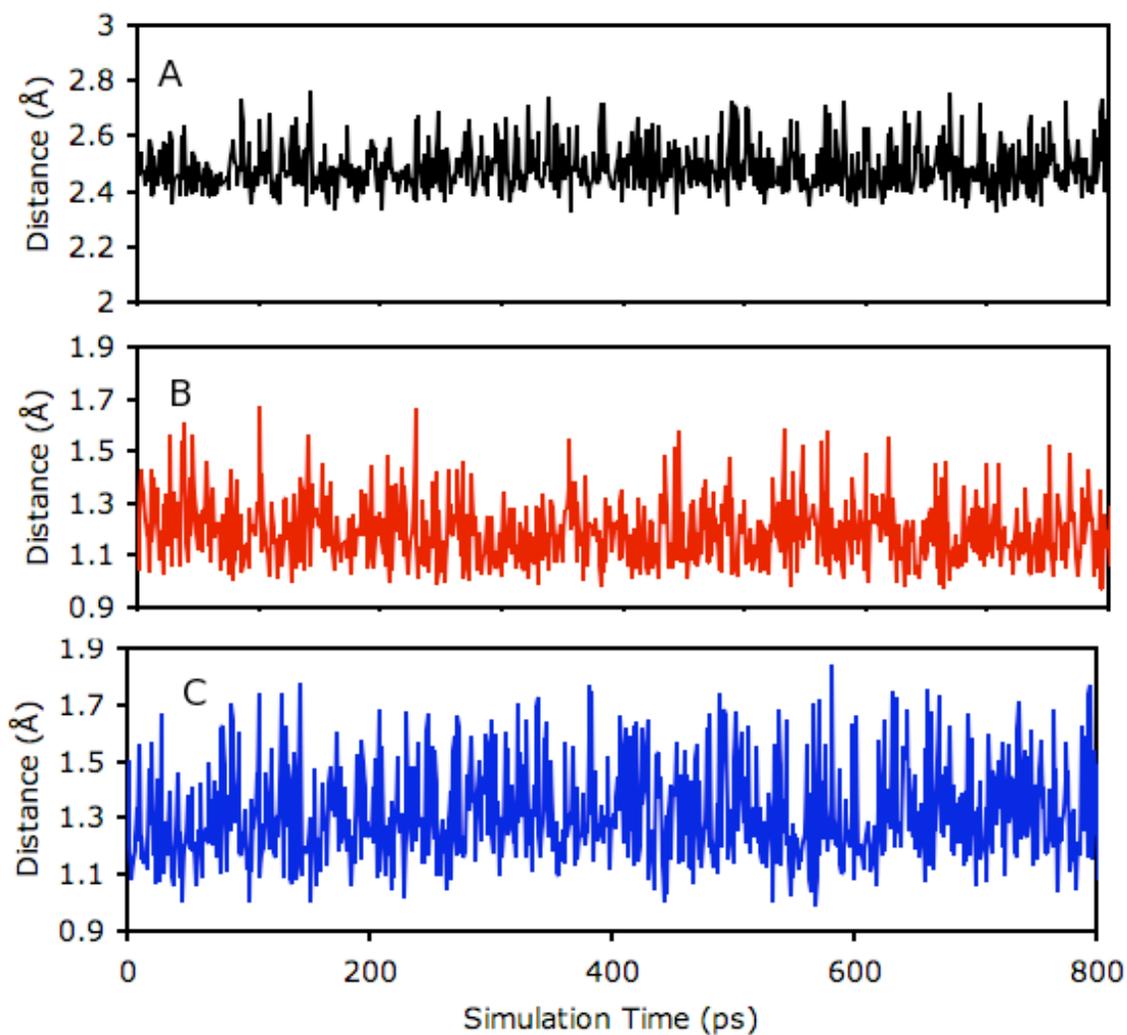


Figure 4-4. Bond distances of snapshots along the MD trajectory of the yCD-DHP complex simulation. (A) The hydrogen bond distance between Glu64-O^{ε2} and DHP-O4. (B) The bond distance between Glu64-O^{ε2} and HO4. (C) The bond distance between DHP-O4 and HO4.

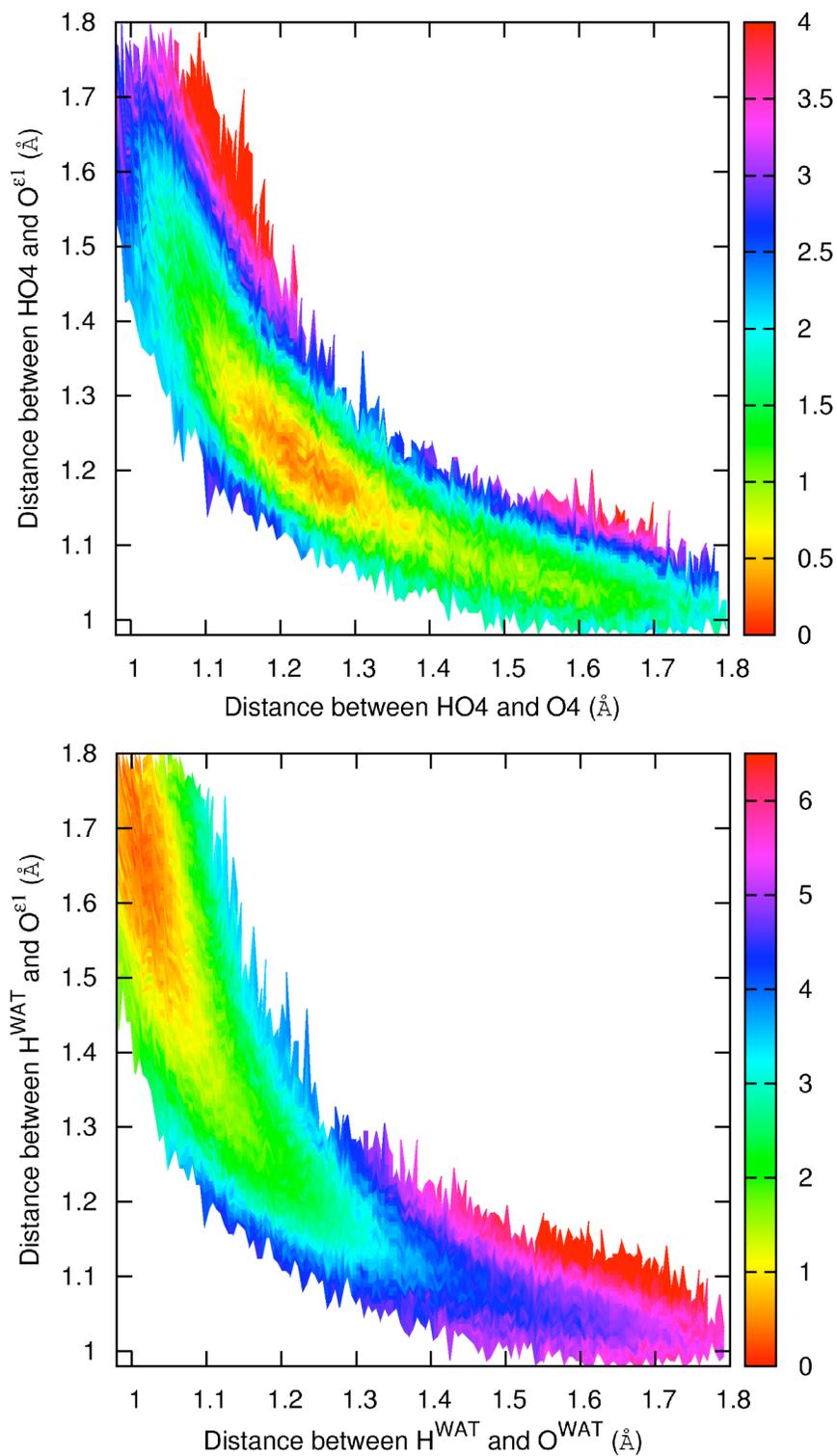


Figure 4-5. Energy profiles for proton transfer. (top) H^{WAT} is transferred from the zinc-coordinated water to O^{ε2} of Glu64 in the free from yCD. (bottom) HO4 is transferred from O4 of DHP to O^{ε2} of Glu64 in the yCD-DHP complex.

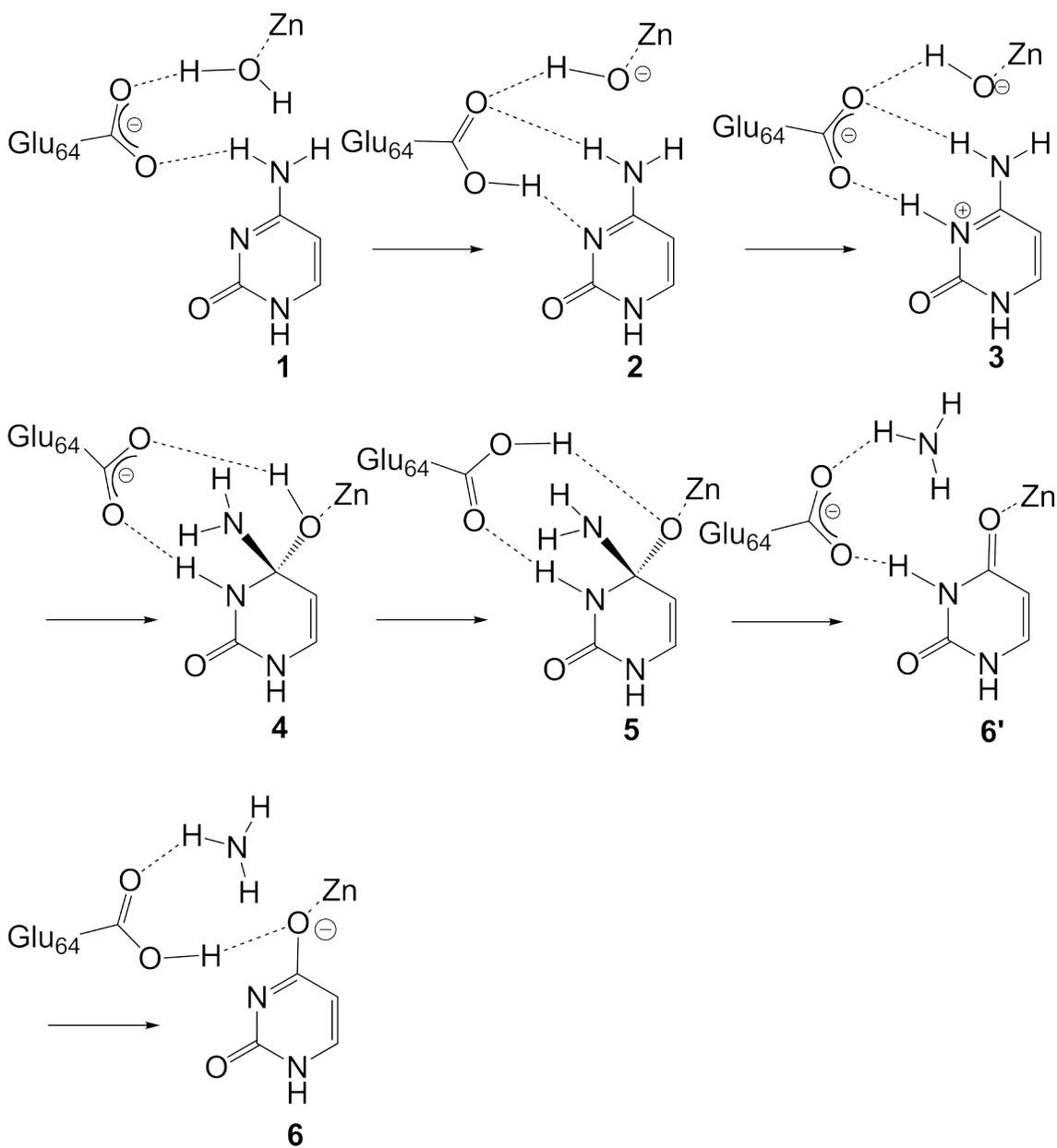


Figure 4-6. Proposed reaction pathway for the conversion of cytosine to uracil in the QM/MM MD simulations. (1) yCD water/cytosine complex. (2) yCD hydroxide/Glu64H cytosine complex. (3) yCD hydroxide/Glu64 cytosine-H complex. (4) yCD protonated tetrahedral intermediate complex. (5) yCD deprotonated tetrahedral intermediate complex. (6') yCD zinc bound uracil/ammonia complex. (6) yCD Glu64H deprotonated zinc bound uracil/ammonia complex.

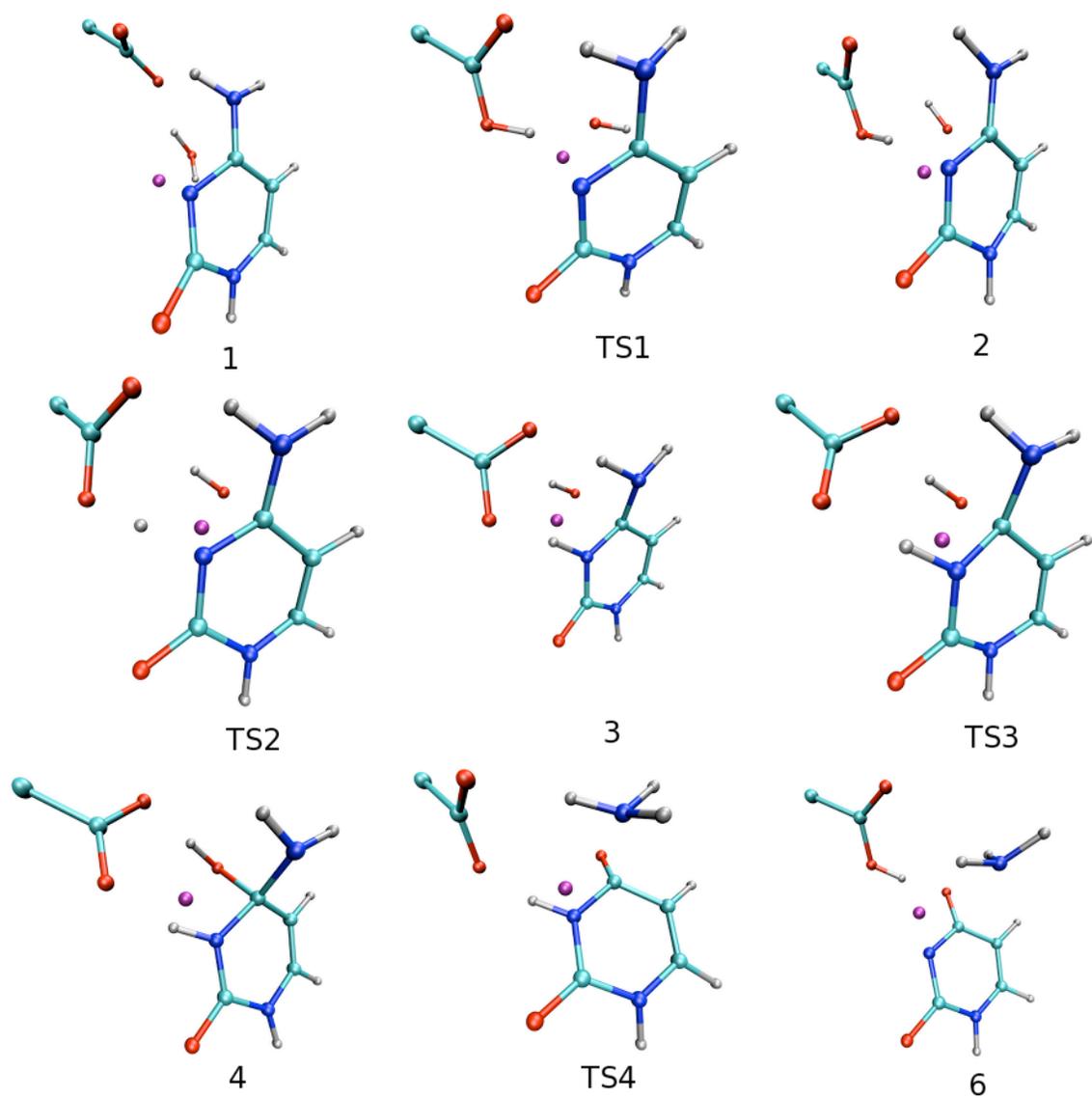


Figure 4-7. Snapshots of the reaction species along the reaction pathway for the conversion of cytosine to uracil in the QM/MM MD simulations.

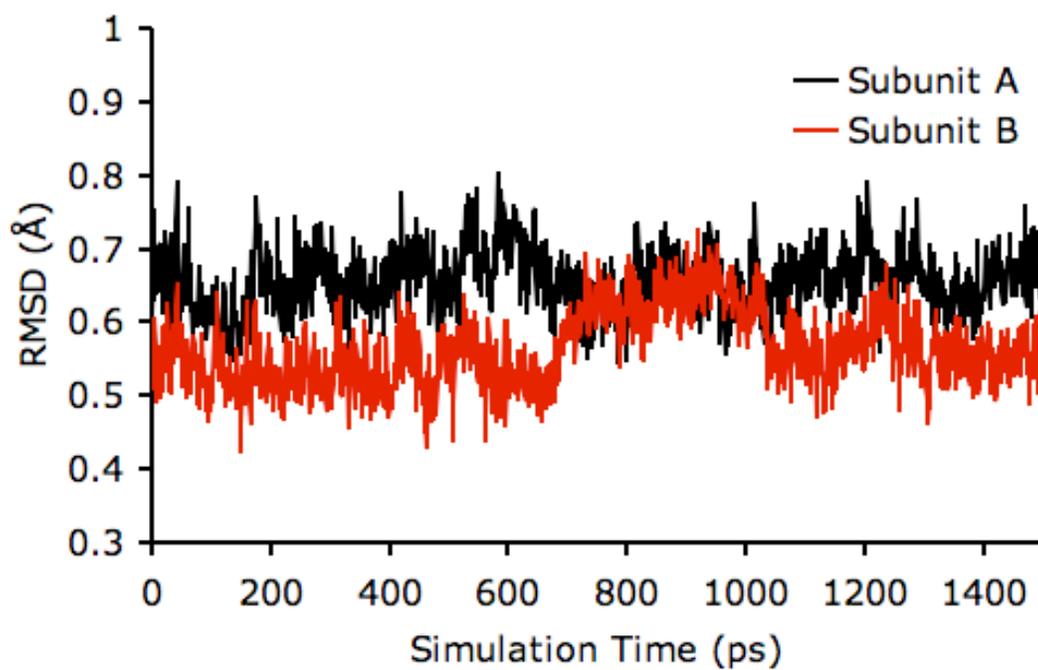


Figure 4-8. RMSDs of the backbone $C\alpha$ atoms relative to the crystal structure during the QM/MM MD simulation for the yCD water/cytosine complex (**1**). Black curve is for subunit A and red curve is for subunit B. The first 500 ps was regarded as the equilibrium stage. The data analysis was based on the last 1000 ps trajectory.

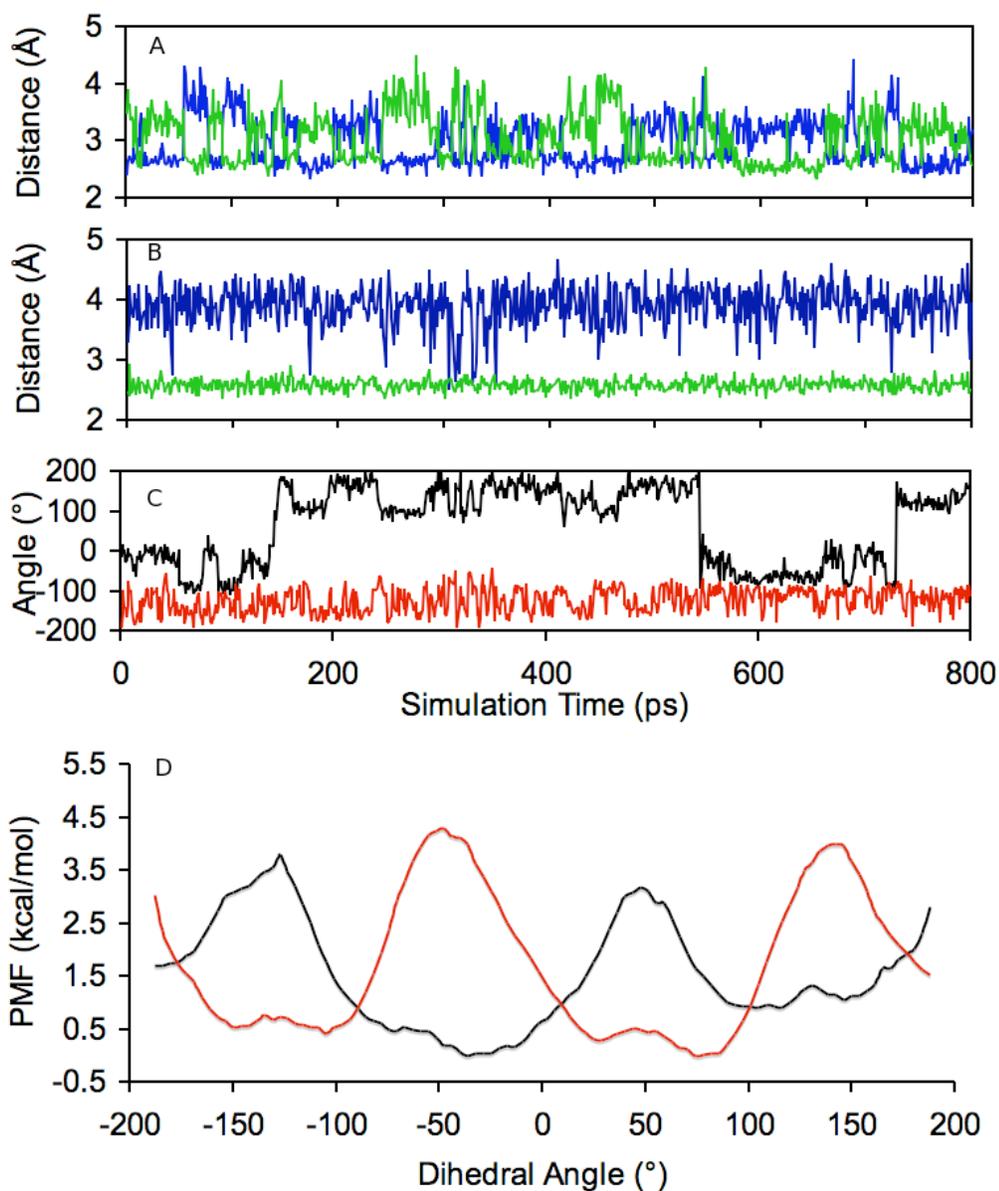


Figure 4-9. Rotation of the sidechain of Glu64 ($\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\epsilon 2}$) in the free form of yCD and the yCD substrate complex simulations. (A) Distance between $O^{\epsilon 1}$ or $O^{\epsilon 2}$ of Glu64 and the zinc-coordinated water (O^{Zn}) in the yCD free form simulation. The curve for the distance between $O^{\epsilon 1}$ and O^{Zn} is colored blue and between $O^{\epsilon 2}$ and O^{Zn} is colored green. (A) Distance between $O^{\epsilon 1}$ or $O^{\epsilon 2}$ of Glu64 and the zinc-coordinated water (O^{Zn}) in yCD substrate form simulation. The curves are colored as the same as in (A). (C) Rotation of the side chain of Glu64 ($\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\epsilon 2}$). The curve is colored black for the free form simulation and colored red for the yCD substrate complex. (D) PMF for rotation of the side chain of Glu64 ($\angle C^{\beta}-C^{\gamma}-C^{\delta}-O^{\epsilon 2}$). The curves are colored as the same as in (C).

