THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Master of Science (Chemistry)

DEGREE

Chemistry	Chemistry
FIELD	**DEPARTMENT**

TITLE: Hybrid Quantum Mechanical/Molecular Mechanical Studies on the Reaction Mechanism of *cis,cis*-Muconate Lactonizing Enzymes

NAME: Ms. Tuanjai Somboon

THIS THESIS HAS BEEN ACCEPTED BY

THESIS ADVISOR

(Associate Professor Supa Hannongbua, Dr.rer.nat.)

THESIS CO-ADVISOR

(Mr. Matthew Paul Gleeson, Ph.D.)

THESIS CO-ADVISOR

(Assistant Professor Chak Sangma, Ph.D.)

DEPARTMENT HEAD

(Associate Professor Supa Hannongbua, Dr.rer.nat.)

APPROVED BY THE GRADUATE SCHOOL ON _____

DEAN

(Associate Professor Gunjana Theeragool, D.Agr.)

THESIS

HYBRID QUANTUM MECHANICAL/MOLECULAR MECHANICAL STUDIES ON THE REACTION MECHANISM OF *CIS,CIS*-MUCONATE LACTONIZING ENZYMES

TUANJAI SOMBOON

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Master of Science (Chemistry)
Graduate School, Kasetsart University
2011

Tuanjai Somboon 2011: Hybrid Quantum Mechanical/Molecular Mechanical Studies on the Reaction Mechanism of *cis,cis*-Muconate Lactonizing Enzymes. Master of Science (Chemistry), Major Field: Chemistry, Department of Chemistry. Thesis Advisor: Associate Professor Supa Hannongbua, Dr.rer.nat. 80 pages.

Muconate lactonizing enzymes (MLEs) derived from *Mycobacterium smegmatis* and *Pseudomonas fluorescens* share ~76% identity and have a very similar arrangement of catalytic residues in their respective active site. MLEs catalyze the conversion of *cis,cis*-muconate to the same achiral product, muconolactone. However, studies in deuterated solvent surprisingly show that the cyclo-isomerization proceeds along stereochemically distinct mechanistic routes. The theoretical calculation performed here have helped to confirm the identity of the basic residues involved in the MLE reaction originally proposed by Sakai et al through indirect means from an analysis of the products produced in experiments performed in deuterated solvents. The results show that although the basic Lys residues located on the 6^{th} strand in both MLEs are almost equally well positioned to accept a proton, however, it is the base located on the 2^{nd} strand that is thermodynamically and kinetically more favorable, as well as giving the optimized QM/MM product closest to the original X-ray structure.

_____ _____ ___ / ___ / ___

Student's signature Thesis Advisor's signature

ACKNOWLEDGEMENTS

I am sincerely grateful to Associated Professor Dr. Supa Hannongbua, my thesis advisor, for her advice, valuable suggestion and continuous support for