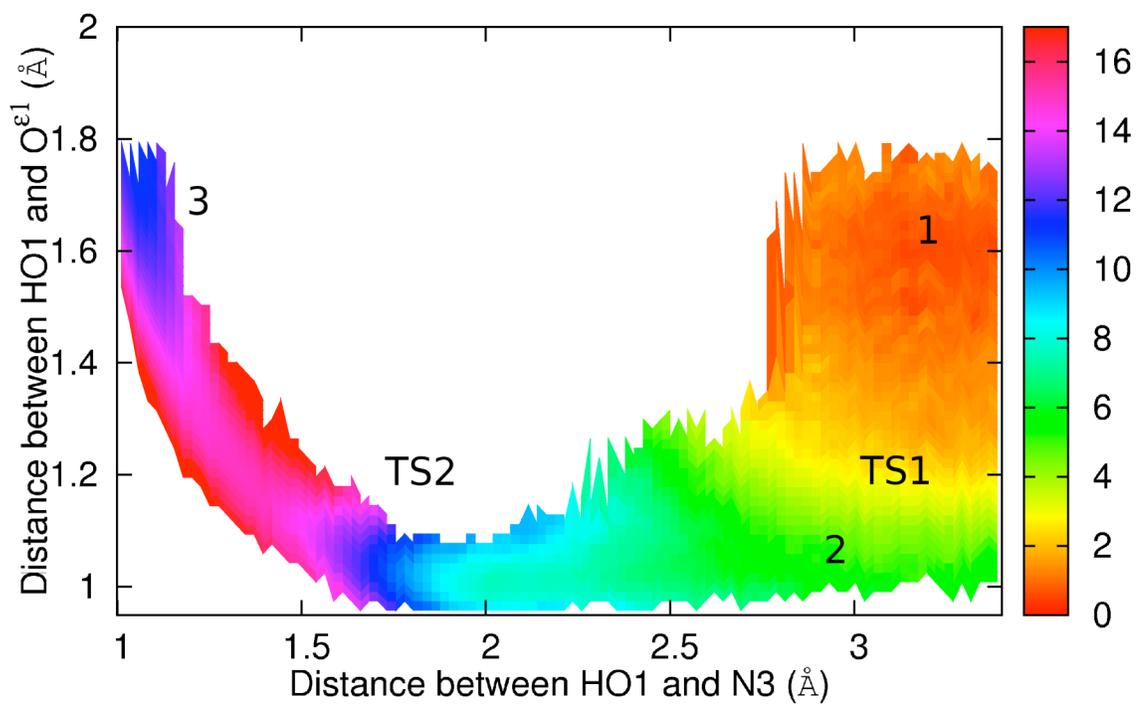


Figure 4-10. Potential of mean force for the proton transfer from the zinc-coordinated water to $O^{\epsilon 2}$ of Glu64 and to N3 of cytosine.

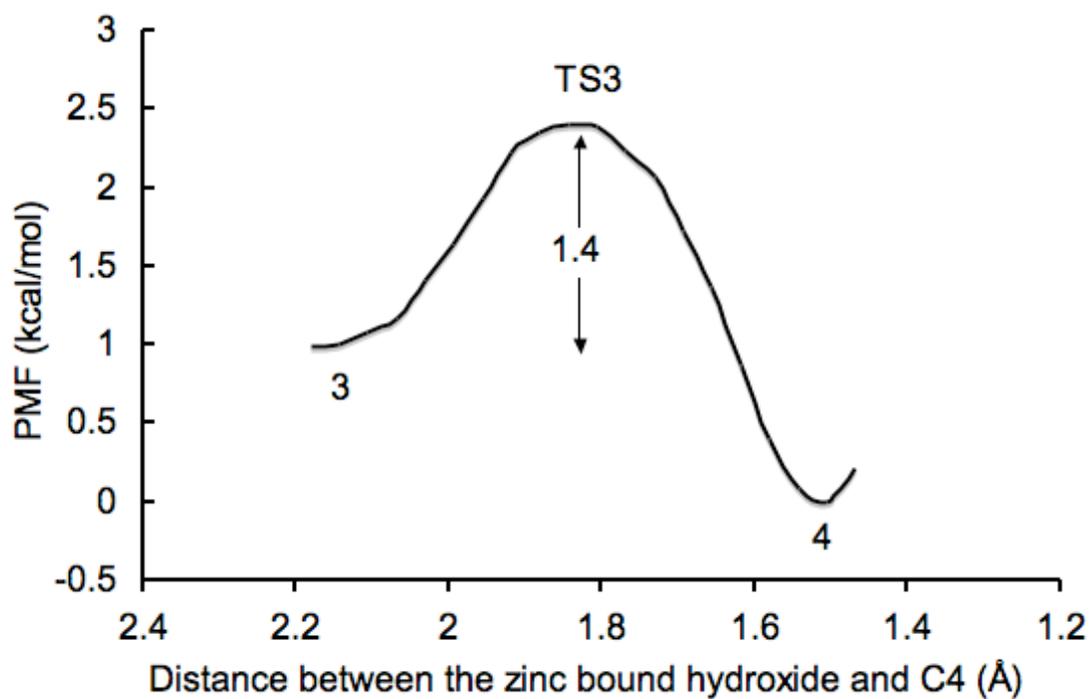


Figure 4-11. Potential of mean force for nucleophilic attack on C4 of cytosine. The tetrahedral intermediate complex (4) is 1.0 kcal/mol more stable than the hydroxide/Glu64 cytosine-H complex (3). The energy barrier is calculated to be 1.4 kcal/mol.

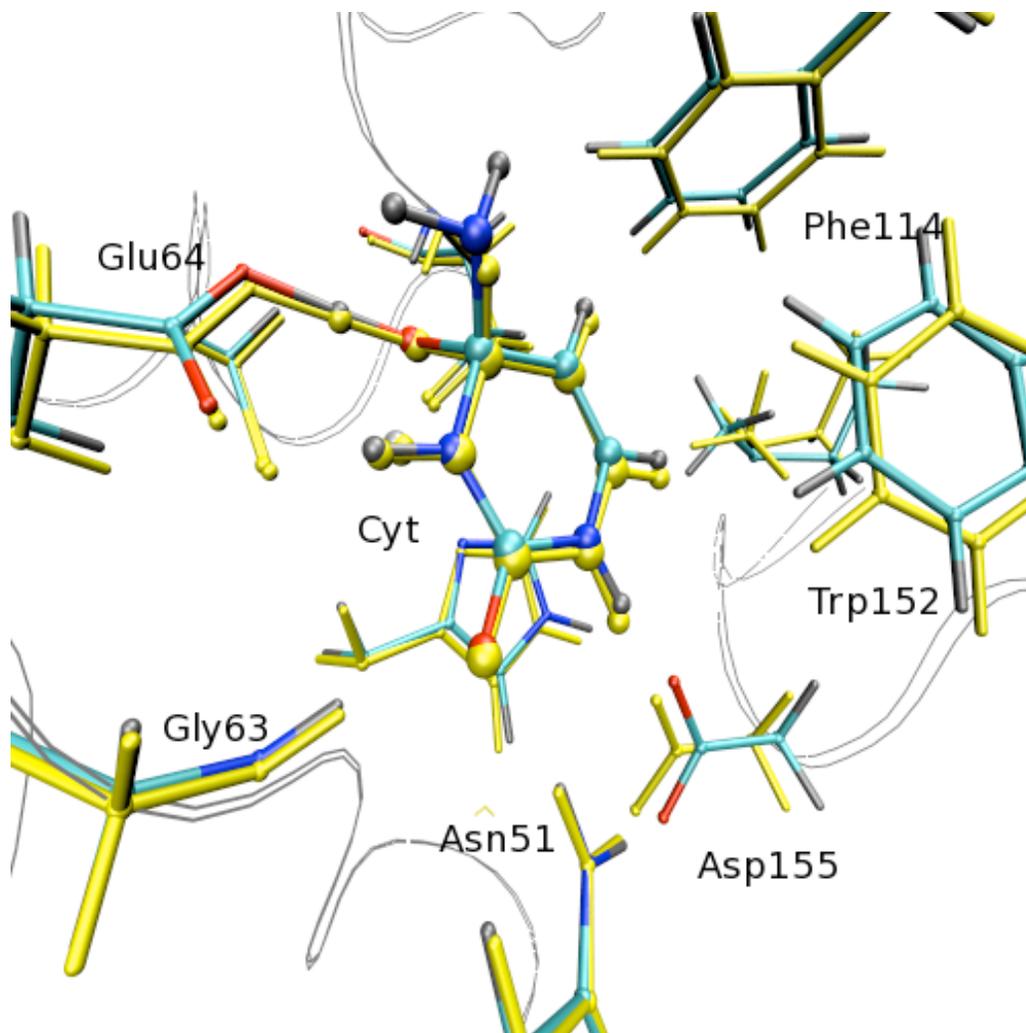


Figure 4-12. Superposition of the average structure of snapshots in the MD trajectory for the complex 4 simulation onto that of the yCD-DHP simulation.

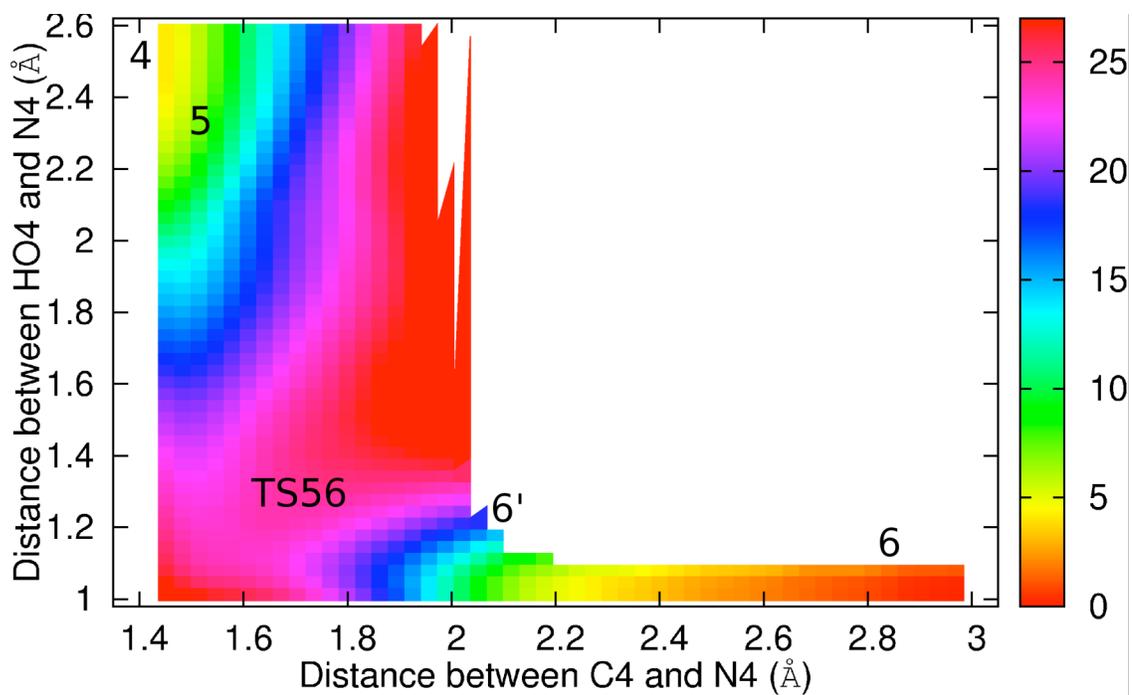


Figure 4-13. Potential of mean force for the C4-N4 bond cleavage.

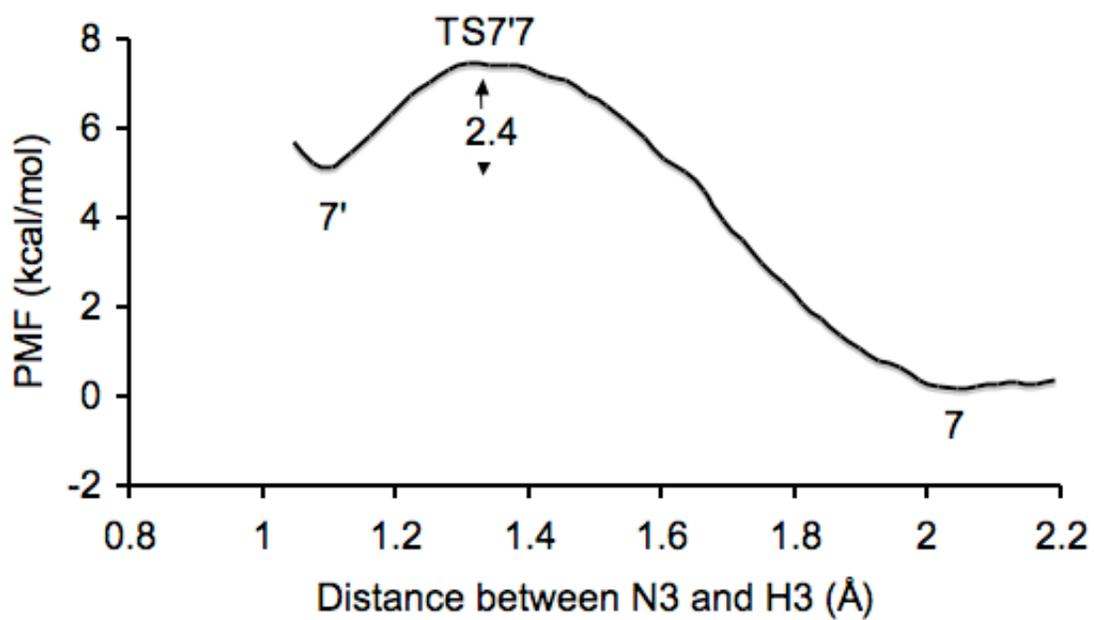


Figure 4-14. Potential of mean force for proton transfer from N3 to O ϵ^2 . The transition state is found when C3-H3 distance is 1.3 Å. The energy barrier is 2.4 kcal/mol.

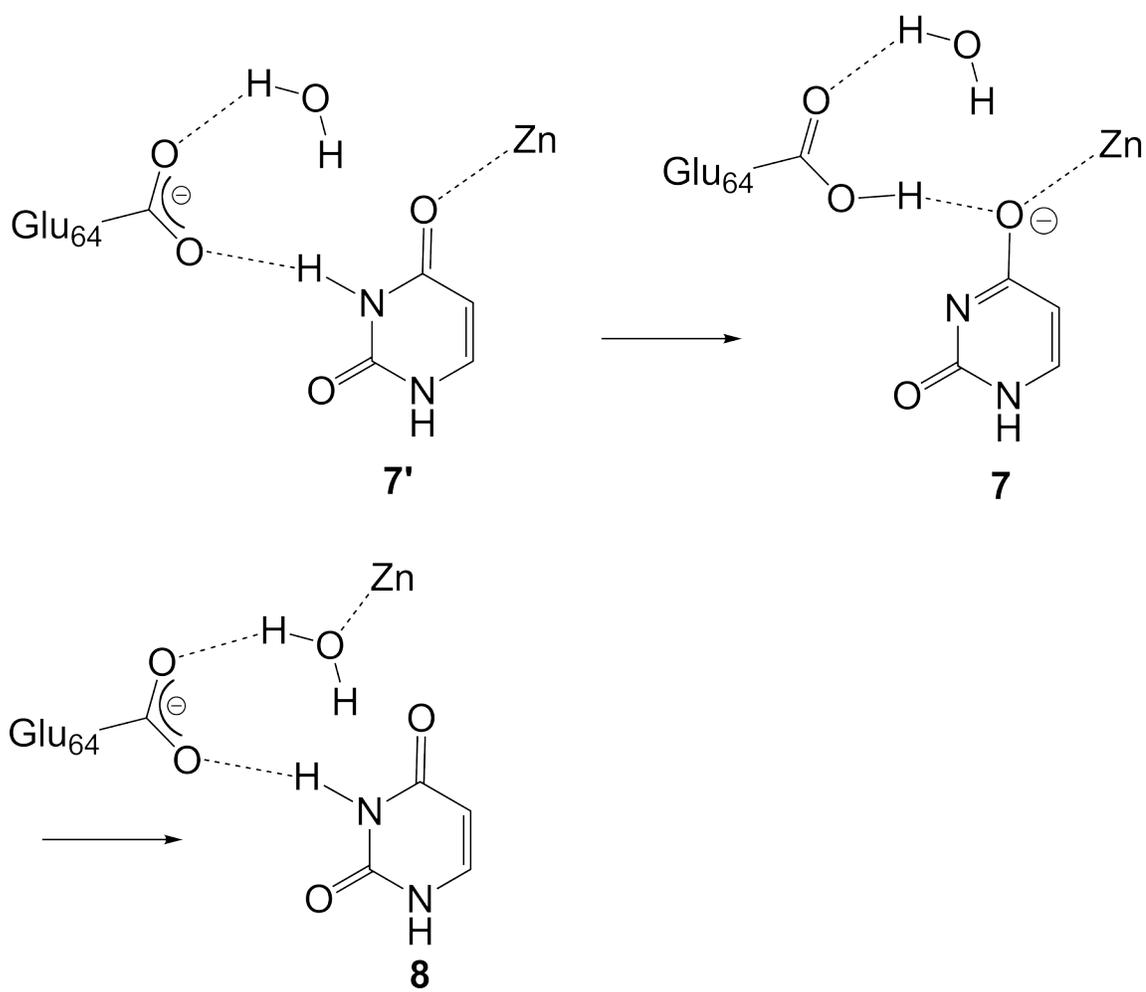


Figure 4-15. Reaction pathway for the $O^{\text{Zn}}\text{-Zn}$ bond breakage.

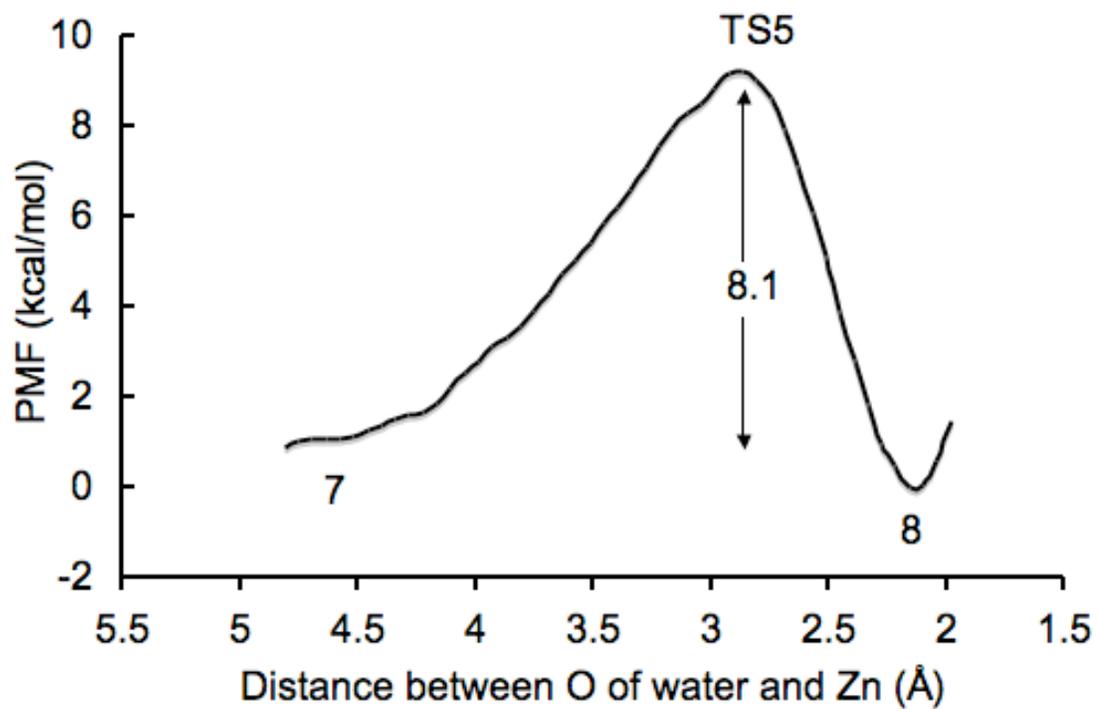


Figure 4-16. Potential of mean force for the $\text{O}^{\text{Zn}}\text{-Zn}$ bond breakage. The transition state is found when $\text{O}^{\text{Zn}}\text{-Zn}$ distance is 2.9 Å. The energy barrier is 8.1 kcal/mol.

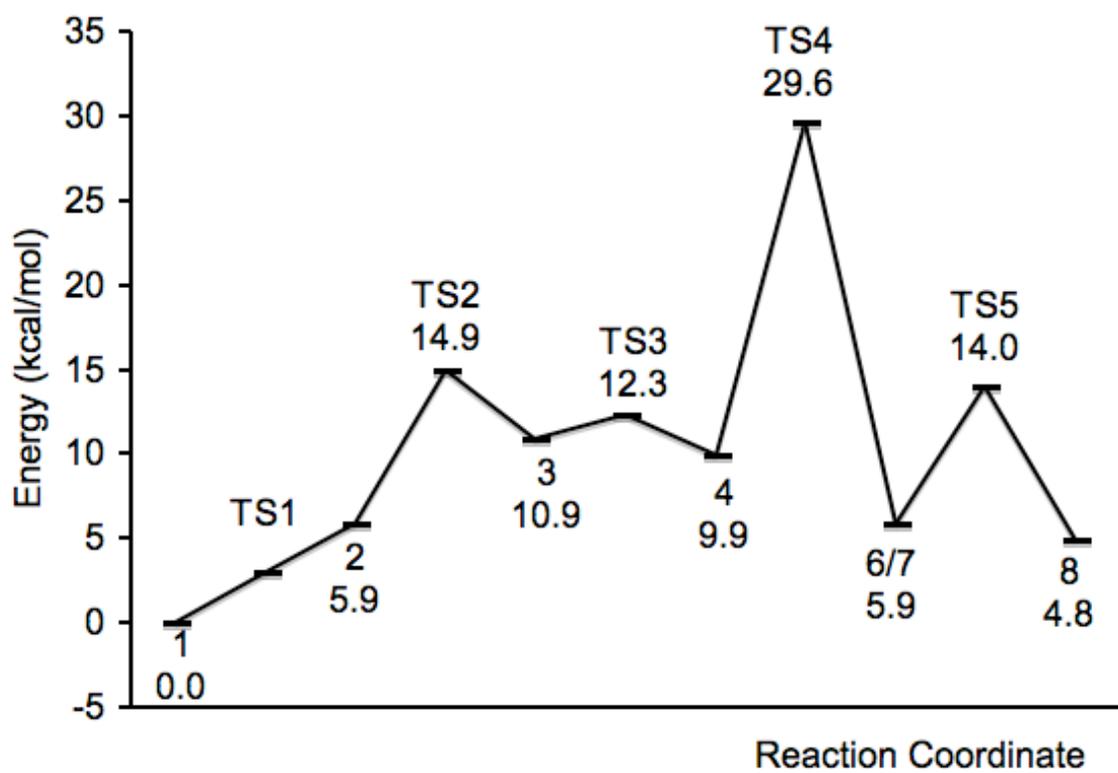


Figure 4-17. Overall energy profile for the conversion of cytosine to uracil by yCD. The overall rate-limiting step is found to be the C4-N4 bond cleavage step, coupled with multiple proton transfers.

REFERENCES

REFERENCES

1. Greco, O., and Dachs, G. U. (2001) Gene directed enzyme/prodrug therapy of cancer: Historical appraisal and future perspectives, *J Cell Physiol* 187, 22-36.
2. Aghi, M., Hochberg, F., and Breakefield, X. (2000) Prodrug activation enzymes in cancer gene therapy, *J Gene Med* 2, 148-164.
3. Ko, T., Lin, J., Hu, C., Hsu, Y., Wang, A., and Liaw, S. (2003) Crystal structure of yeast cytosine deaminase - Insights into enzyme mechanism and evolution, *J Biol Chem* 278, 19111-19117.
4. Ireton, G., Black, M., and Stoddard, B. (2003) The 1.14 angstrom crystal structure of yeast cytosine deaminase: Evolution of nucleotide salvage enzymes and implications for genetic chemotherapy, *Structure* 11, 961-972.
5. Sklenak, S., Yao, L., Cukier, R., and Yan, H. (2004) Catalytic mechanism of yeast cytosine deaminase: An ONIOM computational study, *J Am Chem Soc* 126, 14879-14889.
6. Torrent, M., Vreven, T., Musaev, D. G., Morokuma, K., Farkas, O., and Schlegel, H. B. (2002) Effects of the protein environment on the structure and energetics of active sites of metalloenzymes. ONIOM study of methane monooxygenase and ribonucleotide reductase, *J Am Chem Soc* 124, 192-193.
7. Walker, R. C., Crowley, M. F., and Case, D. A. (2008) The implementation of a fast and accurate QM/MM potential method in Amber, *J Comput Chem* 29, 1019-1031.
8. Seabra, G. D. M., Walker, R. C., Elstner, M., Case, D. A., and Roitberg, A. E. (2007) Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package, *J Phys Chem A* 111, 5655-5664.
9. Hu, H., Lu, Z., and Yang, W. (2007) QM/MM minimum free energy path: Methodology and application to triosephosphate isomerase, *J Chem Theory Comput* 3, 390-406.
10. Okamoto, T., Yamada, K., Koyano, Y., Asada, T., Koga, N., and Nagaoka, M. (2011) A minimal implementation of the AMBER-GAUSSIAN interface for *ab initio* QM/MM-MD simulation, *J Comput Chem* 32, 932-942.
11. Torras, J., Seabra, G. D. M., Deumens, E., Trickey, S. B., and Roitberg, A. E. (2008) A versatile AMBER-Gaussian QM/MM interface through PUPIL, *J Comput Chem* 29, 1564-1573.
12. Woodcock, H. L., Hodoscek, M., Gilbert, A. T. B., Gill, P. M. W., Schaefer, H. F., and Brooks, B. R. (2007) Interfacing Q-Chem and CHARMM to perform QM/MM reaction path calculations, *J Comput Chem* 28, 1485-1502.
13. Riccardi, D., König, P., Prat-Resina, X., Yu, H., Elstner, M., Frauenheim, T., and Cui, Q. (2006) "Proton Holes" in long-range proton transfer reactions in solution and enzymes: A theoretical analysis, *J Am Chem Soc* 128, 16302-16311.

14. Gao, J., and Truhlar, D. G. (2002) Quantum mechanical methods for enzyme kinetics, *Annu Rev Phys Chem* 53, 467-505.
15. Yao, L., Yan, H., and Cukier, R. (2006) A combined ONIOM quantum chemical-molecular dynamics study of zinc-uracil bond breaking in yeast cytosine deaminase, *J Phys Chem B* 110, 26320-26326.
16. Matsubara, T., Dupuis, M., and Aida, M. (2007) The ONIOM molecular dynamics method for biochemical applications: Cytidine deaminase, *Chem Phys Lett* 437, 138-142.
17. Matsubara, T., Dupuis, M., and Aida, M. (2008) An insight into the environmental effects of the pocket of the active site of the enzyme. *Ab initio* ONIOM-molecular dynamics (MD) study on cytosine deaminase, *J Comput Chem* 29, 458-465.
18. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J Comput Chem* 4, 187-217.
19. Xu, Q., Guo, H., Gorin, A., and Guo, H. (2007) Stabilization of a transition-state analogue at the active site of yeast cytosine deaminase: Importance of proton transfers, *J Phys Chem B* 111, 6501-6506.
20. Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S., and Seifert, G. (1998) Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties, *Phys Rev B* 58, 7260-7268.
21. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins* 65, 712-725.
22. Yao, L., Sklenak, S., Yan, H., and Cukier, R. (2005) A molecular dynamics exploration of the catalytic mechanism of yeast cytosine deaminase, *J Phys Chem B* 109, 7500-7510.
23. Case, D. A., Darden, T. A., Cheatham, I., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., and Kollman, P. A. (2008) AMBER 10. University of California, San Francisco.
24. Loncharich, R. J., Brooks, B. R., and Pastor, R. W. (1992) Langevin dynamics of peptides: The frictional dependence of isomerization rates of *N*-acetylalanyl-*N'*-methylamide, *Biopolymers* 32, 523-535.
25. Ryckaert, J., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J Comput Phys* 23, 327-341.

26. Elstner, M. (2005) The SCC-DFTB method and its application to biological systems, *Theor Chem Acc* 116, 316-325.
27. Riccardi, D., and Cui, Q. (2007) pK_a analysis for the zinc-bound water in human carbonic anhydrase II: Benchmark for "multiscale" QM/MM simulations and mechanistic implications, *J Phys Chem A* 111, 5703-5711.
28. Nam, K., Gao, J., and York, D. M. (2005) An efficient linear-scaling Ewald method for long-range electrostatic interactions in combined QM/MM calculations, *J Chem Theory Comput* 1, 2-13.
29. Cui, Q., Elstner, M., Kaxiras, E., Frauenheim, T., and Karplus, M. (2001) A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method, *J Phys Chem B* 105, 569-585.
30. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, *J Chem Phys* 98, 10089-10092.
31. Roux, B. (1995) The calculation of the potential of mean force using computer simulations, *Comput Phys Commun* 91, 275-282.
32. Grossfield, A. WHAM: the weighted histogram analysis method, version 2.0, <http://membrane.urmc.rochester.edu/content/wham>.
33. Tee, O. S., and Paventi, M. (1980) Kinetics and mechanism of the reaction of 2(IH)-pyrimidinone with bromine. Reaction via its covalent hydrate, *J Org Chem* 45, 2072-2076.
34. Cleland, W., and Kreevoy, M. (1994) Low-barrier hydrogen bonds and enzymic catalysis, *Science* 264, 1887-1890.
35. Fersht, A. (1999) Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding. W. H. Freeman and Company, New York.
36. Warshel, A., and Papazyan, A. (1996) Energy considerations show that low-barrier hydrogen bonds do not offer a catalytic advantage over ordinary hydrogen bonds, *Proc Natl Acad Sci U S A* 93, 13665-13670.
37. Perrin, C. L., and Nielson, J. B. (1997) "Strong" hydrogen bonds in chemistry and biology, *Annu Rev Phys Chem* 48, 511-544.
38. Tuckerman, M. E., Marx, D., Klein, M. L., and Parrinello, M. (1997) On the quantum nature of the shared proton in hydrogen bonds, *Science* 275, 817-820.
39. Cleland, W. W., Frey, P. A., and Gerlt, J. A. (1998) The low barrier hydrogen bond in enzymatic catalysis, *J Bio Chem* 273, 25529 -25532.
40. Schutz, C. N., and Warshel, A. (2004) The low barrier hydrogen bond (LBHB) proposal revisited: The case of the Asp ... His pair in serine proteases, *Proteins* 55, 711-723.

41. Yamaguchi, S., Kamikubo, H., Kurihara, K., Kuroki, R., Niimura, N., Shimizu, N., Yamazaki, Y., and Kataoka, M. (2009) Low-barrier hydrogen bond in photoactive yellow protein, *Proc Natl Acad Sci U S A* 106, 440 -444.
42. Kimura, E., Kitamura, H., Koike, T., and Shiro, M. (1997) Facile and selective electrostatic stabilization of uracil N(1)⁻ anion by a proximate protonated amine: A chemical implication for why uracil N(1) is chosen for glycosylation site, *J Am Chem Soc* 119, 10909-10919.

CHAPTER 5

ROLE OF GLUTAMATE 64 IN THE ACTIVATION OF THE PRODRUG 5FC

INTRODUCTION

A major problem in cancer treatment is the toxicity of anticancer drugs to normal cells. Gene directed enzyme prodrug therapy (GDEPT) is aimed to address this critical problem by activating nontoxic prodrugs in tumors, and thereby killing cancer cells while minimizing the side effects of anticancer drugs (1, 2). One of enzyme prodrug combinations frequently used in GDEPT is cytosine deaminase (CD) and 5-fluorocytosine (5FC). CD is a metalloenzyme that catalyzes the deamination of cytosine to generate uracil as well as the deamination of 5FC to generate 5-fluorouracil (5FU). 5FU is an anticancer drug used for the treatment of colorectal, breast, stomach, and pancreatic cancers, but, like many anticancer drugs, the drug is highly toxic, causing side effects such as myelosuppression, mucositis, dermatitis, and diarrhea. 5FC is fairly nontoxic, as CD is absent in human. By activating the prodrug 5FC in the tumor, the CD/5FC-based GDEPT minimizes the side effects of 5FU.

Two CDs have been used as the activating enzyme in CD/5FC-based GDEPT. One is from *E. coli* (eCD), and the other from yeast (yCD). The two enzymes have no

detectable homology and arise apparently by convergent evolution. The *E. coli* enzyme is a homohexameric protein with a molecular weight of 47.5 kDa per subunit (3), whereas the yeast enzyme is a homodimeric protein with a molecular weight of 17.5 kDa per subunit (4-6). Although eCD was used in early GDEPT studies, more recent studies showed that yCD is preferable for GDEPT, because yCD has a much higher catalytic efficiency for 5FC than eCD (7-9).

Because of its biomedical significance, there have been quite a few experimental and computational studies of the structure and catalytic mechanism of yCD. The structure of yCD has been determined by X-ray crystallography at high resolution, including the substrate-free form (10) and the complex form with the inhibitor 2-pyrimidinone (Py) (10, 11). The bound Py is in the hydrated form, 4(*R*)-hydroxyl-3,4-dihydropyrimidinone (DHP), mimicking the transition state for the deamination of cytosine. The key elements of the catalytic apparatus of yCD are believed to be the zinc ion coordinated with DHP and Glu64 with two hydrogen bonds to DHP (Figure 1-3). The catalytic mechanism of yCD has been investigated by both experimental and computational approaches (12-16). Transient kinetic and NMR studies have shown that product release is rate-limiting in the activation of 5FC by yCD (17). A complete reaction pathway for the deamination of cytosine by yCD has been proposed by two-layered quantum mechanics calculations using the ONIOM method (16) and combined molecular dynamics simulations and ONIOM calculations (14). In the proposed reaction pathway, Glu64 plays a critical role in catalysis, serving as a proton shuttle for both the generation of the tetrahedral reaction intermediate from substrate cytosine and the conversion of the intermediate to the product

uracil. However, the role of Glu64 in the enzymatic reaction has not been investigated experimentally.

In the present work, we have investigated the role of Glu64 in the activation of 5FC by combined experimental and computational studies. The results show that Glu64 is important not only for chemical transformation but also for substrate binding in the activation of 5FC. The role of Glu64 in substrate binding is also of great importance for pyrimidine salvage, the physiological role of yCD.

EXPERIMENTAL PROCEDURES

Materials. Cytosine, uracil, 5-fluorocytosine and 5-fluorouracil were purchased from Sigma. [6-³H]-5-fluorocytosine was purchased from Moravek. 5-fluoropyrimidinone was synthesized in our lab. Restriction enzymes and T4 ligase were purchased from New England Biolabs. *Pfu* DNA polymerase and the pET-17b vector were purchased from Stratagene and Novagen, respectively.

Cloning and Site-Directed Mutagenesis. The pET-17bHR vector was constructed by adding DNA encoding a His tag and a TEV protease cleavage site to the pET-17b vector (pET17bHR). The expression construct of the His-tagged yCD with a TEV protease cleavage site (pET17bHR-yCD) was made by PCR. The primers for PCR were 5'-GGG ATC CAT ATG GCA AGC AAG TGG GAT CAG-3' and 5'-GGA ATT CTA CTC ACC AAT ATC TTC AAA CC-3'. The PCR product was digested with *NdeI* and *EcoRI* restriction enzymes and ligated with the vector pET-17bHR digested with the same restriction enzymes. The ligation mixture was transformed into *E. coli* strain DH5 α . The correct coding sequence of the cloned yCD gene was verified by DNA sequencing.

The DNA of the expression construct (pET17bHR-yCD) was then transformed into *E. coli* strain BL21(DE3) pLysS.

The expression construct encoding the E64A mutant protein (pET17b-E64A) was made by PCR-based site-directed mutagenesis using the construct pET17b-yCD described elsewhere (17). The forward and reverse primers for the E64A mutant construct are 5'-CC ACA CAT GGT GCG ATC TCC ACT TTG GAA AAC-3' and 5'-GTT TTC CAA AGT GGA GAT CGC ACC ATG TAG TGT GG-3'. The construct was transformed into *E. coli* strain DH5 α . The mutations were verified by DNA sequencing. The mutant DNA constructs were transformed into *E. coli* strain BL21(DE3) pLysS.

The expression construct encoding the His-tagged E64A mutant enzyme with a TEV protease cleavage site (pET17bHR-E64A) was made by ligating the pET17b-E64A construct and the pET-17bHR vector, both digested with *Nde*I and *Eco*RI restriction enzymes. The construct was transformed into *E. coli* strain DH5 α . The mutations were verified by DNA sequencing. The mutant DNA constructs were transformed into *E. coli* strain BL21(DE3) pLysS.

The expression construct encoding the His-tagged E64D mutant enzyme with a TEV protease cleavage site (pET17bHR-E64D) was made by PCR-based site-directed mutagenesis using the construct pET17bHR-yCD. The forward and reverse primers for the E64D mutant construct were 5'-CC ACA CTA CAT GGT GAC ATC TCC ACT TTG GAA AAC-3' and 5'-GTT TTC CAA AGT GGA GAT GTC ACC ATG TAG TGT GG-3'. The construct was transformed into *E. coli* strain DH5 α . The mutations were verified by DNA sequencing. The mutant DNA constructs were transformed into *E. coli* strain BL21(DE3) pLysS.

Expression and Purification. yCD and E64A and E64D mutant enzymes were expressed using the same protocol implemented for the wild type protein, as described elsewhere (17). Three liters of *E. coli* cells in LB medium were grown for each protein. The cells were harvested by centrifugation, washed once with 50 mM phosphate buffer (pH 8.0) and kept at -20 °C until they were used.

The E64A mutant protein without the His tag was purified using the same protocol for the wild type yCD protein as described elsewhere (17). The His-tagged yCD and E64A and E64D mutant enzymes were purified as described below. The frozen bacterial paste was thawed at room temperature and suspended in 100 mL of precooled buffer A [50 mM potassium phosphate buffer with 300 mM NaCl (pH 8.0)]. DNase I and MgCl₂ were added to reach a final concentration of 1 µg/ml and 20 mM, respectively. The cells were disrupted with a French press. The resulting lysate was centrifuged for 20 min at ~27,000g and 4 °C. Polyethyleneimine was added to the pooled supernatant to a final concentration of 0.1% to precipitate RNA. The solution was centrifuged immediately after mixing at 15,000g for 30 min. The supernatant was loaded onto a nickel-nitriloacetic acid (Ni-NTA) column equilibrated with 10 mM imidazole in buffer A. The column was washed with 10 mM imidazole in the same buffer until the OD₂₈₀ of the effluent was <0.05. The column was eluted with a linear imidazole gradient (from 10 mM to 250 mM) in buffer A. Fractions containing the yCD protein were identified by SDS-PAGE and concentrated to ~15 mL with an Amicon concentrator using an YM10 membrane. The protein solution was then subjected to TEV protease digestion for 4 h at room temperature in the presence of 0.5 mM EDTA and 1 mM DTT. The digestion mixture was dialyzed against buffer A and loaded onto another Ni-NTA column

equilibrated with 10 mM imidazole in buffer A. The column was eluted with the same buffer. The yCD protein was washed out because it cannot bind with the resin after the His tag was removed. The flow-through fraction was collected and dialyzed against 2 mM phosphate buffer, pH 8.0 and lyophilized. The powder was kept at -80 °C until it was used.

Kinetic Measurements. Steady-state kinetic parameters were measured by a radiometric assay or an HPLC assay. The reaction was set up in a total volume of 40 μ l in 50 mM phosphate buffer with 50 mM NaCl at pH 7.5. The final enzyme concentration was 1 μ M, 500 μ M, and 34 μ M for the wild-type, E64A, and E64D yCDs, respectively. 5FC concentrations were 0.08 - 0.68 mM for the wild-type yCD, 10-100 mM for E64A yCD, and 6-64 mM for E64D yCD, respectively. A trace amount of [$6\text{-}^3\text{H}$]-5FC was used to quantify the conversion of 5FC to 5FU. The reactions were initiated at 20 °C by the addition of the enzymes. The reaction time was 10 s, 30 h and 8 h for the wild type, E64A and E64D yCDs, respectively. The reactions were quenched and 5FC and 5FU were separated by TLC as previously described (17). Kinetic parameters were obtained by nonlinear least squares fitting of the data to the standard Michaelis-Menten equation. The kinetic parameters of the E64A enzyme were also measured by an HPLC assay. In the HPLC assay, the reactions were quenched by the addition of an equal volume of 0.1 N HCl. The quenched solutions were injected onto a reverse-phase C18 column (22 mm \times 250 mm) and the column was eluted with 20 mM phosphate buffer, pH 2.0. The elution was monitored by UV absorption at 256 nm. The retention times for 5FC and 5FU were 3.6 and 5.8 min, respectively. The peak area of 5FU was integrated and compared with standard samples of known concentrations.

Fluorometric Measurements. The dissociation constant for the binding of 5FPy to the wild-type yCD was measured by fluorometry using a FluoroMax-2 fluorometer. A series of 5FPy solutions of different concentrations were made in 50 mM phosphate buffer, pH 7.5. A yCD stock solution in the same buffer was added to these 5FPy solutions and the mixtures were incubated at room temperature for 10 min in order for them to reach equilibrium. yCD and 5FPy concentrations were 1 μ M and 0 - 6.4 μ M, respectively. Fluorescence intensities were measured at an emission wavelength of 386 nm with a slit of 4 nm. The excitation wavelength and slit were 320 nm and 4 nm, respectively. A set of controls was obtained in the same manner by adding the buffer instead of the protein stock solution. The K_d value was obtained by nonlinear least-squares fitting of the difference data to the following equation:

$$\Delta F = \frac{df \times (P_t + L_t + K_d - \sqrt{(P_t + L_t + K_d)^2 - 4P_t L_t})}{2P_t} \quad (5-1)$$

where ΔF is the fluorescence change upon ligand binding, df is the fluorescence constant difference between the bound form and the free form ligand, P_t and L_t are the total concentration of the enzyme and 5FPy, respectively and K_d is the dissociation constant of 5FPy with the yCD enzyme.

Inhibition Assay. The inhibition of the 5FC deamination reaction by 5FPy was measured by a radiometric assay as described earlier. IC_{50} values were obtained by nonlinear least squares fitting of the inhibition assay data to the equation below:

$$v = k_{\min} + \frac{k_{\max} - k_{\min}}{1 + \frac{[I]}{IC_{50}}} \quad (5-2)$$

where v is the initial reaction rate, k_{\min} the minimum rate, k_{\max} the maximum rate, and $[I]$ the concentration of the inhibitor. K_i values were calculated by the Cheng-Prusoff equation (18):

$$K_i = \frac{IC_{50}}{1 + [S]/K_m} \quad (5-3)$$

where $[S]$ is the concentration of the substrate 5FC.

¹⁹F NMR Spectroscopy. NMR samples were prepared by dissolving the wild-type yCD or the E64A enzyme in 100 mM potassium phosphate buffer (pH 7.5) made in D₂O. An aliquot of a 5FPy stock solution was added to the protein solutions and the mixtures were incubated for 10 min to reach equilibrium. ¹⁹F NMR experiments were performed at 10 or 20 °C on a Varian Inova 600 MHz NMR spectrometer. NMR spectra were acquired with a spectral width of 10,000 Hz, 20,000 complex data points, and 1,024 transients and processed with a 10 Hz of line broadening. Relaxation delay was 6 s. The ¹⁹F chemical shifts were referenced to CF₃C₆H₅, which was set to -63.73 ppm.

ONIOM Calculations. The ONIOM method (19) as implemented in the Gaussian program (Gaussian 03) (20) was used to investigate the effects of the E64A substitution on yCD catalysis. The molecular set up and protocol for the ONIOM calculations were essentially the same as previously described (16). Briefly, the molecular system for the calculations was based on the 1.14-Å crystal structure of the

yCD-DHP complex (PDBID: 1p6o) (10). The enzyme is represented by the catalytic Zn and the active site and neighboring residues, including Ile33, Asn51, Thr60, Leu61, His62, Gly63, Glu64, Ile65, Leu88, Ser89, Pro90, Cys91, Asp92, Met93, Cys94, Thr95, Phe114, Trp152, Phe153, Glu154, Asp155, and Ile156. The system was divided into two layers. The inner layer is composed of the sidechains of residues Glu64, His62, Cys91, Cys94, Asn51, Asp155, zinc and the zinc-bound water or substrate and its derivatives. The outer layer composed of the rest of the molecular system. The inner layer was treated by the density functional method B3LYP and the 6-31G** basis set and the outer layer by the semi-empirical PM3 method. The total ONIOM energy E^{ONIOM} is given by (19, 21):

$$E^{\text{ONIOM}} = E(\text{high, model}) + E(\text{low, real}) - E(\text{low, model}) \quad (5-4)$$

where $E(\text{high, model})$ and $E(\text{low, model})$ are the energies of the inner layer at the high level of theory (B3LYP and the 6-31G** basis set) and at the low level of theory (PM3), respectively, and $E(\text{low, real})$ is the energy of the entire system at the level of PM3.

RESULTS

Biochemical Analysis. To investigate the role of Glu64 in the activation of the anticancer prodrug 5FC, we replaced Glu64 with alanine and aspartate. Preliminary biochemical characterization of the mutant enzymes without a His tag indicated that the purified mutant enzymes contained a minute amount of *E. coli* CD, even though the *E. coli* enzyme is not homologous to the yeast enzyme, and differs in both size and quaternary structure. To get rid of the *E. coli* CD contamination, genes encoding the

wild-type yCD and mutant proteins were sub-cloned into a vector for the production of His-tagged proteins. The His-tagged proteins were purified with a Ni-NTA column and the His-tag was removed by TEV protease digestion. Our analysis was focused on the effects of the E64A substitution, which completely removed the carboxyl functional group. The steady-state kinetic parameters of the wild-type yCD and the E64A and E64D enzymes are summarized in Table 5-1. Because of the relatively large errors of the radiometric assay in measuring a small amount of 5FU in the presence of a large amount of 5FC, the kinetic parameters of the E64A mutant enzyme were also determined by HPLC. Both substitutions caused a dramatic decrease in k_{cat} and a dramatic increase in K_{m} . The E64A substitution caused a decrease in k_{cat} by about 5 orders of magnitude and an increase in K_{m} by about 2 orders of magnitude, resulting in a decrease in $k_{\text{cat}}/K_{\text{m}}$ by about 8 orders of magnitude. The effects of the E64D substitution were only slightly milder. The kinetic results indicated that Glu64 is involved in both binding and catalysis in the activation of 5FC by yCD, and the longer side-chain is important, as Glu64 cannot be replaced with aspartate with the same functional group but a shorter side-chain.

To further investigate the role of Glu64 in the activation of 5FC, we measured the binding and inhibitory activities of the putative transition state analogue 5FPy. The results are summarized in Table 5-2. The inhibitory activity of this compound on yCD was measured for the E64D mutant enzyme but not for E64A protein, because the activity of the latter mutant enzyme was too low for such an analysis. The E64D substitution caused a decrease in K_{i} by about 4 orders of magnitude, indicating that the mutant enzyme has a much lower affinity for 5FPy than the wild-type yCD. The K_{d} value

for the binding of 5FPy could be measured by fluorometric titration for the wild-type yCD but not for either of the two mutant enzymes, because the method is not suitable for the weak binding of 5FPy to the mutant enzymes. The K_d value for the binding of 5FPy to the E64A mutant enzyme was estimated by NMR as described below.

^{19}F NMR Analysis. The bound Py in the crystal structure of yCD is in the hydrated form DHP (10, 11). 5FPy differs from Py in having a fluorine atom at position 5 instead of a hydrogen. In order to determine whether the bound 5FPy is hydrated like Py and to examine the role of Glu64 on the hydration, ^{19}F -NMR spectra were acquired for 5FPy in absence of and the presence of the wild-type yCD or the E64A mutant enzyme (Figure 5-1). In the absence of the enzymes, 5FPy showed a sharp peak at -160.6 ppm. In the presence of a lower concentration of the wild-type yCD, 5FPy showed an additional broad peak due to the bound 5FPy at around -161.1 ppm, upfield of the peak of free 5FPy. In the presence of the same concentration of the E64A mutant enzyme, 5FPy also showed a broad but much smaller peak due to the bound 5FPy at -158.1 ppm, downfield of the peak of free 5FPy. Using model compounds, it has been shown that the ^{19}F NMR signal is a good indicator of the hydration of 5FPy (22). The hydration of 5FPy results in a change in the bonding of C4 from *sp*² to *sp*³, causing a ~1 ppm upfield shift of the ^{19}F NMR signal. This upfield shift of the ^{19}F NMR signal has been used for the determination of the hydration state of 5-fluoropyrimidin-2-one ribonucleoside (22). The dramatic opposite shift of the ^{19}F NMR signal of 5FPy in Figure 5-1 indicates that 5FPy

is hydrated when it binds to the wild-type yCD and that it is not hydrated when it binds to the E64A mutant enzyme. In either case, however, free 5FPy is in slow exchange with the bound 5FPy on the NMR chemical shift timescale. In such an exchange regime, chemical shift does not change with protein or ligand concentration and cannot be used for the estimation of K_d values. However, the relative peak intensities of the NMR signals of the free form and the bound form can be used to estimate the K_d value if the protein and ligand concentrations are the same order of magnitude as the K_d value. The protein and 5FPy concentrations were suitable for the estimation of the K_d value for the E64A mutant enzyme but not for the wild-type yCD. The K_d value for the binding of 5FPy to the E64A mutant enzyme was estimated to be ~10 mM based on the ^{19}F NMR data, about 5 orders of magnitude higher than that for the binding of 5FPy to the wild-type enzyme. As 5FPy is the hydrated form when it binds to the wild-type yCD and the unhydrated form is favored over the hydrated form by a factor of 2.9×10^3 , the K_d for the binding of the hydrated 5FPy, a transition state analogue, to the wild-type CD was estimated to be 2.3×10^{-10} M.

Computational Analysis. Because 5FPy is hydrated when bound to the wild-type yCD, but not hydrated when bound to the E64A mutant enzyme, we investigated the hydration by yCD by ONIOM calculations taking advantage of the availability of the high resolution crystal structures of the enzyme in complex with DHP, the hydrated form of Py. Py was docked into the active site by superposition with DHP in the X-ray crystal

structure (10). After optimization, Py shifted slightly from the initial position to give Complex **1** (Figure 5-2 and Figure 5-3). Complex **1** is stabilized through a hydrogen bond network and a π -stacking interaction. One new interaction is present between the zinc-coordinated water O^{Zn} and N3 of Py. Glu64 is in the deprotonated form and forms a strong hydrogen bond with the other hydrogen of the zinc-coordinated water. The carbonyl oxygen (O2) of Py is hydrogen bonded to the side chain amide of Asn51 and the backbone amide of Gly63. The amide group at position 1 (N1) of Py is hydrogen bonded to the carboxyl group of Asp155. The latter is also hydrogen bonded to the side chain amide of Asn51. The hydrogen bond between the amide group ($N^{\epsilon 2}$) of His62 and the backbone carbonyl oxygen ($O^{\delta 1}$) of Asp155 helps in maintaining appropriate orientation of the imidazole ring of His62 that stacks with the cytosine ring. Before the hydration occurs, the water is converted to a hydroxide with the assistance of Glu64. First, the hydrogen bond between O^{Zn} and N3 is broken and a new hydrogen bond between O^{Zn} and Glu64 is formed (not the one already present). The new conformation is given in Complex **2** (Figures 5-2 and 5-3). In this complex, the hydrogen bond $O^{Zn} \dots N3$ is not completely disrupted. One important feature of this conformation is that the distance between O^{Zn} and C4 is shortened because the change of the H^{OZn} position reduces the steric hindrance. Complex **2** is calculated to be less stable than **1** by 2.9 kcal/mol (Figure 5-4). However, **2** is more stable than **1** by 5.4 kcal/mol at the E(high, model) energy level. The difference of 8.3 kcal/mol indicates the surrounding protein destabilizes the complex. The corresponding barrier is 3.2 kcal/mol and there is no barrier found at the

E(high, model) energy surface. Second, H^{OZn} is transferred from O^{Zn} to Glu64, as seen in Complex **3** (Figures 5-2 and 5-3). In this process, the hydrogen bond between O^{Zn} and N3 is completely broken. After the proton transfer, a strong hydrogen bond is formed between the Zn-coordinated hydroxide and Glu64-OH. The distance between O^{Zn} and C4 is further reduced. Complex **3** is slightly more stable than Complex **2**. The energy barrier is calculated to be 5.6 kcal/mol at the ONIOM energy level. Again, there is no barrier found at the E(high, model) level. In the next step, the hydrogen bond between O^{Zn} and Glu64-OH is broken and a new hydrogen bond is formed between Glu64-OH and N3, which can be seen in Complex **4** (Figures 5-2 and 5-3). The distance between O^{Zn} and C4 is shortened to 2.17 Å. The energy barrier is 0.9 kcal/mol at the ONIOM energy level and 1.7 at the E(high, model) energy level. Complex **4** is more stable than the initial Complex **1** by 7.5 kcal/mol at the ONIOM energy level and by 11.8 kcal/mol at the E(high, model) energy level. The Zn-coordinated hydroxide in Complex **4** is well positioned for the nucleophilic attack on C4. The distance between the hydroxide and C4 decreases from 3.01 Å in Complex **1** to 2.66 Å in Complex **2**, to 2.44 Å in Complex **3**, and 2.17 Å in Complex **4**. The nucleophilic attack on C4 is concomitant with the proton transfer from Glu64-OH to N3, resulting in the formation of the tetrahedral product (DHP) as shown in Complex **5** (Figures 5-2 and 5-3). The energy barrier is 1.0 kcal/mol at the ONIOM level and 0.7 kcal/mol at the E(high, model) level. The bond lengths, bond angles, and dihedral angles of DHP in complex **5** are in close agreement with those of DHP in the crystal structure. The yCD-DHP complex is more stable than the initial yCD-Py complex by 23.2 kcal/mol at the ONIOM level and 30.9 kcal/mol at the E(high, model) level. The

difference of 7.7 kcal/mol indicates that the model system is relatively destabilized by the outer layer. **5** is considerably more stable than **1** because of the intrinsic stability of the model system, which suggests that the tight binding of the inhibitor is mainly due to the high affinity of DHP for the wild-type yCD.

To investigate the effects of the E64A substitution on the hydration, an E64A-Py complex was constructed in the same manner as for the wild-type yCD complex with Py (Complex **6** in Figures 5-5 and 5-6), assuming that the structure of the E64A mutant enzyme is the same as that of the wild-type yCD. The conformation of the optimized complex **6** is similar to that of the corresponding wild-type yCD complex. The Zn-coordinated water, however, forms only one hydrogen bond (with N3 of Py), and the other hydrogen bond in the yCD-Py complex is absent, as Glu64 is not present in the mutant enzyme complex. Because of the lack of Glu64, the Zn-coordinated water cannot be converted to hydroxide before the nucleophilic attack on C4. The distance between O^{Zn} and C4 is 3.08 Å, longer than that in the yCD-Py complex, making it even harder to attack C4. The nucleophilic attack of the Zn-coordinated water on C4 of Py is concomitant with the transfer of one of its protons to N3 of Py to form the tetrahedral product as in Complex **7** (Figures 5-5 and 5-6). The approximate geometry of transition state **TS67** was found by scanning the H^{OZn} -N3 and O^{Zn} -C4 distances from complex **6** to Complex **7**. The transition state was found where the O^{Zn} -C4 distance is 2.58 Å. The energy barrier is 14.4 kcal/mol at the ONIOM energy level and 36.1 kcal/mol at the E(high, model) level (Figure 5-7), indicating that the protein environment (the outer layer) compensates for the barrier by 21.7 kcal/mol. The energy barrier is much higher

than the corresponding energy barrier for the hydration catalyzed by the wild-type yCD. It is noted that the proton is transferred slightly ahead of the formation of the O^{Zn} -C4 bond. Complex **7** is calculated to be less stable than **6** by 1.7 kcal/mol at the ONIOM level, indicating that Py is predominantly in the unhydrated form.

To investigate further the effects of the E64A substitution on the deamination reaction, a complex of the E64A mutant enzyme with the substrate cytosine was constructed as for the construction of the wild-type enzyme-substrate complex (*16*), assuming again that the structure of the mutant enzyme is the same as that of the wild-type enzyme. A water molecule was added in place of the carboxyl group of Glu64. After optimization, cytosine shifted slightly from the initial position to give Complex **8** (Figures 5-8 and 5-9). The interactions between the substrate and enzyme in complex **8** are similar to those in the corresponding wild-type enzyme complex (*16*) except for the interaction involving the Zn-coordinated water and Glu64. In the wild-type enzyme complex, Glu64 forms a hydrogen bond with the Zn-coordinated water and another with the amino group of cytosine. In the mutant enzyme complex, the Zn-coordinated water forms a hydrogen bond with N3 of cytosine, and the amino group of cytosine forms a hydrogen bond with the additional water molecule. As the Zn-coordinated water is hydrogen bonded to N3 and cannot be deprotonated by the additional water molecule, the Zn-coordinated hydroxyl is generated by a direct proton transfer to N3. The energy barrier for the proton transfer is 7.7 kcal/mol at the ONIOM energy level and 13.6 kcal/mol at the E(high, model) level. But the hydroxide is not as well positioned for the nucleophilic attack on C4 as in the corresponding wild-type complex (*16*). The distance between O^{Zn} and C4 is only 2.24 Å in the wild-type enzyme complex but is 2.9 Å in the

mutant enzyme complex. The transition state is found when the distance between O^{Zn} and C4 is 2.48 Å. The energy barrier for the nucleophilic attack is calculated to be 11.2 kcal/mol at the ONIOM energy level and 22.1 kcal/mol at the E(high, model) level. For the wild-type yCD-catalyzed reaction, the formation of the tetrahedral intermediate is through three steps, in which the highest energy barrier is only 1.9 kcal/mol (16). The stabilization of the tetrahedral intermediate is also affected, because the energy level of the tetrahedral intermediate is 6.1 kcal/mol higher than that of the substrate complex in the mutant enzyme-catalyzed reaction. However the energy level of the same intermediate is 10.8 kcal/mol lower than that of the substrate complex during the wild-type yCD catalyzed reaction (16).

Once the tetrahedral intermediate is formed, another proton transfer from O^{Zn} to N4 is needed before the C4-N4 bond is cleaved (16). Two possible pathways were explored. In pathway one, the proton is transferred directly to N4 via a four-member ring transition state. In pathway two, the proton transfer is mediated by the additional water molecule. The water donates a proton to N4 and extracts a proton from O^{Zn} in a concerted manner. The C4-N4 bond distance increases to 1.64 Å after the proton is transferred to N4. The energy barrier in second pathway (12.1 kcal/mol at the ONIOM energy level) is dramatically lower than that in the first pathway (22.3 kcal/mol at the ONIOM energy level), indicating that the O^{Zn} proton is transferred to N4 via the second pathway. The energy barrier is high when compared with the wild-type yCD catalyzed deamination, in which the energy barrier is only 7.8 kcal/mol at the ONIOM energy level.

DISCUSSION

The results of the combined experimental and computational studies reported here indicate that Glu64 is important for both binding and chemical steps in the activation of the anticancer prodrug 5FC. Assuming that the activation of 5FC follows a mechanism similar to that of the deamination of cytosine, it involves the generation of a tetrahedral reaction intermediate from the substrate and the conversion of the intermediate to the product. A critical role of Glu64 in the chemical steps is supported by the results of the biochemical analysis of the mutant enzymes. First, the k_{cat} of the E64A mutant enzyme decreases by 5 orders of magnitude relative to that of the wild-type yCD. Our previous transient kinetic and NMR studies showed that product release is rate-limiting in the activation of 5FC by the wild-type yCD (17). Consequently, the k_{cat} of the wild-type enzyme is significantly smaller than the rate constant for the chemical transformation. On the other hand, the rate-limiting step is most likely the chemical transformation in the activation of 5FC by the mutant enzyme, because of its severely impaired catalytic apparatus. Therefore, the k_{cat} of the E64A mutant enzyme should be compared with the corresponding rate constant of the wild-type yCD for the chemical transformation, and thus the E64A mutation slows down the chemical transformation by 6 orders of magnitude. Glu64 contributes to the stabilization of the transition state of the activation of 5FC by ~ 9 kcal/mol. Second, inhibition and binding studies indicate that the K_i of the putative transition state analogue 5FPy for the inhibition of the E64D mutant enzyme increases by 3 orders of magnitude relative to that for the inhibition of the wild-type yCD and the K_d of 5FPy for the binding of the E64A mutant enzyme increases by 4 orders of

magnitude relative to that for the binding of the wild-type yCD, corresponding to ~6 kcal/mol.

A more precise role of Glu64 in the activation of 5FC is revealed by the NMR analysis of the binding of 5FPy to the wild-type and mutant enzymes and the computational analysis of the hydration of Py by the wild-type yCD and the deamination of cytosine by the E64A mutant enzyme. The NMR analysis shows that 5FPy is hydrated when bound to the wild-type yCD but is unhydrated when bound to the E64A mutant enzyme, suggesting that Glu64 plays a critical role in the formation and/or stabilization of the hydrated 5FPy. The ONION calculations of the hydration of Py indicate that Glu64 functions as a proton shuttle for the transfer of a proton from the Zn-coordinated water to N3 of Py and the hydration follows a sequential mechanism with N3 protonated first followed by the nucleophilic attack of C4 by the Zn-coordinated hydroxide. Furthermore, the complex of the hydrated Py is much more stable than that of the unhydrated Py, based on the dramatic difference in the ONION energies of the two complexes. Without the proton shuttle Glu64 in the E64A mutant enzyme, the hydration follows a concerted mechanism with a significantly higher energy barrier and the ONION energy of the complex of the hydrated Py is higher than that of the complex of the unhydrated Py. The results of the ONION calculations suggest that Glu64 facilitates the hydration of Py and stabilizes the hydrated Py. Since the hydration reaction mimics the formation of the reaction intermediate in the deamination reaction and 5FPy is predominantly in the unhydrated form in the active center of the E64A mutant enzyme, Glu64 may play a critical role in the formation of the reaction intermediate in the activation of the anticancer prodrug 5FC.

The ONION calculations of the deamination reaction catalyzed by the E64A mutant enzyme show that Glu64 is important for both the formation of the tetrahedral reaction intermediate from the substrate and the conversion of the reaction intermediate to the product. Without Glu64 shuttling protons, the barriers for both the formation of the tetrahedral reaction intermediate and its conversion to the product are much higher than those in the wild-type enzyme reaction. Furthermore, the ONION energy of the reaction intermediate complex is 10.8 kcal/mol lower than that of the substrate complex in the wild-type enzyme reaction but is 6.4 kcal/mol higher in the mutant enzyme reaction. The combination of these effects makes the mutant enzyme a very sluggish catalyst.

Glu64 is also important for the binding of the anticancer prodrug 5FC. The K_m of the E64A mutant enzyme for 5FC increases more than 200-fold in comparison with that of the wild-type enzyme. Because the rate-limiting step in the activation of 5FC by the wild-type yCD is product release, the K_m of the enzymatic reaction is a complex parameter, smaller than the K_d for the binding of 5FC (17). On the other hand, the rate-limiting step for the activation of 5FC by the E64A mutant enzyme must be the chemical transformation. Consequently, 5FC is in rapid equilibrium with its complex with the E64A mutant enzyme and thus the K_m and K_d values should be similar. Therefore, the K_m value of the mutant enzyme is a good measure of the affinity of the mutant enzyme for 5FC and should be compared with the K_d value of the wild-type enzyme for the binding of 5FC, which was determined to be 0.19 mM. The ratio of the mutant enzyme's K_m to the wild-type enzyme's K_d is ~170, indicating that Glu64 contributes to the

binding of 5FC by ~ 3 kcal/mol. The biochemical result is consistent with the complex of the enzyme with cytosine (16) built based on the crystal structure of the complex of yCD with the hydrated Py (10, 11). After energy minimization, Glu64 forms a hydrogen bond with the amino group of cytosine. Disruption of the hydrogen bond by mutagenesis is likely to decrease the affinity of the enzyme for the substrate.

yCD belongs to a family of purine/pyrimidine deaminases with an active center featuring a catalytic zinc coordinated with cysteine and histidine residues and a strictly conserved glutamate. Interestingly, yCD is not homologous to eCD, which has a very different size, a different quaternary structure, and an active center containing a ferrous iron ion coordinated with four histidine residues, an aspartate residue, and a water molecule. Among the family of purine/pyrimidine deaminases, yCD is most similar to *E. coli* cytidine deaminase (CDA), having a catalytic Zn of the same coordination chemistry and a superimposable glutamate (10, 11). However, the functional role of Glu64 of yCD is not the same as that of the corresponding residue Glu104 of *E. coli* CDA, based on our studies on yCD presented here and the published mutagenesis study of *E. coli* CDA (23). Both glutamate residues are critical for catalyzing their respective reactions. The E104A mutation of *E. coli* CDA decreases the k_{cat} of the enzyme by 8 orders of magnitude (23). The effect is more dramatic than that of the E64A mutation of yCD, which decreases the k_{cat} of yCD by 5 orders of magnitude and the rate constant for the chemical transformation by 6 orders of magnitude. Similarly, the fluorinated transition-state analogue 5-fluoro-pyrimidin-2-one riboside is in the hydrated form when bound to the wild-type *E. coli* CDA but is in the unhydrated form when bound to the E104A mutant

enzyme (22), suggesting that Glu104 also plays a critical role in the generation of the reaction intermediate in the deamination reaction catalyzed by *E. coli* CDA.

The roles of the two glutamate residues in substrate binding are different. Glu64 of γ CD contributes to the stability of the γ CD complex with 5FC by ~ 3 kcal/mol, as discussed earlier. In contrast, Glu104 of *E. coli* CDA destabilizes the substrate complex of the enzyme *E. coli* CDA (23). Thus, the K_m of the E104A mutant enzyme decreases by a factor of ~ 30 . Interestingly, the K_d for the binding of the fluorinated transition-state analogue 5-fluoro-pyrimidin-2-one riboside to the E104A mutant enzyme is even lower than that for its binding to the wild-type enzyme, although the fluorinated transition-state analogue is in the unhydrated form when bound to the mutant enzyme, suggesting that Glu104 of *E. coli* CDA dramatically destabilizes the complex of the unhydrated form of the analogue with the enzyme. It has been proposed that Glu104 of *E. coli* CDA decreases the energy barrier of the deamination reaction not only by stabilizing the transition state but also by destabilizing the enzyme-substrate complex (22). The destabilization of the enzyme-substrate complex is probably due to the unfavorable interaction between the negatively charged Glu104 and the lone pair of electrons of N3 of cytidine, which becomes a favorable interaction when N3 is protonated during the deamination reaction. *E. coli* CDA can afford this sacrifice of substrate binding energy for the reduction of the energy barrier of the deamination reaction, because the ribose moiety of cytidine contributes to the binding energy for the formation of the enzyme-substrate complex and consequently the enzyme is still able to maintain its K_m in the sub-millimolar range even with the unfavorable interaction between the carboxyl group of

Glu104 and N3 of cytidine. On the other hand, without a ribose moiety, cytosine needs a favorable interaction with the carboxyl group of Glu64 of yCD, i.e. a hydrogen bond between the amino group of cytosine and the carboxyl group of Glu64 as described earlier, to keep its K_m in the millimolar range. Without this favorable interaction, the K_m increases by more than an order of magnitude as shown this mutagenesis study and a CD with a K_m in the ten millimolar range would not be useful for pyrimidine salvage.

In conclusion, the combined experimental and computational studies show that Glu64 plays a critical role in the activation of the anticancer prodrug 5FC. It is important not only for the deamination reaction but also for the binding of the substrate. In the deamination reaction, Glu64 facilitates both the formation of the tetrahedral intermediate and its conversion to the product by shuttling protons. Furthermore, it stabilizes the tetrahedral intermediate and the hydrated form of the inhibitor 5FPy, which mimics the transition state and/or the reaction intermediate. In contrast to Glu104 of *E. coli* CDA, which destabilizes the enzyme-substrate complex, Glu64 of yCD is also important for substrate binding. Its contribution to substrate binding is of great physiological significance. Without Glu64, the K_m of the enzyme would be too high for pyrimidine salvage.

TABLES AND FIGURES

Table 5-1: Kinetic constants of the wild-type yCD and the E64 mutant enzymes for the activation of 5FC.

Protein	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}s^{-1}$)
WT	0.16 ± 0.01	17 ± 0.4	106
E64A	89 ± 30^a	$(1.3 \pm 0.3) \times 10^{-4}^a$	$1.4 \times 10^{-6}^a$
	33 ± 6^b	$(7.3 \pm 0.6) \times 10^{-5}^b$	$2.2 \times 10^{-6}^b$
E64D	19 ± 6	$(5.1 \pm 1.0) \times 10^{-3}$	2.6×10^{-4}

a: Determined by a radiometric assay.

b: Determined by HPLC.

Table 5-2: Inhibitory and binding properties of 5FPy.

Protein	K_i (μM)	K_d (μM)
WT	0.86 ± 0.16	0.65 ± 0.08^a
E64A		$1.0 \times 10^4^b$
E64D	1060 ± 60	

a: Measured by fluorometric titration.

b: Estimated by ^{19}F NMR.

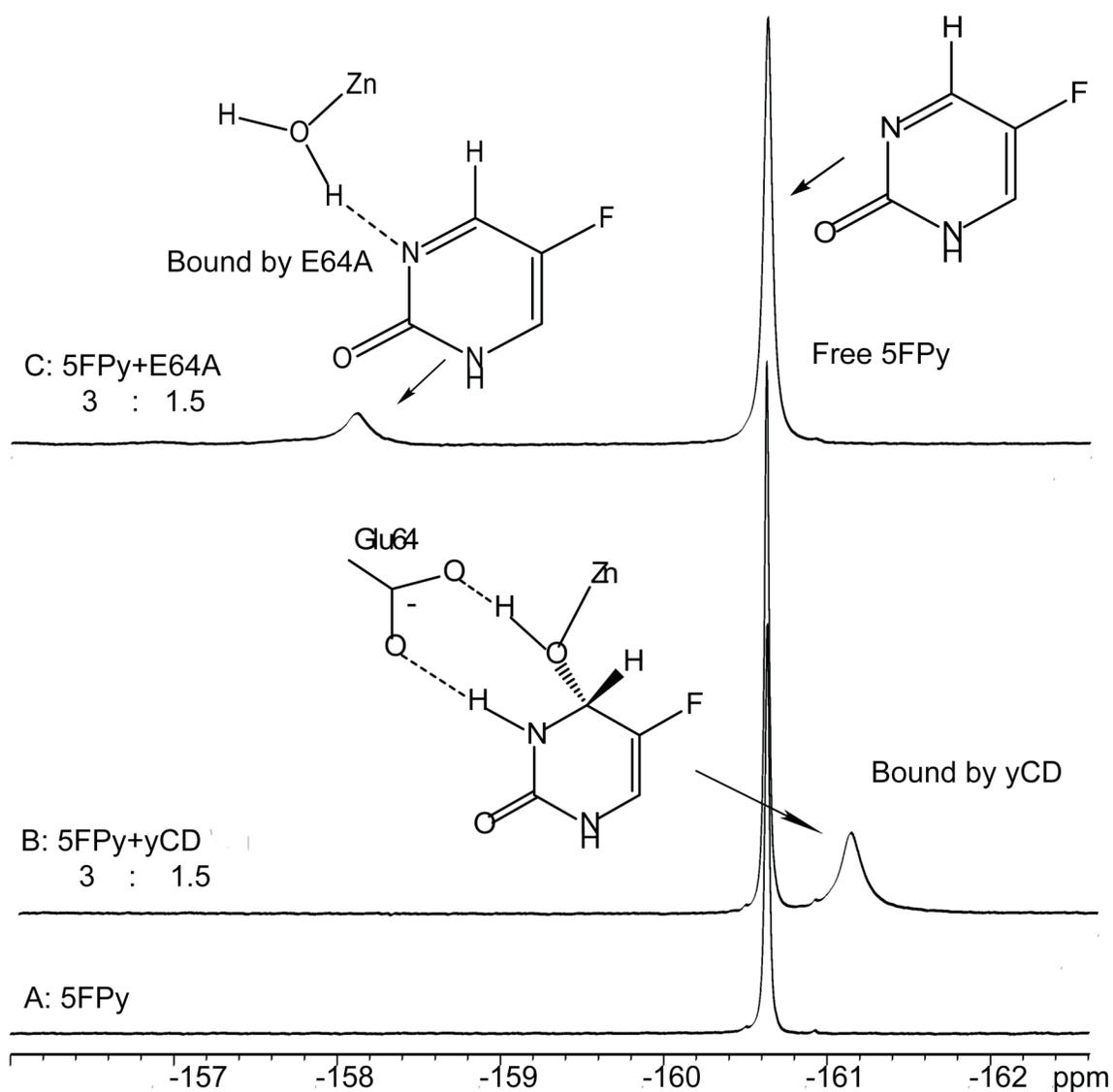


Figure 5-1. ^{19}F -NMR spectra of the E64A-5FPy complex and the wild-type yCD-5FPy complex. All samples were in a D_2O buffer containing 100 mM potassium phosphate, pH 7.5. (A) Spectrum of 5FPy (4.0 mM) in the absence of the enzyme; (B) spectrum of 5FPy (3.0 mM) in the presence of the wild-type yCD (1.5 mM); (C) spectrum of 5FPy (3.0 mM) in the presence of E64A (1.5 mM).

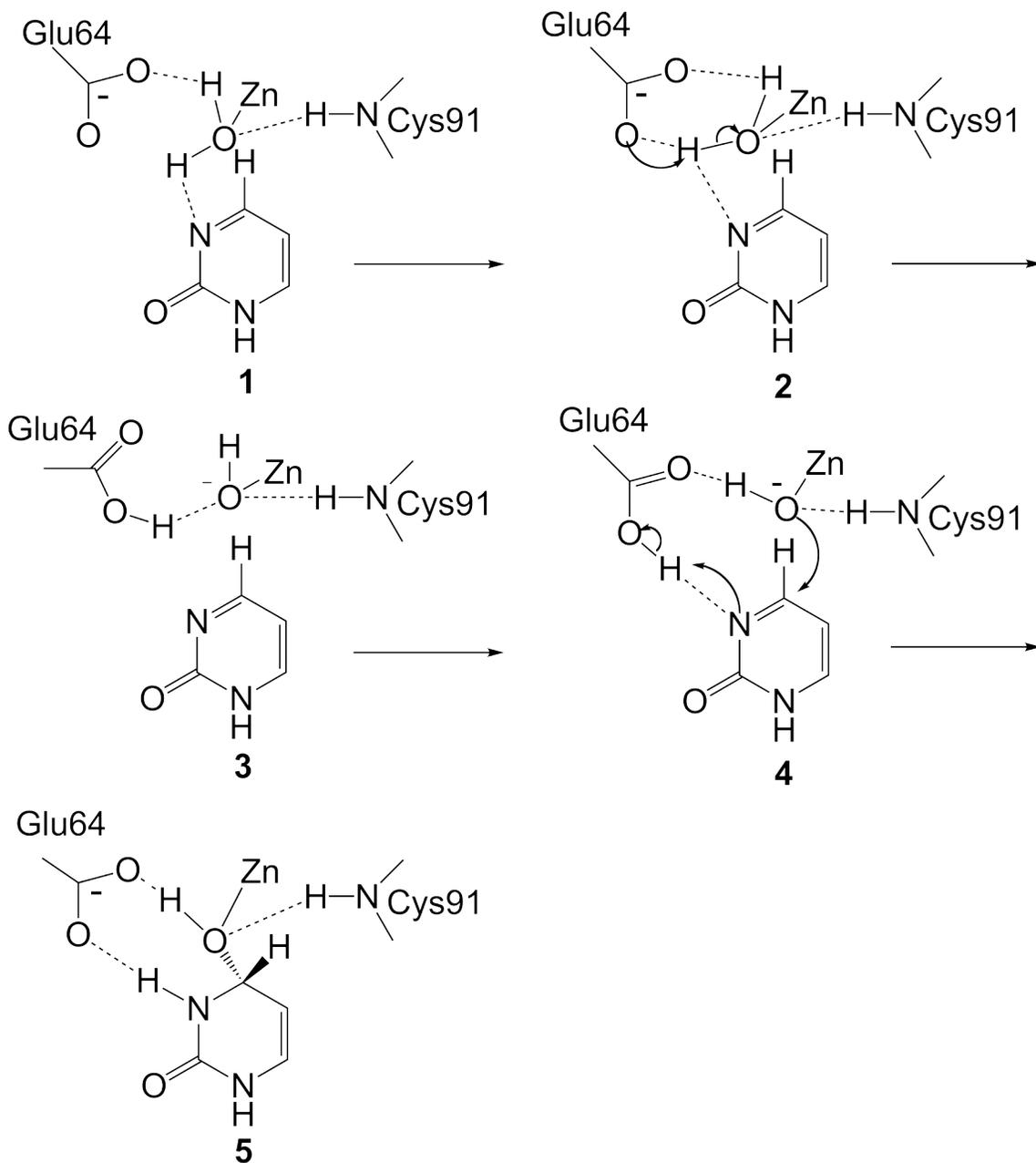


Figure 5-2. Proposed reaction pathway from Py to DHP catalyzed by the wild type yCD.

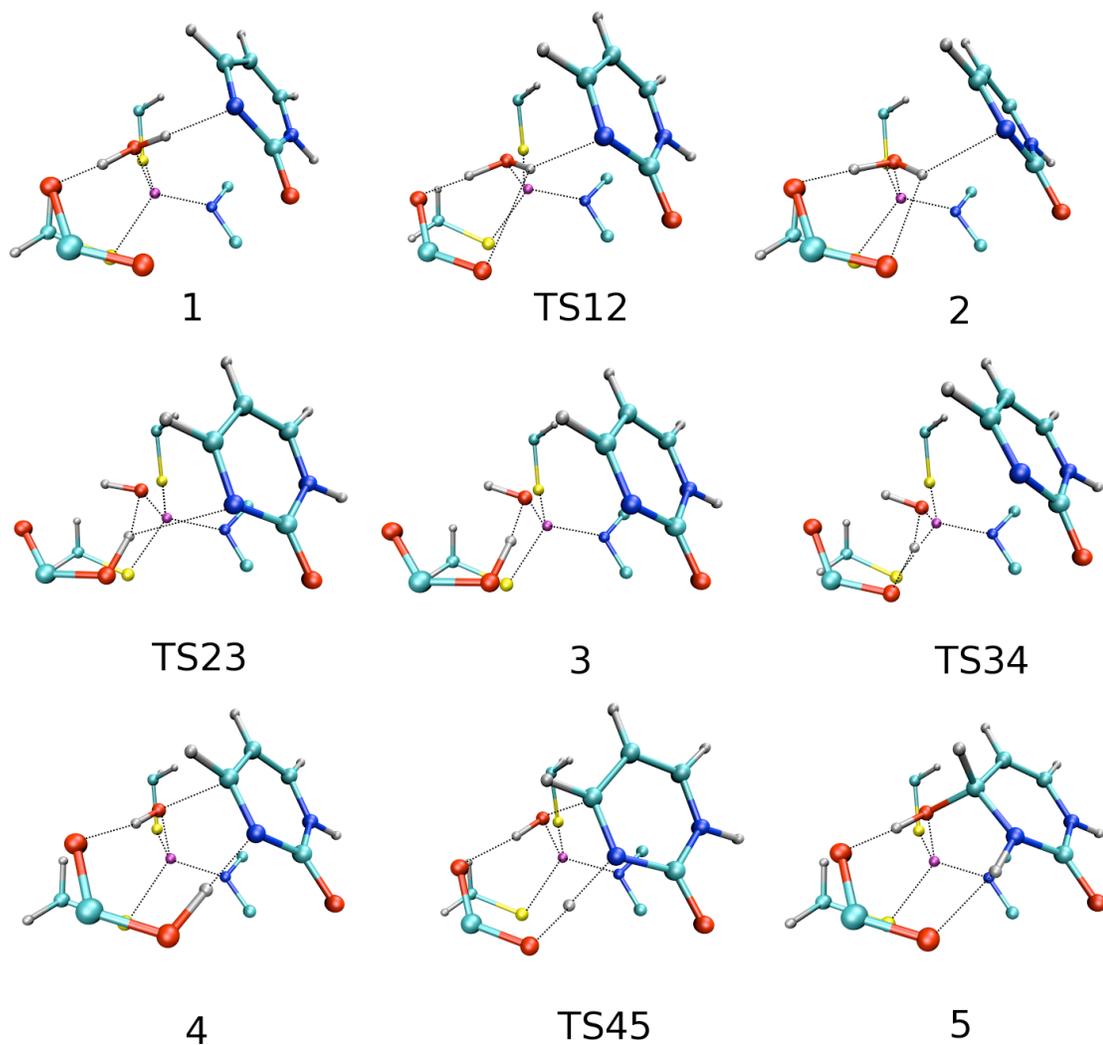


Figure 5-3. ONIOM optimized structures for the conversion of Py to DHP by the wild type yCD. (1) complex 1, (2) TS12 between complex 1 and 2, (3) complex 2, (4) TS23 between complex 2 and 3, (5) complex 3, (6) TS34 between complex 3 and 4, (7) complex 4, (8) TS45 between complex 4 and 5, (9) complex 5.

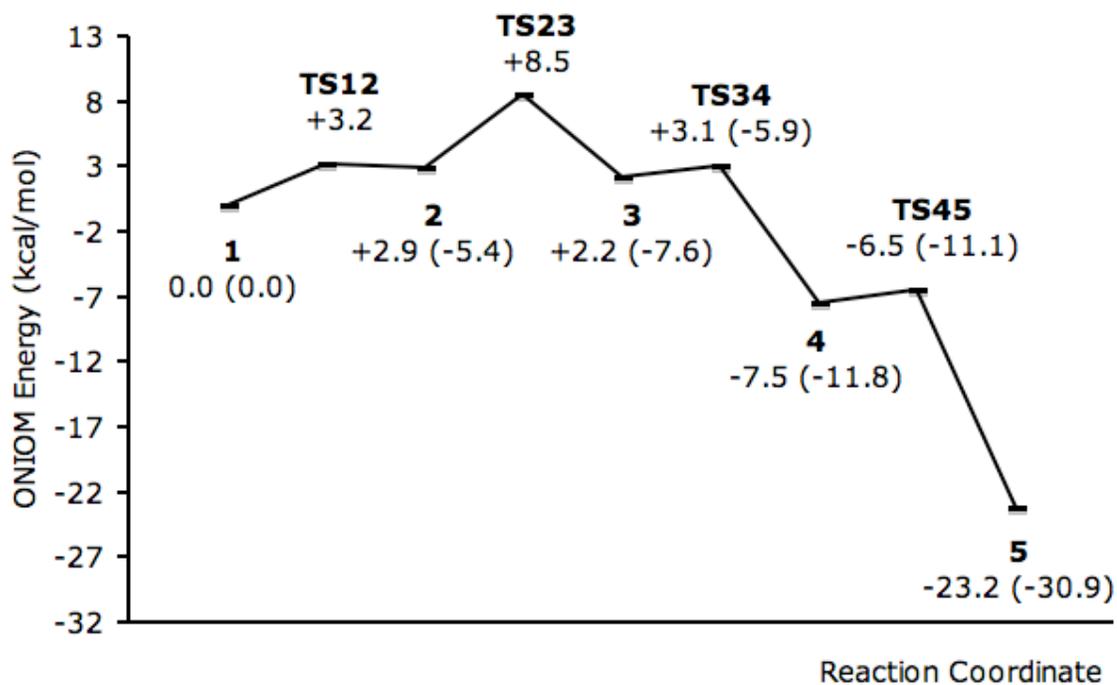


Figure 5-4. Schematic E^{ONIOM} and $E(\text{High,model})$ energy profile for the conversion of Py to DHP catalyzed by the wild-type yCD.

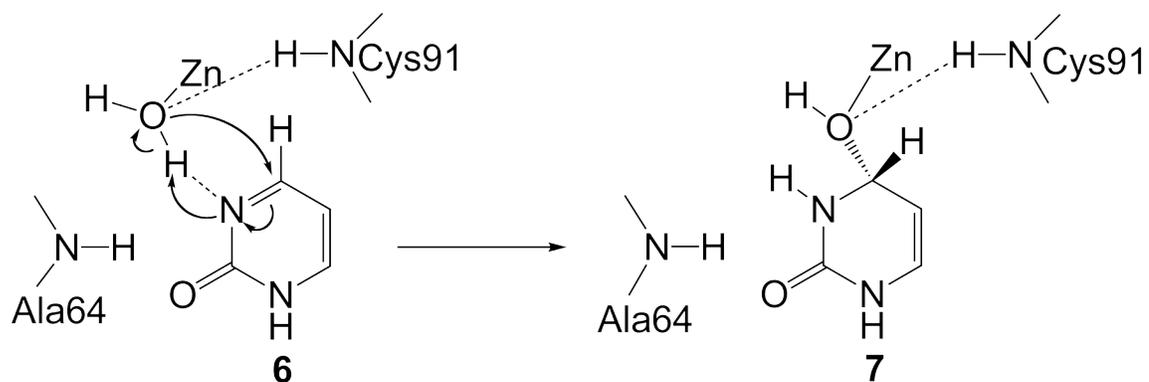


Figure 5-5. Proposed reaction pathway from Py to DHP catalyzed by the E64A mutant enzyme.

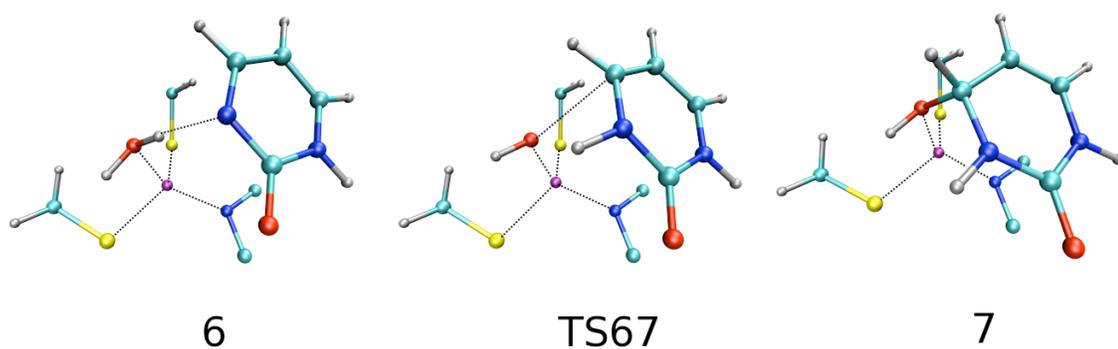


Figure 5-6. ONIOM optimized structures for the conversion of Py to DHP by the E64A mutant enzyme. (1) complex **6**, (2) **TS67** between complex **6** and **7**, (3) complex **7**.

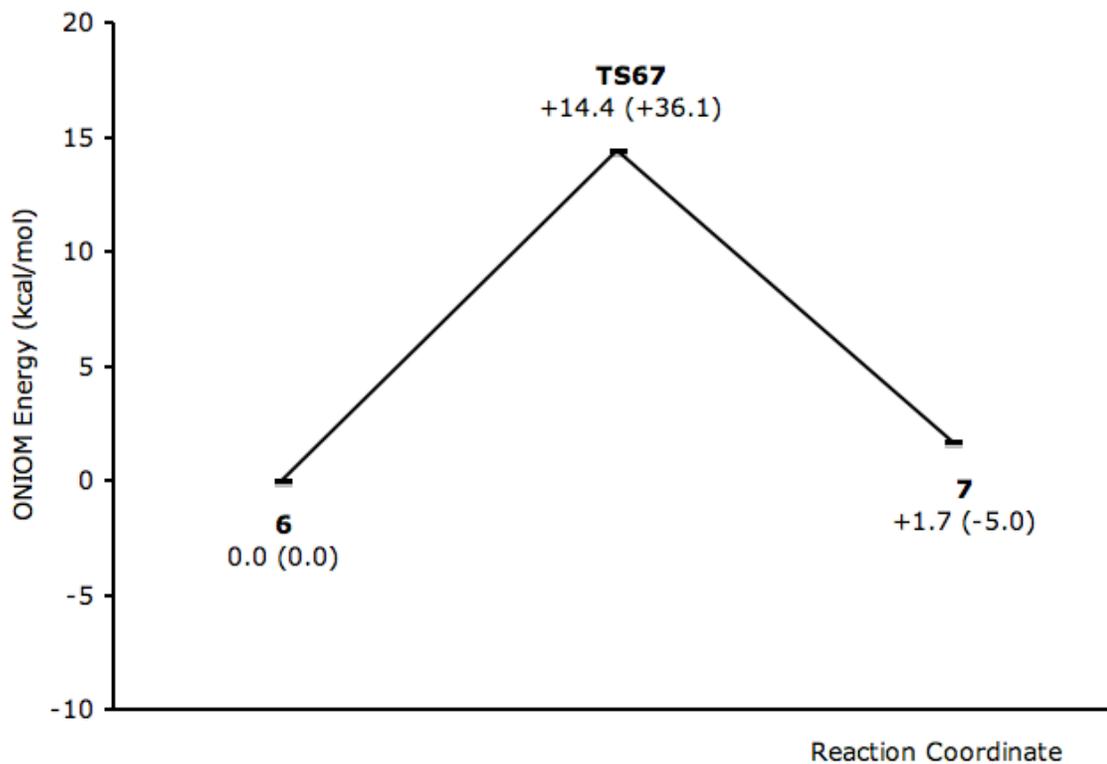


Figure 5-7. Schematic E^{ONIOM} and $E(\text{High,model})$ energy profile for the conversion from Py to DHP catalyzed by the E64A mutant enzyme.

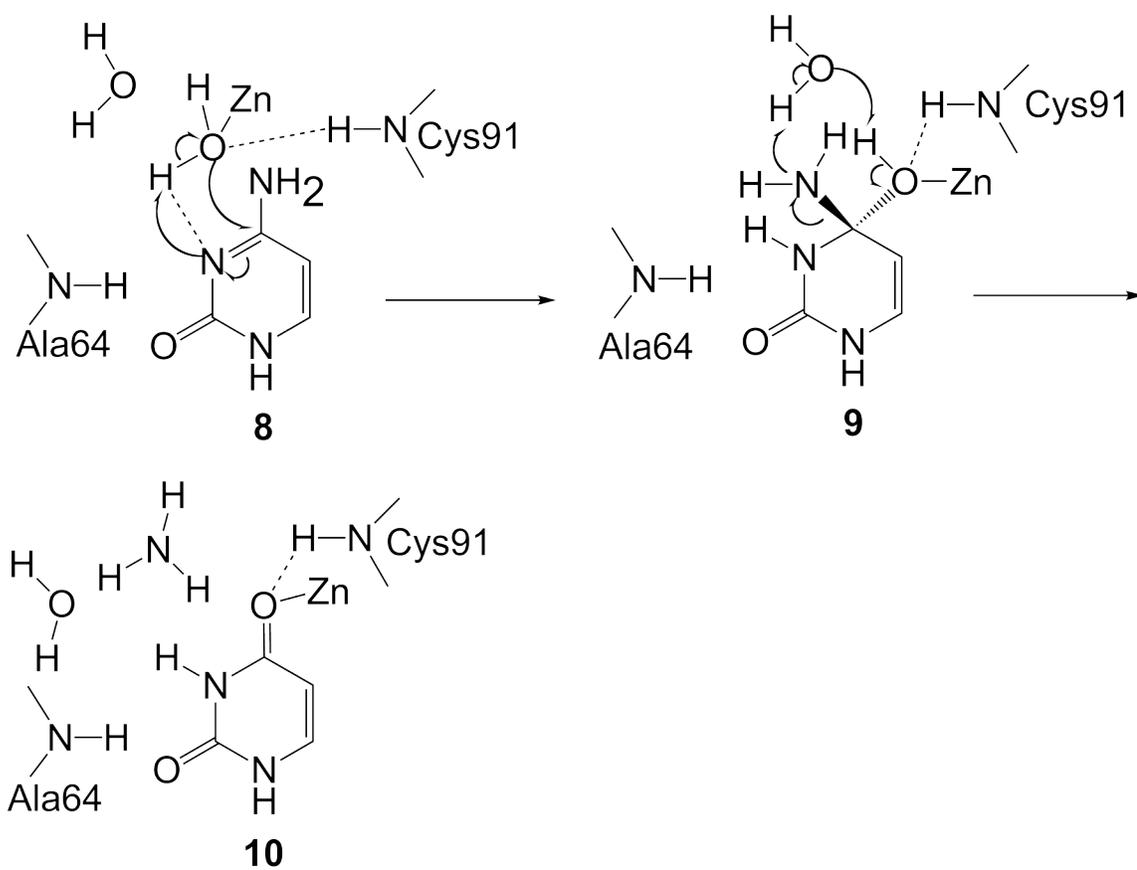


Figure 5-8. Conversion of cytosine to the zinc-coordinated uracil by the E64A mutant enzyme.

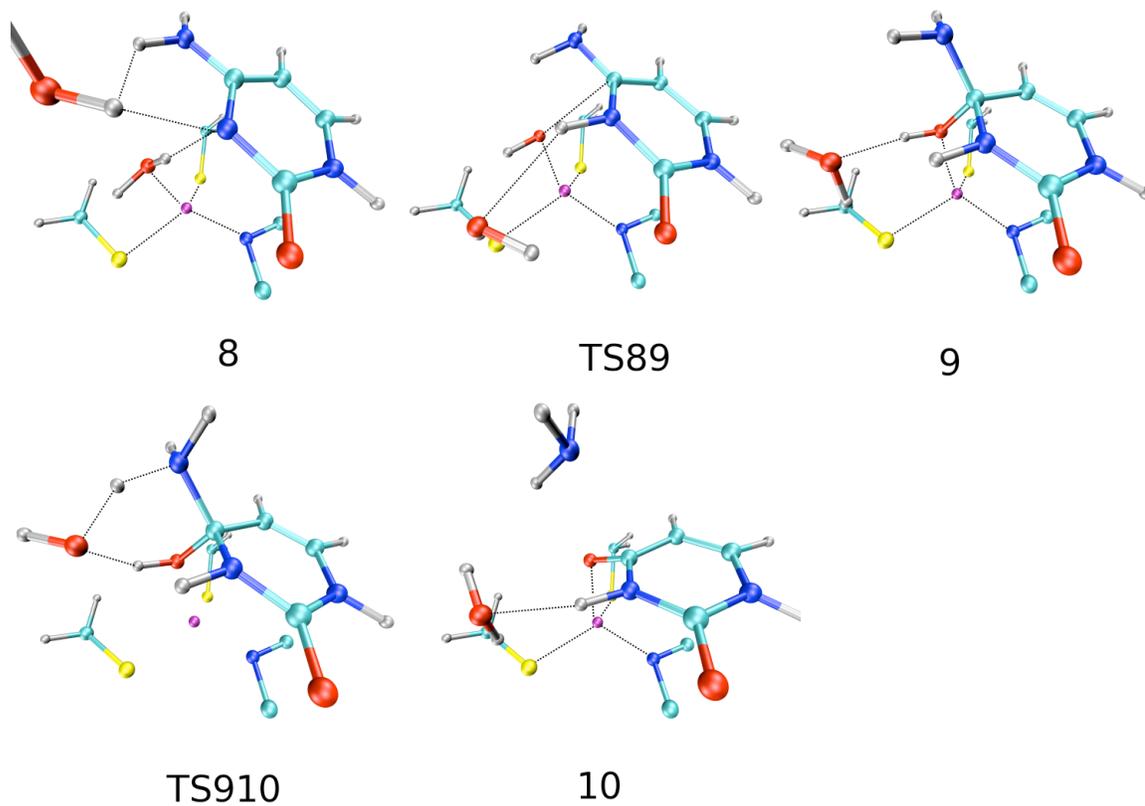


Figure 5-9. ONIOM optimized structures for the conversion from cytosine to the zinc-coordinated uracil. (1) complex **8**, (2) **TS89** between complex **8** and **9**, (3) complex **9**, (4) **TS910** between complex **9** and **10** and (5) complex **10**.

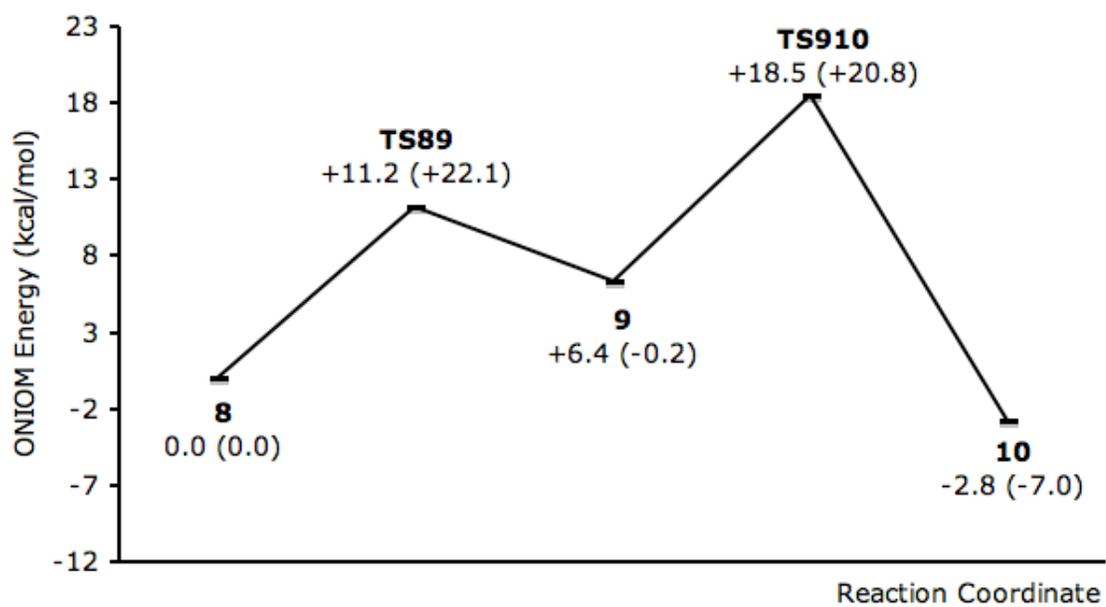


Figure 5-10. Schematic E^{ONIOM} and $E(\text{High,model})$ energy profile for the conversion from cytosine to uracil catalyzed by the E64A mutant enzyme.

REFERENCES

REFERENCES

1. Greco, O., and Dachs, G. U. (2001) Gene directed enzyme/prodrug therapy of cancer: historical appraisal and future perspectives, *J Cell Physiol* 187, 22-36.
2. Aghi, M., Hochberg, F., and Breakefield, X. (2000) Prodrug activation enzymes in cancer gene therapy, *J Gene Med* 2, 148-164.
3. Ireton, G. C., McDermott, G., Black, M. E., and Stoddard, B. L. (2002) The structure of *Escherichia coli* cytosine deaminase, *J Mol Biol* 315, 687-697.
4. Ipata, P., Marmocch, F., Magni, G., Feliciol, R., and Polidoro, G. (1971) Bakers yeast cytosine deaminase - Some enzymic properties and allosteric inhibition by nucleosides and nucleotides, *Biochemistry* 10, 4270-4276.
5. Hayden, M., Linsley, P., Wallace, A., Marquardt, H., and Kerr, D. (1998) Cloning, overexpression, and purification of cytosine deaminase from *Saccharomyces cerevisiae*, *Protein Expres Purif* 12, 173-184.
6. Hsu, Y., Hu, C., Lin, J., and Liaw, S. (2003) Crystallization and preliminary crystallographic analysis of yeast cytosine deaminase, *Acta Crystallogr D* 59, 950-952.
7. Kievit, E., Bershad, E., Ng, E., Sethna, P., Dev, I., Lawrence, T., and Rehemtulla, A. (1999) Superiority of yeast over bacterial cytosine deaminase for enzyme/prodrug gene therapy in colon cancer xenografts, *Cancer Res* 59, 1417-1421.
8. Kievit, E., Nyati, M., Ng, E., Stegman, L., Parsels, J., Ross, B., Rehemtulla, A., and Lawrence, T. (2000) Yeast cytosine deaminase improves radiosensitization and bystander effect by 5-fluorocytosine of human colorectal cancer xenografts, *Cancer Res* 60, 6649-6655.
9. Hamstra, D., Rice, D., Fahmy, S., Ross, B., and Rehemtulla, A. (1999) Enzyme/prodrug therapy for head and neck cancer using a catalytically superior cytosine deaminase, *Hum Gene Ther* 10, 1993-2003.
10. Ireton, G., Black, M., and Stoddard, B. (2003) The 1.14 angstrom crystal structure of yeast cytosine deaminase: Evolution of nucleotide salvage enzymes and implications for genetic chemotherapy, *Structure* 11, 961-972.
11. Ko, T., Lin, J., Hu, C., Hsu, Y., Wang, A., and Liaw, S. (2003) Crystal structure of yeast cytosine deaminase - Insights into enzyme mechanism and evolution, *J Biol Chem* 278, 19111-19117.
12. Xu, Q., Guo, H., Gorin, A., and Guo, H. (2007) Stabilization of a transition-state analogue at the active site of yeast cytosine deaminase: Importance of proton transfers, *J Phys Chem B* 111, 6501-6506.
13. Yao, L., Sklenak, S., Yan, H., and Cukier, R. (2005) A molecular dynamics exploration of the catalytic mechanism of yeast cytosine deaminase, *J Phys Chem B*

109, 7500-7510.

14. Yao, L., Yan, H., and Cukier, R. (2006) A combined ONIOM quantum chemical-molecular dynamics study of zinc-uracil bond breaking in yeast cytosine deaminase, *J Phys Chem B* 110, 26320-26326.
15. Yao, L., Yan, H., and Cukier, R. (2007) A molecular dynamics study of the ligand release path in yeast cytosine deaminase, *Biophys J* 92, 2301-2310.
16. Sklenak, S., Yao, L., Cukier, R., and Yan, H. (2004) Catalytic mechanism of yeast cytosine deaminase: An ONIOM computational study, *J Am Chem Soc* 126, 14879-14889.
17. Yao, L., Li, Y., Wu, Y., Liu, A., and Yan, H. (2005) Product release is rate-limiting in the activation of the prodrug 5-fluorocytosine by yeast cytosine deaminase, *Biochemistry* 44, 5940-5947.
18. Cheng, Y., and Prusoff, W. H. (1973) Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I_{50}) of an enzymatic reaction, *Biochem Pharmacol* 22, 3099-3108.
19. Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., and Morokuma, K. (1996) ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder Reactions and $\text{Pt}(\text{P}(t\text{-Bu})_3)_2 + \text{H}_2$ oxidative addition, *J Phys Chem* 100, 19357-19363.
20. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2004) Gaussian 03, Revision D.01. Gaussian, Inc., Wallingford CT.
21. Humbel, S., Sieber, S., and Morokuma, K. (1996) The IMOMO method: Integration of different levels of molecular orbital approximations for geometry optimization of large systems: Test for *n*-butane conformation and S_N2 reaction: $\text{RCl} + \text{Cl}^-$, *J Chem Phys* 105, 1959-1967.

22. Carlow, D. C., Short, S. A., and Wolfenden, R. (1996) Role of glutamate-104 in generating a transition state analogue inhibitor at the active site of cytidine deaminase, *Biochemistry* 35, 948-954.
23. Carlow, D. C., Smith, A. A., Yang, C. C., Short, S. A., and Wolfenden, R. (1995) Major contribution of a carboxymethyl group to transition-state stabilization by cytidine deaminase: mutation and rescue, *Biochemistry* 34, 4220-4224.

CHAPTER 6

CONTRIBUTIONS OF CONSERVED RESIDUES IN THE ACTIVE SITE TO SUBSTRATE BINDING AND CATALYSIS

INTRODUCTION

Yeast cytosine deaminase (yCD), an enzyme participating in the pyrimidine salvage pathway in fungi, catalyzes the deamination of cytosine to uracil. yCD is of great biomedical interest because it also catalyzes the conversion of the prodrug 5-fluorocytosine (5FC) to 5-fluorouracil (5FU), which is a widely used drug in the treatment of breast, rectal, stomach, colon, and pancreatic cancers (1, 2). 5FU has high gastrointestinal and hematological toxicities. The undesirable side effect is minimized (3) by the combination of yCD/5FC in gene-directed enzyme prodrug therapy (GDEPT) through which the administered nontoxic 5FC is converted to the toxic and lethal 5FU in the cancer cells by yCD. This strategy takes advantage of the fact that there is no cytosine deaminase expressed in mammalian cells (4). Thus, in principle, only the tumor cells are exposed to the toxic 5FU and killed, while the normal cells are not affected.

Crystal structures of yCD have been determined at high resolution, including the substrate-free form (PDB id: 1ox7) and the complex form (PDB id: 1p6o and 1uaq) with 2-pyrimidinone (Py), a transition state analog (4, 5). yCD is a homodimeric

metalloenzyme. Each monomer is tightly packed into a mixed α/β fold, which consists of a central five stranded β -sheet sandwiched by α -helices on both sides. A catalytic zinc atom is found in the active site, covalently bonded with His62 N $^{\delta 1}$, Cys91 S $^{\gamma}$, Cys94 S $^{\gamma}$ and a water molecule in the free form yCD. The inhibitor is bound in the hydrated form 4(*R*)-hydroxyl-3,4-dihydropyrimidine (DHP) and coordinates with the zinc atom through the 4-hydroxyl group. In both the free form and the complex form enzyme, the active site is completely covered by the same cluster of residues from the C-terminal α -helix and the loop between $\beta 4$ and αD and appears to be inaccessible to the substrate.

The yCD-DHP complex is stabilized by a hydrogen bond network. At one end of the pyrimidine ring of DHP, N1 forms a hydrogen bond with the carboxyl group (O $^{\delta 1}$) of Asp155 and O2 forms hydrogen bonds with the side chain amide (N $^{\delta 2}$) of Asn51 and with the backbone amide of Gly63, respectively. At the other end, the zinc-coordinated hydroxide (O4) is hydrogen bonded with the backbone amide group of Cys91. Glu64 is believed to be critical for the stabilization of DHP (see Chapter 5) through two hydrogen bonds, one with O4 and the other with N3. DHP is stabilized by hydrophobic interactions as well. A cluster of residues, including Phe114, Trp152, Ile156, Trp152, Ile33 and Leu88, protect DHP from being exposed to the solvent.

In the previous chapter, the function of Glu64 was investigated. The results show that Glu64 does not merely function in catalysis, but also contributes significantly to substrate binding. In the closely related enzyme *E. coli* cytidine deaminase, the counterpart of Glu64 seems not to affect substrate binding (6). Instead, the authors suggest that the glutamate may destabilize the substrate binding. In this chapter, I

continue investigating functions of conserved residues in the active site. A series of mutant proteins, including substitutions involving N51A, D155A, E154G, E154D, R73A and D151A, were constructed. Their functions were investigated by kinetic, NMR, and computational methods.

EXPERIMENTAL PROCEDURES

Materials. Cytosine, uracil, 5FC and 5FU were purchased from Sigma. [$6\text{-}^3\text{H}$]-5FC was purchased from Moravek. 5-fluoro-pyrimidinone was synthesized in our lab. Restriction enzymes and T4 ligase were purchased from New England Biolabs. *Pfu* DNA polymerase and the pET-17b vector were purchased from Stratagene and Novagen, respectively.

Site-Directed Mutagenesis and Protein Expression. The pET-17bHR and pET17bHR-yCD vectors were constructed as described in Chapter 5. The expression constructs encoding the His-tagged N51A, D155A, E154G, E154D, R73A, and D151A proteins with a TEV protease cleavage site were made by PCR-based site-directed mutagenesis using the construct pET17bHR-yCD. The forward and reverse primers for the N51A construct were 5'-CTC GGT CGT GGT CAC GCC ATG AGA TTT CAA AAG GG-3' and 5'-CCC TTT TGA AAT CTC ATG GCG TGA CCA CGA CCG AG-3', respectively. The forward and reverse primers for the D155A construct were 5'-CT CAG GAT TGG TTT GAA GCT ATT GGT GAG TAG AAT TCC-3' and 5'-GGA ATT CTA CTC ACC AAT AGC TTC AAA CCA ATC CTG AG-3', respectively. The forward and reverse primers for the E154G construct were 5'-GA CCT CAG GAT TGG TTT GGT GAT ATT GGT GAG TAG AAT TC-3' and 5'-GAA TTC TAC TCA CCA ATA TCA CCA AAC CAA TCC TGA GGT C-3', respectively. The forward and reverse

primers for the E154D construct were 5'-GA CCT CAG GAT TGG TTT GAT GAT ATT GGT GAG TAG AAT TC-3' and 5'-GAA TTC TAC TCA CCA ATA TCA TCA AAC CAA TCC TGA GGT C -3', respectively. The forward and reverse primers for the R73A construct were 5'-CT TTG GAA AAC TGT GGG GCA TTA GAG GGC AAA GTG TAC-3' and 5'-GTA CAC TTT GCC CTC TAA TGC CCC ACA GTT TTC CAA AG-3', respectively. The forward and reverse primers for the D151A construct were 5'-C GAT GAA AGA CCT CAG GCT TGG TTT GAA GAT ATT GG-3' and 5'-CCA ATA TCT TCA AAC CAA GCC TGA GGT CTT TCA TCG -3', respectively.

The constructs were transformed into *E. coli* strain DH5 α . The mutations were verified by DNA sequencing. The mutant DNA constructs were transformed into *E. coli* strain BL21(DE3) pLysS. The His-tagged mutant enzymes were purified as described in Chapter 5.

Kinetic Measurements. Steady-state kinetic parameters for 5FC were determined by radiometric assays. The reaction was set up in a total volume of 40 μ l in 50 mM phosphate buffer with 50 mM NaCl at pH 7.5. The final enzyme concentrations were 1 μ M, 40 μ M, 40 μ M, 5 μ M, 5 μ M, 5 μ M and 0.1 μ M for the wild-type, N51A, D155A, E154G, E154D, R73A, and D151A enzymes, respectively. 5FC concentrations were 0.08 - 0.68 mM for the wild-type enzyme, 2.0 - 24 mM for N51A, E154G, E154D and R73A enzyme, 2.0 -67 mM for D155A enzyme and 0.05 – 1.0 mM for D151A enzyme. A trace amount of [6-³H]-5FC was used to quantify the conversion of 5FC to 5FU. The reactions were initiated at 20 °C by the addition of the enzymes. The reaction times were 10 s, 80 min, 300 min, 60 s, 60 s, 60 s, and 30 s for the wild type, N51A, D155A, E154G, E154D, R73A and D151A enzymes, respectively. The reactions were quenched and 5FC and 5FU

were separated by TLC as previously described (7). Kinetic parameters were obtained by nonlinear least squares fitting of the data to the standard Michaelis-Menten equation.

Dissociation Constant of 5FPy. The dissociation constant of 5FPy with yCD was obtained by measuring the change in fluorescence intensity upon binding, using a FluoroMax-2. A series of 5FPy solutions at different concentrations were made by dissolving in 50 mM phosphate buffer, pH 7.5. yCD solution in the same buffer was added into these solutions and incubated at room temperature for 10 min in order to reach equilibrium. Fluorescence intensities were measured at an emission wavelength of 386 nm with a slit width of 4 nm. The excitation wavelength and slit width were 320 nm and 4 nm, respectively. A set of controls was obtained in the same manner by adding buffer instead of protein solution. The K_d value was obtained by nonlinear least-squares fitting. The same protocol was applied to the binding of 5FPy with the N51A and D155A mutant enzymes, however, no fluorescence change was observed.

^{19}F NMR Spectroscopy. NMR samples were prepared by dissolving the wild type yCD and the mutant enzymes in 100 mM potassium phosphate buffer (pH 7.5) made in D_2O . 5FPy was added to the protein solutions and incubated at room temperature for 10 min to reach equilibrium. ^{19}F NMR experiments were performed at 20 °C on a Varian Inova 600 MHz NMR spectrometer. The carrier frequency was 564.342 Hz. The NMR spectra were acquired with a spectral width of 10,000 Hz, 20,000 data points, exponential multiplication (10 Hz) and 1,024 transients for each spectrum. Relaxation delay was set as 2s. The ^{19}F chemical shifts were referenced to $\text{CF}_3\text{C}_6\text{H}_5$ (chemical shift of -63.73 ppm).

Molecular Dynamics Simulations. Molecular dynamics (MD) simulations were carried out on the yCD-Py complex, N51A yCD-Py complex, D155A yCD-Py complex, yCD cytosine complex, N51A yCD cytosine complex, and D155A yCD cytosine complex. The starting structure was taken from the 1.14 Å resolution crystal structure (PDB code: 1p6o) (5). The mutant enzymes were made by replacing Asn51 or Asp155 with Ala in InsightII. The ligand Py or cytosine was docked in by superimposing with DHP in the active site.

The procedure for the conventional MD simulation was described in Chapter 4. For each complex, 8 ns MD simulation trajectory was obtained and analyzed in the molecular mechanics Poisson Boltzmann surface area (MM-PBSA) calculations.

MM-PBSA Computations. The molecular mechanics Poisson Boltzmann surface area (MM-PBSA) technique (8, 9) was used to estimate the free energy changes of initial binding of cytosine or Py with the wild type, N51A, and D155A yCD enzymes. Since there are no chemical steps involved after the initial binding of the ligand in the MD simulation, the calculated binding free energy only reflects the initial binding of the ligand. According to the binding process



the initial binding free energy

$$\Delta G_{binding} = G(EL) - G(E) - G(L) \quad (6-1)$$

where $G(EL)$, $G(E)$ and $G(L)$ are the free energies for the initial complex, the receptor and the ligand, respectively. The single MD trajectory protocol was utilized, i.e. the snapshots of complex, the receptor and the ligand were extracted from an MD trajectory of the complex. Hence, the binding free energy was estimated according to

$$\Delta G_{binding} = \Delta E_{MM} + \Delta G_{sol} - T \Delta S_{MM} \quad (6-2)$$

where ΔE_{MM} is the difference of the molecular mechanical energies, including electrostatic and van der Waals interaction energies, since the internal energies are canceled out. ΔG_{sol} was estimated by summing up the contributions from the electrostatic solvation energy ΔG_{ele} and the nonpolar solvation energy ΔG_{np} . The electrostatic contribution was calculated by solving the Poisson-Boltzmann equations (10) with the grid spacing set at 0.5 Å and the internal and external dielectric constant set at 1.0 and 80.0 respectively. The nonpolar contribution (11) was calculated according to

$$\Delta G_{np} = \gamma SABA + b \quad (6-3)$$

where SABA is the solvent accessible surface area and γ was set at 0.00542 kcal/Å² and b was set at 0.92 kcal/mol. The entropy term was not considered in the calculation because it is computationally time-consuming. So, strictly the calculated free energy is not the true binding free energy. But in this study, because we are comparing the binding free energy between very similar systems, the entropy term should be close to ΔG and cancels out each other when $\Delta \Delta G$ is calculated.

QM/MM MD Simulations. Self-consistent charge density functional tight-binding (SCC-DFTB) method is a semi-empirical method developed by Elstner *et al* (12, 13), which has been applied on a number of enzymatic systems (14-18). In this study, QM/MM MD simulations were performed for yCD using the SCC-DFTB method implemented in the amber molecular dynamics package (15). The molecular systems for the yCD water/cytosine complex, N51A water/cytosine complex and D155A water/cytosine complex were built as described in Chapter 4. The active site of subunit A

was modeled in the QM region, including the side chains of His62, Glu64, Cys91 and Cys94 and the catalytic zinc atom, and the ligands for the respective complexes. The rest of the protein was modeled in the MM region. SCC-DFTB and PARM99SB force field was used for the QM and MM region, respectively. SHAKE was turned on for all bonds except those in the QM region to allow proton transfer.

The molecular systems were minimized and heated up to 300 K using the same procedure in the conventional MD, except that the residues in the QM region were calculated by the SCC-DFTB method. The structures heated up to 300 K were used as the starting structures for the equilibration and production MD simulations. As an alternative, representative structures from the conventional production MD simulations were tested as the starting structures. The results were similar.

During the equilibration and production MD simulation, temperature was regulated by Langevin dynamics with $\gamma = 1.0 \text{ ps}^{-1}$. Tight convergence on the density matrix in the SCF was used. The electrostatic cutoff was set at 12.0 Å for QM/MM electrostatic interactions. Long-range QM-QM and QM-MM electrostatic interactions were calculated by Ewald sum (17, 19) and a QM compatible PME approach (14). The time step was 1 fs. MD simulations were carried out using the parallel version Sander module in Amber 10. All of the MD results were analyzed with the Ptraj module of Amber 10.

The free energy profile along a reaction coordinate was obtained from the potential of mean force (PMF) by adopting an umbrella sampling technique (20). The proper amount of windows was used to sample enough space along the reaction coordinate. The force constant was set in the range of 100 to 200 kcal/mol. The data from the simulations were analyzed by the weighted histogram analysis method (WHAM) (21,

22) to compute the PMF. Data from a 200 ps trajectory were collected after 200 ps of equilibration for each window.

RESULTS

Steady State Kinetics. By site-directed mutagenesis, plasmids encoding the N51A, D155A, E154G, E154D, R73A and D151A proteins were constructed and the proteins were expressed and purified. Steady-state kinetic parameters of the conversion of 5FC to 5FU by the wild type yCD and the mutant enzymes were measured by initial velocity analysis as described in the methodology section.

Asn51 forms a hydrogen bond with O2 of DHP in the yCD DHP complex, indicating it may contribute to substrate binding. The results show that substituting Asn51 with Ala greatly reduces the catalytic efficiency. As shown in Table 6-1, the k_{cat}/K_m value of N51A yCD decreases about 1×10^7 -fold, relative to that of wild-type yCD, from $106 \text{ mM}^{-1} \text{ s}^{-1}$ to $1.1 \times 10^{-5} \text{ mM}^{-1} \text{ s}^{-1}$. The reduced activity is not only due to weakened affinity to the substrate 5FC ($K_m = 4.1 \text{ mM}$), but it is also a result of the low turnover rate ($k_{\text{cat}} = 4.6 \times 10^{-3} \text{ s}^{-1}$).

Similar behavior was observed for the D155 enzyme. Asp155 forms a hydrogen bond with N1 of DHP in the yCD DHP complex. Replacing Asp155 with Ala reduces k_{cat}/K_m about 2×10^7 -fold. The K_m of 5FC increases to 17 mM and the k_{cat} decreases to $9.8 \times 10^{-3} \text{ s}^{-1}$. The results suggest that both Asn51 and Asp155 are important for ligand binding. They also participate in catalysis, probably by stabilizing the transition state.

E154G, E154D, and R73A proteins were constructed to explore the function of the C-terminal α -helix. The C-terminal α -helix is suggested to serve as a “gate” controlling access to the active site. Glu154 in this region forms a strong salt bridge with Arg73 from the other subunit. Removal of the interaction may increase the flexibility of the C-terminal α -helix and hence increase the product release rate. Three mutant enzymes, with E154G, E154D, and R73A substitutions, were made and steady state kinetics experiments were performed. The results show that the values of K_m of all three enzymes increase, relative to that of the wild type enzyme. The value of K_m for E154G, E154D and R73A proteins is 4.2, 2.4 and 3.2 mM, respectively. This finding could be an indication of a more opened active site; however, the alterations may reduce the stability of the transition state. The k_{cat} of these enzymes decreased several fold to 8.7, 2.6, and 1.7 s^{-1} for E154G, E154D and R73A proteins, respectively.

Asp151 is also a residue in the C-terminal α -helix. The negatively charged side chain does not interact with the inhibitor in the yCD DHP complex. Rather, it points out of the active site and is likely surrounded by solvent molecules. The activity of the D151A enzyme is close to that of the wild type enzyme, with a similar K_m . The k_{cat} only decreases 2-fold to 6.3 s^{-1} .

Dissociation Constants of 5FPy. The dissociation constants of 5FPy with the wild type and the N51A and D155A enzymes were measured. It turned out that 5FPy bound yCD very tightly with an apparent K_d value of $0.65 \mu\text{M}$ (Table 5-2 in Chapter 5 or Table 6-2). However, the change in fluorescence signal upon ligand binding was not

observed in the binding assays for either mutant enzyme, suggesting that they have low binding affinity with 5FPy.

The dissociation constants of 5FPy with the mutant enzymes were estimated by comparing the peak area of 5FPy in the bound form and in the free form by ^{19}F NMR spectroscopy. ^{19}F -NMR spectra were acquired for the wild-type yCD (1.5 mM) in the presence of 3.0 mM 5FPy and the N51A and D155A enzymes (1.5 mM) in the presence of 20.0 mM 5FPy, as shown in Figure 6-1. All enzyme spectra had a peak as the free form inhibitor at around -160.6 ppm, because the 5FPy concentration was much higher than that of the enzyme. Both mutant enzyme spectra also had a broad peak corresponding to the bound inhibitor. The peaks were well separated because the exchange rate is slow on the NMR time scale. However, the K_d of 5FPy with these two mutant enzymes is harder to obtain because the intensity of the peak is so low, even when more 5FPy was added. These results indicate that both residues are critical for the binding of the inhibitor. This result is consistent with the steady state data in which K_m for both enzymes are increased due to the low binding affinity with 5FC.

Binding Free Energy Analysis by MM-PBSA/GBSA. To understand how Asn51 and Asp155 contribute to ligand binding, MM-PBSA/GBSA was performed to estimate the binding free energy of Py. Contributions from other conserved residues in the active site were analyzed as well.

An 8 ns production MD trajectory for the yCD Py complex simulation was collected. The stability of the system was analyzed by plotting the root-mean-square displacements (RMSDs) of the backbone $\text{C}\alpha$ atoms along the reaction coordinate,

relative to the starting structure (Figure 6-2). RMSDs show that the system was stable and finally reached the equilibrium state after 1.5 ns.

The interactions between Py and the surrounding residues include a set of hydrogen bonds and hydrophobic interactions. O2 of Py forms a hydrogen bond with N^{δ2} of Asn51 and with the backbone amide of Gly63 with distances of 2.88 and 3.14 Å, respectively. N1 of Py forms a hydrogen bond with O^{δ1} of Asp155 with a distance of 2.73 Å. O^{δ2} of Asp155 forms a hydrogen bond with N^{δ2} of Asn51 with a distance of 2.88 Å. N^{ε2} of His62 forms a hydrogen bond with the backbone carbonyl oxygen of Asp155 with a distance of 2.89 Å. One new interaction is between the zinc-coordinated water O^{Zn} and N3 of Py with a distance of 3.09 Å. Glu64 is in the deprotonated form and forms a strong hydrogen bond with the other hydrogen from the zinc-coordinated water with 2.60 Å. Residues such as Ile33, His62, Leu88, Phe114, Trp152 and Ile156 contribute to the interaction by hydrophobic effects.

100 snapshots in the MD trajectory between 5 ns and 8 ns, at 30 ps intervals, were extracted and analyzed. The data in Table 6-3 show the energy terms accounting for the binding free energy by the MM-PBSA technique. ΔE_{MM} represents the change of molecular mechanics interaction energy upon inhibitor binding, including electrostatic energies and van der Waals interaction energies. The negative value (-51.6 kcal/mol) suggests the binding is favored by the solutes in the gas phase. ΔG_{sol} represents the change of solvation energies, including polar and nonpolar energy. The positive value (33.2 kcal/mol) indicates that the binding is not favored due to the presence of the

solvent. The polar solvation energy is dominant compared with the nonpolar solvation energy. The overall binding free energy suggests the binding of Py with yCD is favored by 18.4 kcal/mol. The binding free energy was also calculated by MM-GBSA. The same conclusion can be drawn, though the value is more negative. Entropy was not considered because it is computationally time consuming and the results normally have large deviation.

Free Energy Decomposition. Per-residue free energy decomposition was performed to evaluate the contributions from individual residues. The energy contributions were plotted in Figure 6-3. A number of residues, including Val31, Ile33, Asn51, His62, Gly63, Leu88, Trp152, Asp155 and Ile156, contribute to the binding of Py.

Both Asn51 and Asp155 make a significant contribution (Table 6-4) to the total free energy. The individual energies are 2.7 and 0.7 kcal/mol for Asn51 and Asp155, respectively. The contribution is mainly from the gas phase energy of the side chain. The solvation effects are quite different for these two residues. The Asp155 side chain contributes to the solvation energy by 5.3 kcal/mol and hence the overall contribution from Asp155 decreases to 0.7 kcal/mol. Similarly the side chain of Glu64 contributes to solvation energy (1.47 kcal/mol). The presence of Glu64 is an unfavorable force for the inhibitor binding. Interestingly, the E64A enzyme does not increase the affinity to Py (see Chapter 5). Instead, the binding affinity is greatly reduced. Per-residue energy reflects the contribution of a given residue in the wild type protein complex with Py. When Glu64 is changed to Ala, the behavior of the mutant enzyme was quite different. MM-PBSA on the E64A-Py complex showed that the enzyme indeed loses affinity to Py (data not shown).

Gly63 forms a hydrogen bond with O2 of Py through the main chain amide, by which Gly63 contributes to the gas phase energy (-2.5 kcal/mol). Similarly, since the backbone NH is exposed to solvents in the active site, the solvation energy cost is relatively high too. The overall effects favor the inhibitor binding.

His62 interacts with Py on one side by π -stacking interaction and residues Val31, Ile33, Thr86, Leu88, Ser89, Pro90, Cys91, Trp152 and Ile156 interact on the other side by hydrophobic effects. His62 is far more important than the other nonpolar residues because of the strong π -stacking interaction. This residue contributes to the gas phase energy with 4.5 kcal/mol and a total binding free energy of 3.5 kcal/mol. The individual contribution from the other hydrophobic residues is small. However, the collective effects are not negligible, contributing roughly 3.9 kcal/mol to the total binding energy.

Binding Free Energy with Mutant Enzymes. Kinetics and binding assays suggest that the substitutions greatly reduce the binding affinity of 5FC/5FPy with the mutant enzymes. Energy decomposition analysis shows that Asn51 and Asp155 contribute to the total binding free energy with 2.7 kcal/mol and 0.7 kcal/mol, respectively. To get a better estimation of the free energy, MM-PBSA analysis was performed on the mutant enzyme Py complexes.

An 8 ns production MD trajectory was collected for each mutant enzyme using the same procedure as for the wild type yCD Py complex. Similarly, stabilities of the systems were analyzed by plotting the RMSDs for the backbone C α atoms in the MD trajectory of the production MD simulation, relative to the starting structure (Figure 6-2). RMSDs show that the systems were not as stable as the wild type complex but eventually reached equilibrium after 3 ns. For the MM-PBSA analysis, 100 configurations from the MD trajectory starting from 5ns to 8 ns, at 30 ps intervals, were extracted and analyzed.

Similar results were obtained when different time windows were used between 3 to 8 ns or more snapshots were analyzed (data not shown).

The results are summarized in Table 6-2. Both GB and PB methods were used to calculate the free energy difference $\Delta G_{\text{binding}}$. The value calculated from PBSA is a little larger than that from GBSA, but in both cases binding of the inhibitor is favored. The binding free energy of Py with the wild type yCD is calculated to be -18.4 kcal/mol. When Asp51 is replaced by Ala, ΔG increases to -7.8 kcal/mol, which means that the N51A enzyme Py complex is 10.6 kcal/mol less stable than the wild type yCD Py complex. When Asp155 is changed to Ala, ΔG increases to -10.8 kcal/mol, which also means the mutant enzyme in complex with Py is less stable than that of the wild type yCD enzyme. These results explain the behavior of the enzymes in the binding assays, which revealed a great increase of K_d for both mutant enzymes.

Since Py may be present in the active site as either an unhydrated form or a hydrated form, the initial binding with Py (with no chemical transformation) alone may not account for the overall binding affinity with these two enzymes. To get a better understanding, ^{19}F NMR was performed, as presented in the previous section. The results suggest that there is no hydrated form inhibitor formed in the complex with either mutant enzyme.

Formation of DHP by the Mutant Enzymes. To explain why the hydrated inhibitor is not preferred by the N51A and D155A enzymes, QM/MM MD simulations and free energy calculations were performed to obtain the energy profiles for the formation of DHP from Py. We proposed the formation of DHP through three steps. First, a proton (HW) from the zinc-coordinated water transfers to the carboxyl group

(O^{ε1}) of Glu64. Second, N3 is protonated by O^{ε1} of Glu64. Third, the zinc-coordinated hydroxide attacks C4 and DHP is formed.

A 2D umbrella sampling technique was performed for the first two proton transfer steps. The PMF was obtained along two reaction coordinates, one is the distance between HW and O^{ε1} and the other is the distance between HW and N3. The energy profiles for the N51A and D155A enzymes, in Figure 6-4, are different from that for the wild type yCD. In complex with yCD, Py is stabilized by a hydrogen bond network, including those with Asn51 and Asp155. It costs about 7.7 kcal/mol for the proton transfer from the zinc-coordinated water to O^{ε1} of Glu64. After HW is transformed to O^{ε1} of Glu64, it still forms a hydrogen bond with N3. The conformation sits in a very shallow energy well. It costs another 4.9 kcal/mol for the proton (HW) transfer from O^{ε1} of Glu64 to N3 of Py. The transition state is found when the distances between HW and N3 and between HW and O^{ε1} are 1.15 and 1.47 Å, respectively. O4-C4 bond is already formed in a concerted manner with the proton transfer from O^{ε1} of Glu64 to N3 of Py. The yCD DHP complex is found to be 1.8 kcal/mol more stable than the yCD Py complex.

The behaviors of the N51A and D155A enzymes in complex with Py are different from that of the wild type yCD. For the N51A enzyme, Py can still form hydrogen bonds with O^{δ1} of Asp155 at a distance of 2.79 Å and with the backbone amide group of Gly63 at a distance of 3.01 Å. However, these interactions cannot hold Py in the same position as in the complex with yCD. The pyrimidine ring of Py rotates around 30° to form a

hydrogen bond with the zinc-coordinated water with a better configuration. Hence, extra energy is needed to transfer a proton from the zinc-coordinated water to O^{ε1} of Glu64.

The overall energy transition state was found when the distances between HW and O^{ε1} and between HW and N3 are 1.50 and 1.12 Å respectively, with an activation energy of 17.1 kcal/mol. For Py in complex with the D155A enzyme, its pyrimidine ring shifts further away from where it is in the complex with yCD, although O2 still forms a hydrogen bond with N^{δ2} of Asn51 with a distance of 2.88 Å. However, the hydrogen bond between the backbone amide of Gly63 and O2 is completely disrupted. N1 forms a new hydrogen bond with the backbone carbonyl oxygen of Asp155 with a distance of 2.89 Å. This hydrogen bond completely disrupts the one between the backbone carbonyl oxygen of Asp155 and N^{ε2} of His62, which is presented in the yCD Py complex to maintain the scaffold of the active site. The energy profile shows that the overall energy barrier is 26.1 kcal/mol, which is much larger than what is needed in the same process by yCD.

Transition state structures for three reactions catalyzed by the wild type, N51A and D155A yCDs were extracted and superpositioned, as shown in Figure 6-5. Although Py orientations in the initial structures are quite different, the transition state structures are very similar. Even the imidazole ring of His62 rotates back to the position, where it stays in the yCD complex. However, without the presence of Asn51 or Asp155 in these two mutant enzymes, the transition state is less stable than the wild type enzyme.

The N51A and D155A enzymes have low affinity for the inhibitor Py and the substrate 5FC, as revealed by steady state kinetics studies and binding assays. Their

contributions were explored by free energy decomposition, MM-PBSA calculations and QM/MM MD simulations from an energetic point of view. Py binds with the mutant enzymes in an alternative orientation compared to the complex with wild type yCD and the activation energy is elevated.

DISCUSSION

To understand the functions of conserved residues in the active site, we examined the N51A, D155A, E64A, G154G, E154D, R73A and D151A variant proteins. In Chapter 5, we already discussed the function of Glu64. The results show that Glu64 indeed participates in substrate binding and transition state stabilization. The other mutant proteins were successfully expressed and purified. These mutants were initially designed for different purposes.

N51A and D155A. Asn51 and Asp155 form direct hydrogen bonds with the ligand. These hydrogen bonds, with the hydrogen bond between the main chain amide group of Gly63 and O2, are important for substrate binding, and probably for transition state stabilization. The steady state kinetics data show that the K_m of the N51A and D155A enzymes for 5FC is greatly increased, indicating their function in substrate binding. However, the variant proteins have much lowered k_{cat} , which means that these residues are not merely used for substrate binding. Their presence may affect the catalysis by stabilizing the transition state.

These residues function in stabilization of the transition state as implied by ^{19}F NMR studies. ^{19}F NMR spectra were acquired to detect what form of the inhibitor was bound in the active site of the wild type yCD and the mutant enzymes. As discussed in

Chapter 5, the inhibitor in the active site of the wild type enzyme is the hydrated form 5FDHP. The results show that the inhibitor is the unhydrated form in complex with the N51A and D155A enzymes. The observed K_d is most likely the true dissociation constant of the unhydrated form of the inhibitor.

A reasonable explanation for why the hydrated form inhibitor cannot be formed by the mutant enzymes was obtained by QM/MM MD simulations and free energy calculations. For the unhydrated inhibitor Py, the pyrimidine ring adopts a different orientation, rotating about 30 degrees along the axis of N1-C4, when it binds with the N51A enzyme. Accordingly, the π -stacking interaction with His62 is attenuated. The energy needed for the formation of the hydrated inhibitor in the variant protein is 3.4 kcal/mol greater than that by the wild type enzyme. In complex with the D155A mutant enzyme, Py shifts further to a new position in which N1 forms a new hydrogen bond with the main chain carbonyl oxygen of Asp155. Hence, extra energy is required to shift Py back to a suitable position for the proton transfer from Glu64 to N3. The overall activation energy is 26 kcal/mol and the hydrated form is 15 kcal/mol greater than the unhydrated form. Thus, the inhibitor most likely stays in the unhydrated form, when binding with both mutant enzymes.

Py is a substrate mimic and shares a similar interaction pattern with the surrounding residues as for cytosine, except that Glu64 forms a hydrogen bond with N4 of cytosine. It is no surprise that substituting Asn51 and Asp155 with Ala results in a lower binding affinity for cytosine. K_d for Py is much lower than K_m for cytosine because Py is converted to DHP in the active site of yCD. For the initial binding, cytosine should be a better ligand due to the additional interaction between Glu64 and the amide group of cytosine. This result was supported by MM-PBSA calculations, which show

that the binding free energy of cytosine with yCD is 7.4 kcal/mol lower than that of Py. Similarly, cytosine is a better ligand for the N51A and D155A mutant enzymes. The reaction rate of the deamination is so slow that K_m is close to the dissociation constant of cytosine with yCD. The same results should hold for the 5FC and 5FPy pair, because the steady-state kinetics and binding assays indicate that K_m for 5FC is lower than K_d for 5FPy for the mutant enzymes.

E154G, E154D and R73A. Besides the N51A and D155A substitutions, other mutations in the C-terminal α -helix were introduced. Glu154 forms a strong salt bridge with Arg73 from the other subunit, which rigidifies the C-terminal α -helix. Previous studies show that product release is the rate-limiting step for the deamination of 5FC. E154G, E154D and R73A were designed to increase the flexibility of the C-terminal α -helix and hopefully to increase the rate of product release. However, the results show that the turnover rates (k_{cat}) were further reduced. The reduced rates are likely due to rate changes of the chemical steps because the flexibility of the C-terminal α -helix is indeed increased, as suggested by MD simulations. Py binds with these mutant enzymes in the hydrated form DHP, as shown by ^{19}F NMR spectra (data not shown), which suggests that the transition state is still stabilized by the mutant enzymes. However, the transition state may not bind as tightly as in the wild type enzyme.

Besides the hydrogen bonds that stabilize DHP, hydrophobic interactions are important too. A cluster of conserved residues Ile33, Leu88, Phe114 and Trp152 and Ile156 interact with DHP through hydrophobic interactions. Phe114, Ile156 and Trp152, separate the active site from being accessed from the solvent. The functions of these

residues need to be investigated to get a clear picture for how the enzyme interacts with its ligand and to understand the catalytic mechanism of enzymes in the CDA family of purine/pyrimidine deaminase.

CONCLUSIONS

Functions of conserved residues in the active site were investigated by site-directly mutagenesis, kinetics, NMR, and computational methods. A series of plasmids encoding mutant proteins were constructed, including the substitutions N51A, D155A, G154G, E154D, R73A, and D151A.

Asn51 and Asp155 form hydrogen bonds with the ligand. Substituting Asn51 or Asp155 with Ala greatly reduces the catalytic efficiency by a factor of 10^7 . The low activity is not only from reduced turnover rate, but also from weakened binding affinity. The k_{cat} value of N51A enzyme decreases about 1×10^4 -fold, relative to that of wild-type yCD, from 17 s^{-1} to $4.6 \times 10^{-3} \text{ s}^{-1}$. The K_{m} value of N51A protein increases about 25-fold, relative to that of wild-type yCD, from 0.16 mM to 4.1 mM. Similarly, for the D155A enzyme, the K_{m} of 5FC increases to 17 mM and the k_{cat} decreases to $9.8 \times 10^{-3} \text{ s}^{-1}$. The results suggest both Asn51 and Asp155 are important for ligand binding, as well as catalysis.

Initial binding of ligand to these enzymes was investigated by MM-PBSA and free energy decomposition. The binding free energies of Py with the mutant enzymes are much lower than that with the wild type enzyme. Formation of hydrated inhibitor by the mutant enzymes was investigated by ^{19}F NMR spectroscopy and QM/MM MD

simulation. The hydrated form was not observed in ^{19}F NMR spectra of the mutant proteins and the activation energy for the formation of the hydrated inhibitor is elevated, compared with the wild type enzyme. These results suggest that Asn51 and Asp155 participate in catalysis by stabilizing the transition state.

TABLES AND FIGURES

Table 6-1: Kinetic constants of the wild type yCD and mutant enzymes for the activation of 5-fluorocytosine.

Protein	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} s^{-1}$)
yCD WT	0.16 ± 0.01	17 ± 0.4	106
N51A	4.1 ± 1.0	$(4.6 \pm 0.6) \times 10^{-3}$	1.1×10^{-5}
D155A	17 ± 2.0	$(9.8 \pm 1) \times 10^{-3}$	5.4×10^{-6}
E154D	2.4 ± 0.3	2.6 ± 0.2	1.0
E154G	4.2 ± 0.7	8.7 ± 0.9	2.1
R73A	3.4 ± 0.7	1.7 ± 0.2	0.5
D151A	0.15 ± 0.03	6.3 ± 1.5	45

Table 6-2: Binding constants of 5FPy with the wild type yCD and mutant enzymes.

Protein	K_d , 5F-Py2O
yCD WT	$0.65 \pm 0.08 \mu M$
N51A	>100 mM
D155A	>100 mM

Table 6-3. Internal energy and solvation energy components for the interactions between residues and Py^a.

	ΔE_{ele}	ΔE_{vdw}	ΔE_{MM}	ΔG_{polar}	ΔG_{np}	ΔG_{sol}	$\Delta G_{\text{binding}}$	$\Delta \Delta G$
yCD- 2Py	-35.3	-16.3	-51.6	35.6	-2.4	33.2	-18.4	
N51A- 2Py	-26.7	-16.4	-43.1	37.9	-2.6	35.3	-7.8	10.6
D155A -2Py	-25.4	-18.0	-43.5	35.2	-2.5	32.7	-10.8	7.6
yCD- cyt	-48.1	-18.7	-66.9	43.6	-2.6	41.0	-25.8	
N51A- cyt	-42.2	-19.2	-61.4	44.0	-2.6	41.4	-20.0	5.8
D155A -cyt	-41.6	-19.1	-60.7	43.1	-2.6	40.5	-20.2	5.6

a: $\Delta E_{\text{MM}} = \Delta E_{\text{ele}} + \Delta E_{\text{vdw}}$; $\Delta G_{\text{sol}} = \Delta G_{\text{polar}} + \Delta G_{\text{np}}$; $\Delta G_{\text{binding}} = \Delta E_{\text{MM}} + \Delta G_{\text{sol}}$.

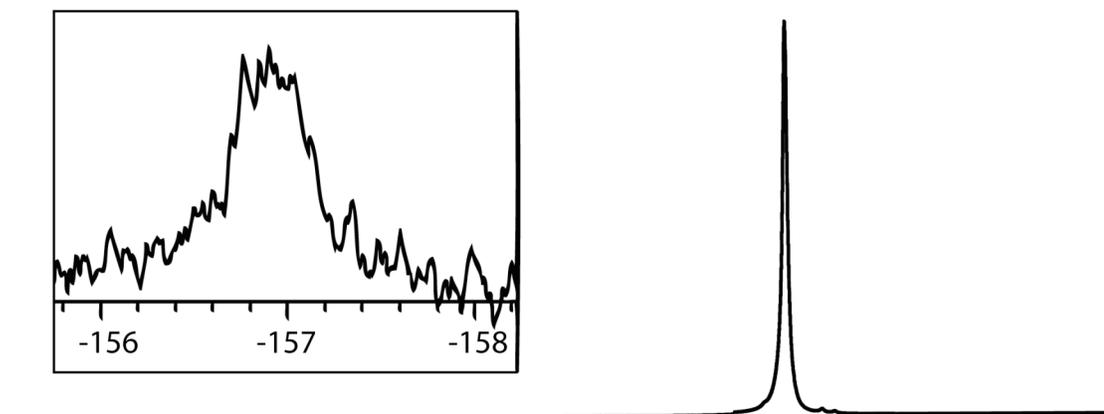
Table 6-4. Individual energy components (kcal/mol) by MM-GBSA energy decomposition^a.

	SGAS	BGAS	SGBsol	BGBsol	SGBtot	BGBtot	TGBtot
Val31	-0.4	0.47	0.12	-0.6	-0.2	-0.1	-0.3
Ile33	-1.3	0.46	0.22	-0.7	-1.1	-0.2	-1.3
Asn51	-3.3	-0.7	0.87	0.48	-2.4	-0.3	-2.7
His62	-2.1	-2.4	0.07	0.89	-2.0	-1.5	-3.5
Gly63	0	-2.5	0	1.29	0	-1.2	-1.2
Glu64	-0.7	0.1	1.47	-0.2	0.79	-0.2	0.59
Thr86	-0.2	-0.2	-0.1	0.09	-0.3	-0.1	-0.3
Leu88	-0.4	0.14	-0.1	-0.2	-0.5	-0.1	-0.6
Ser89	0.19	-1.2	-0.2	1.03	0	-0.2	-0.2
Pro90	0.19	-0.4	-0.2	0.29	-0.1	-0.1	-0.2
Cys91	-1	0	0.6	0	-0.4	0	-0.4
Trp152	-0.5	-0.3	0.17	0.34	-0.3	0.01	-0.3
Asp155	-6.0	-0.4	5.31	0.38	-0.7	-0.01	-0.7
Ile156	0.11	-0.3	-0.4	0.22	-0.2	-0.1	-0.3
Zn	0.7	0	-0.3	0	0.39	0	0.39
H ₂ O	-0.3	0	0.5	0	0.23	0	0.23

a: SGAS and BGAS are the gas phase energy for the sidechain and backbone, respectively. SGBsol and BGBsol are the solvation energy for the sidechain and backbone, respectively. SGBtot and BGBtot are the total binding free energy for the sidechain and backbone, respectively. TGBtot is the total binding free energy for the whole residue.

D: 5FPy + N51A

Free 5FPy



C: 5FPy+D155A

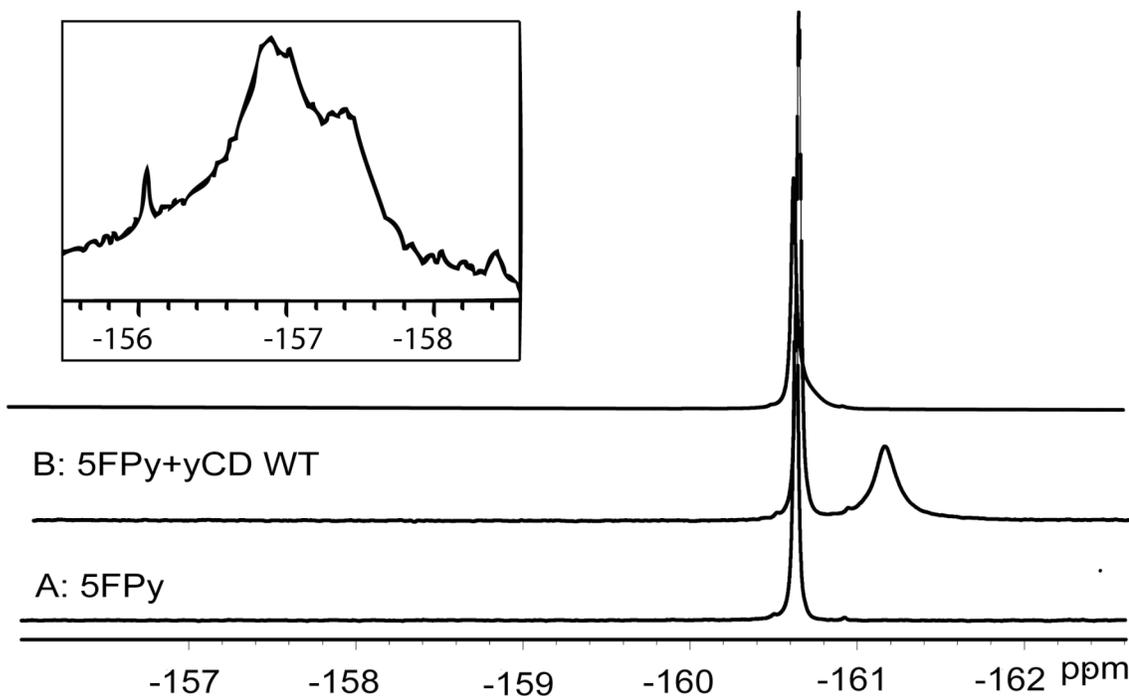


Figure 6-1. ^{19}F -NMR spectra of the mutant enzyme-5FPy complexes and wild type yCD-5FPy complex. (A) 1.5 mM 5FPy in solution. (B) 5FPy (3.0 mM) in complex with yCD (1.5 mM). (C) 5FPy (20.0 mM) in complex with N51A (1.5 mM). The insert shows a small peak at -157 ppm, corresponding to the inhibitor in the hydrated form. (D) 5FPy (200 mM) in complex with D155A (1.5 mM). The insert shows a small peak at -157 ppm, corresponding to the inhibitor in the hydrated form.

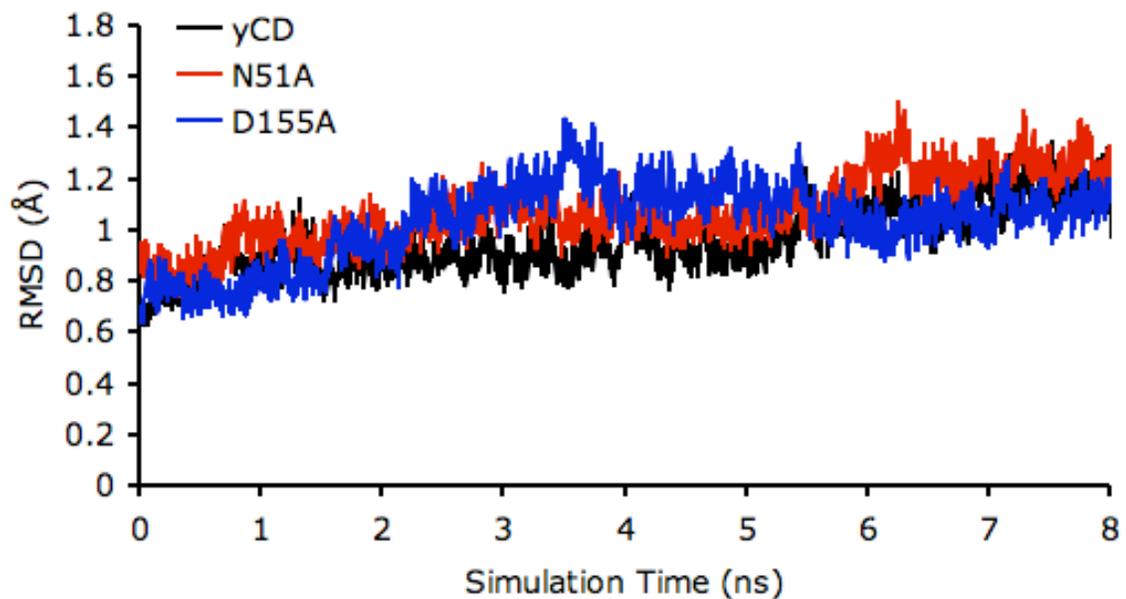


Figure 6-2. RMSDs of the backbone C α atoms relative to the starting structure for the yCD Py complex during the 8 ns production MD simulations.

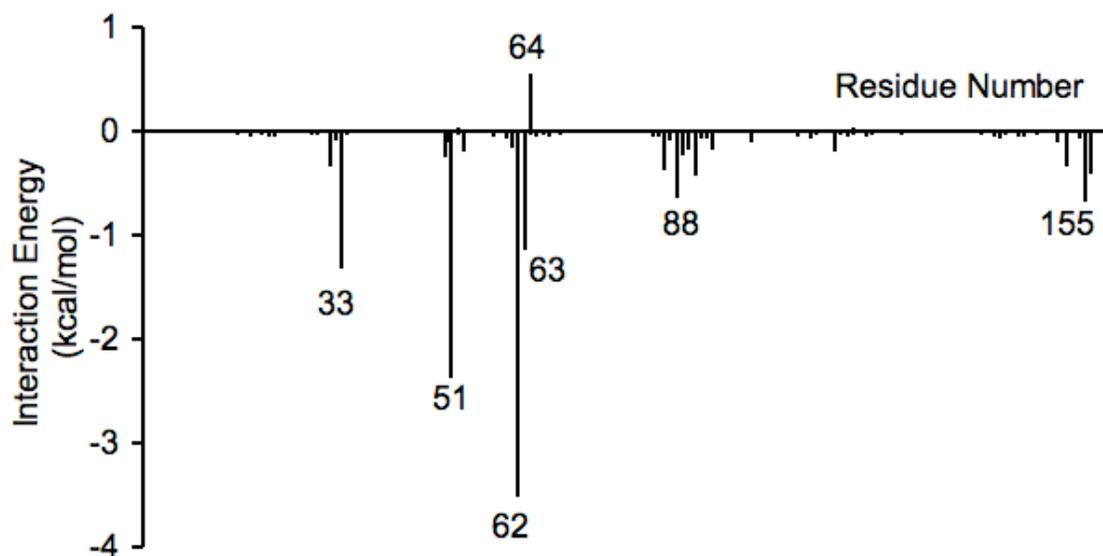


Figure 6-3. Ligand-residue interaction spectrum of the yCD Py complex by MM-GBSA energy decomposition.

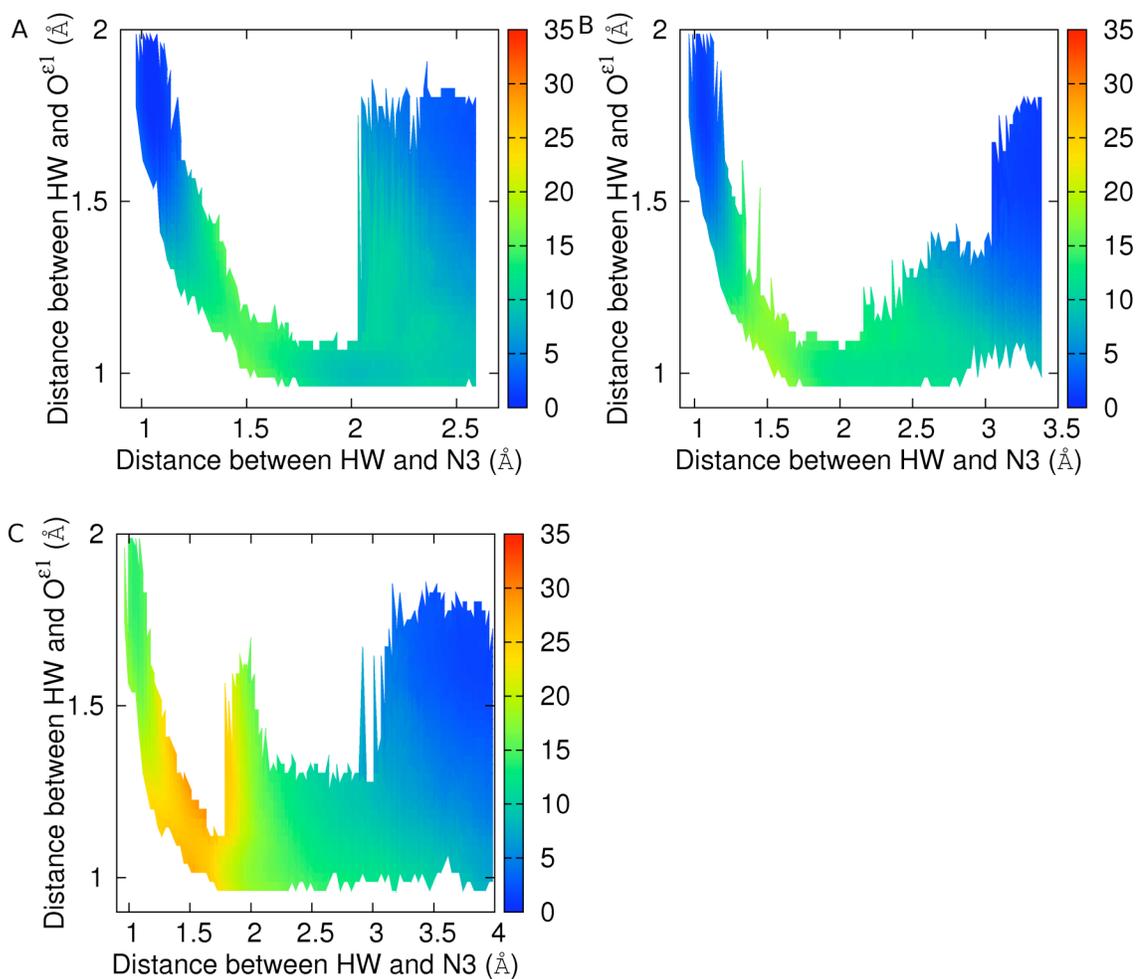


Figure 6-4. 2D energy profiles for proton transfer from the zinc-coordinated water to N3 of Py. (A) The yCD Py complex. (B) The N51A Py complex. (C) The D155A Py complex.

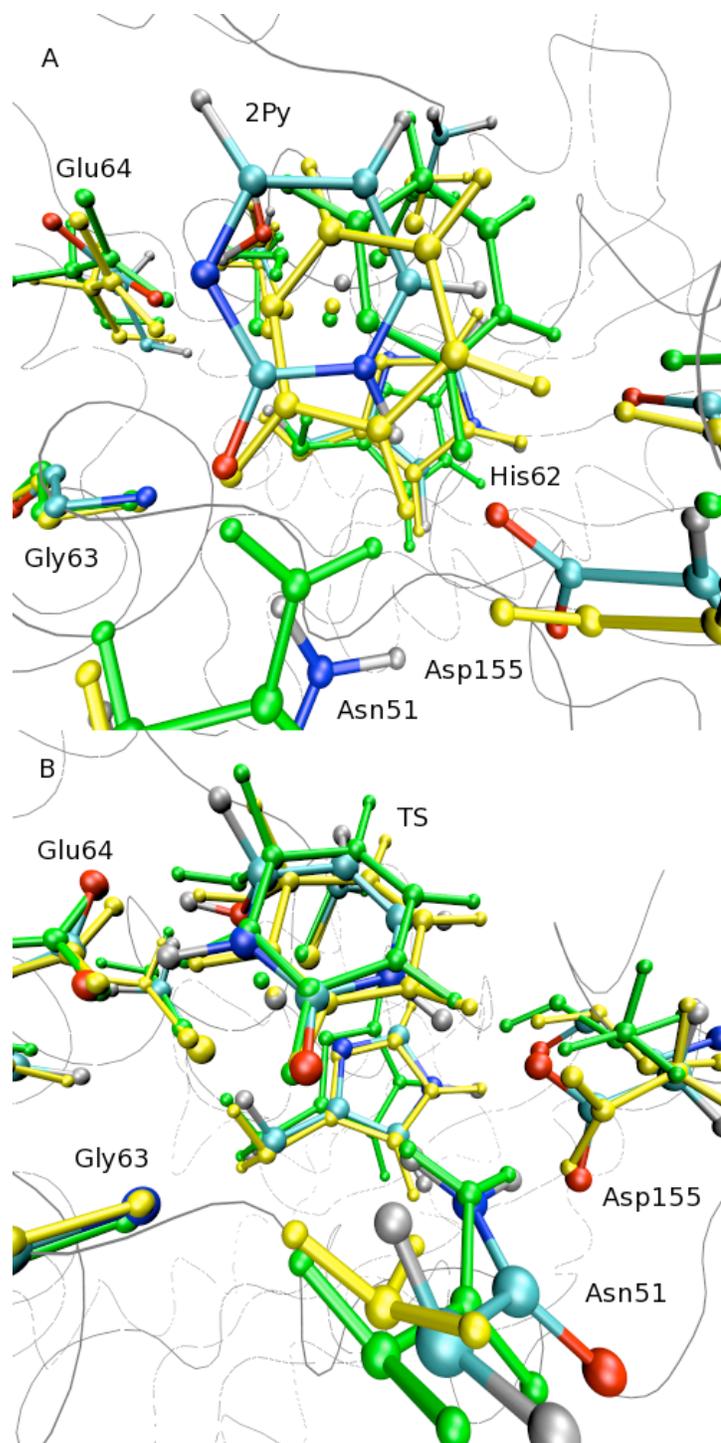


Figure 6-5. Superposition of representative snapshots for the yCD Py complex, N51A Py complex and D155A Py complex. (A) Initial structures. (B) Transition state structures. The residues of the yCD Py complex are colored by atom names; the residues of the N51A Py complex are colored yellow; and the residues of the D155A Py complex are colored green.

REFERENCES

REFERENCES

1. Aghi, M., Hochberg, F., and Breakefield, X. (2000) Prodrug activation enzymes in cancer gene therapy, *J Gene Med* 2, 148-164.
2. Greco, O., and Dachs, G. U. (2001) Gene directed enzyme/prodrug therapy of cancer: Historical appraisal and future perspectives, *J Cell Physiol* 187, 22-36.
3. Erbs, P., Regulier, E., Kintz, J., Leroy, P., Poitevin, Y., Exinger, F., Jund, R., and Mehtali, M. (2000) *In vivo* cancer gene therapy by adenovirus-mediated transfer of a bifunctional yeast cytosine deaminase/uracil phosphoribosyltransferase fusion gene, *Cancer Res* 60, 3813-3822.
4. Ko, T., Lin, J., Hu, C., Hsu, Y., Wang, A., and Liaw, S. (2003) Crystal structure of yeast cytosine deaminase - Insights into enzyme mechanism and evolution, *J Biol Chem* 278, 19111-19117.
5. Ireton, G., Black, M., and Stoddard, B. (2003) The 1.14 angstrom crystal structure of yeast cytosine deaminase: Evolution of nucleotide salvage enzymes and implications for genetic chemotherapy, *Structure* 11, 961-972.
6. Carlow, D. C., Short, S. A., and Wolfenden, R. (1996) Role of glutamate-104 in generating a transition state analogue inhibitor at the active site of cytidine deaminase, *Biochemistry* 35, 948-954.
7. Yao, L., Li, Y., Wu, Y., Liu, A., and Yan, H. (2005) Product release is rate-limiting in the activation of the prodrug 5-fluorocytosine by yeast cytosine deaminase, *Biochemistry* 44, 5940-5947.
8. Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E. (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models, *Acc Chem Res* 33, 889-897.
9. Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A., and Case, D. A. (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices, *J Am Chem Soc* 120, 9401-9409.
10. Carlow, D. C., Smith, A. A., Yang, C. C., Short, S. A., and Wolfenden, R. (1995) Major contribution of a carboxymethyl group to transition-state stabilization by cytidine deaminase: Mutation and rescue, *Biochemistry* 34, 4220-4224.
11. Gilson, M. K., Sharp, K. A., and Honig, B. H. (1988) Calculating the electrostatic potential of molecules in solution: Method and error assessment, *J Comput Chem* 9, 327-335.
12. Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S., and Seifert, G. (1998) Self-consistent-charge density-functional tight-

- binding method for simulations of complex materials properties, *Phys Rev B* 58, 7260.
13. Elstner, M. (2005) The SCC-DFTB method and its application to biological systems, *Theor Chem Acc* 116, 316-325.
 14. Walker, R. C., Crowley, M. F., and Case, D. A. (2008) The implementation of a fast and accurate QM/MM potential method in Amber, *J Comput Chem* 29, 1019-1031.
 15. Seabra, G. D. M., Walker, R. C., Elstner, M., Case, D. A., and Roitberg, A. E. (2007) Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics Package, *J Phys Chem A* 111, 5655-5664.
 16. Riccardi, D., and Cui, Q. (2007) pK_a analysis for the zinc-bound water in human carbonic anhydrase II: Benchmark for "multiscale" QM/MM simulations and mechanistic implications, *J Phys Chem A* 111, 5703-5711.
 17. Nam, K., Gao, J., and York, D. M. (2005) An efficient linear-scaling Ewald method for long-range electrostatic interactions in combined QM/MM calculations, *J Chem Theory Comput* 1, 2-13.
 18. Cui, Q., Elstner, M., Kaxiras, E., Frauenheim, T., and Karplus, M. (2001) A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method, *J Phys Chem B* 105, 569-585.
 19. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, *J Chem Phys.* 98, 10089-10092.
 20. Torrie, G. M., and Valleau, J. P. (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J Comput Phys* 23, 187-199.
 21. Roux, B. (1995) The calculation of the potential of mean force using computer simulations, *Comput Phys Commun* 91, 275-282.
 22. Grossfield, A. WHAM: the weighted histogram analysis method, version 2.0, <http://membrane.urmc.rochester.edu/content/wham>

CHAPTER 7

SUMMARY AND FUTURE WORK

SUMMARY

My efforts have been devoted to understanding the catalytic mechanism of yeast cytosine deaminase by both experimental and computational methods.

I tried to explore the enzymatic mechanism by probing the transition state by ^{13}C kinetic isotope effects and quantum mechanics calculations. ^{13}C KIEs were measured at natural abundance by NMR spectroscopy. The results indicate that the reaction proceeds by a stepwise $\text{A}_\text{N}+\text{D}_\text{N}$ mechanism, with a transition state in the C4-N4 bond cleavage step. To obtain more detailed information about the reaction pathway and facilitate protein design of variant enzymes with increased substrate specificity towards 5FC, a two-layered ONIOM(QM:MM) method was employed to study the deamination of 5FC and the reaction pathway was compared with that of cytosine. The results show that the reaction pathway for the deamination of 5FC is similar to that of cytosine. Energetically favorable tetrahedral intermediates were found for 5FC, which may account for the relatively low K_m for 5FC. However, configurational sampling of proper geometries during the Zn-O4 bond breakage was not promising. Hence, the pathway was re-investigated by QM/MM MD simulations and free energy calculations. The results show

that the C4-N4 bond cleavage is indeed the rate-limiting step with an activation energy of 19.7 kcal/mol. The activation energy for the Zn-O4 bond breakage is only 8.1 kcal/mol.

Functional roles of conserved residues in the active site were investigated by site-directed mutagenesis, biochemical, NMR, and computational methods. Glu64 is believed to serve as a proton shuttle during multiple steps. Steady-state kinetic studies show that changing Glu64 to alanine causes a dramatic decrease in k_{cat} and a dramatic increase in K_{m} . The substitution also causes a dramatic decrease in the affinity of the reaction intermediate analogue 5FPy. ^{19}F NMR experiments show that inability to form a hydrated inhibitor is one of the causes for the low binding affinity. ONIOM calculations also demonstrate that, without the help of Glu64, both proton transfers before and after the formation of the tetrahedral intermediate become partially rate-limiting steps. The results indicate that Glu64 is not only directly involved in catalysis, but also important for substrate binding.

The functional roles of the active site residues, particularly Asn51 and Asp155, were investigated as well. Steady-state kinetic studies show that both the N51A and D155A proteins exhibit a dramatic decrease in k_{cat} and a dramatic increase in K_{m} . The substitutions also cause a decrease in the binding affinity of 5FPy by a factor of >100.

The lowered binding affinity is partially due to weak initial binding, suggested by MM-PBSA and binding free energy calculations, which show that changing Asn51 and Asp155 to alanine reduces the binding affinity of Py by 10.6 and 7.6 kcal/mol, respectively. The binding of 5FPy to the N51A and D155A enzymes causes a downfield shift in the ^{19}F NMR signal, indicating that the inhibitor is in the unhydrated form.

QM/MM MD simulations show that, with the N51A and D155A enzymes, the hydrated form is no longer favored and the activation energy is elevated. The results suggest that similar to Glu64, Asp51, and Glu155 are important not only for the initial binding of the ligand, but also for the stabilization of the transition state.

FUTURE WORK

Starting from what we have learned, other researchers will be able to continue to work on the catalytic mechanism of yCD. Meanwhile, researchers could try to construct thermostable variants of yCD, which are highly desirable in the gene directed enzyme prodrug therapy (GEDPT).

Deuterium and ^{15}N KIEs. ^{13}C KIE studies provide a relatively rapid and convenient probe to detect the transition state of yCD. While it alone may not precisely describe the conformation of the transition state, the use of ONIOM calculations can provide additional information. Deuterium and ^{15}N can be employed to get a better description of the transition state. Deuterium KIEs can be measured by proton inventory analysis of reactions in buffers with different D_2O content. ^{15}N KIEs can be measured for substrate at natural abundance by isotope ratio mass spectrometric analysis. More details in the transition state are expected. For example, the protonation state of N3 can be determined from the ^{15}N KIE of N3 and deuterium KIE of H3. By combining deuterium, ^{15}N and ^{13}C KIEs, one can also determine if the proton transfer from Glu64 to N4 and the C4-N4 bond cleavage are concerted or stepwise.

Function of other residues in the active site. The conserved Glu64 and the zinc apparatus define the enzyme as a deaminase. Asn51 and Asp151 provide specific interactions with the substrate and the transition state. However, these residues may not offer enough driving forces for catalysis. The other residues surrounding the substrate, such as Ile33, Phe114, Leu88, and Trp152, etc, also contribute through hydrophobic interactions. For example, changing Trp152 to leucine, which has a shorter side chain, reduces the catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$) by 2-fold (data not shown). How exactly these residues contribute to catalysis is worth further study.

Protein engineering. Besides understanding the catalytic mechanism, researchers are interested in designing thermostable variants of yCD with high catalytic efficiency. Computational approaches and random mutagenesis approaches have been developed to design stable protein with point mutation in the enzyme core region or the active site. In our laboratory, we identified such variants showing better catalytic efficiency and thermostability in *in vitro* enzymatic assays. New variants can be designed by combining two or more mutations into a single construct. Cell assays or animal models need to be designed to evaluate their performance.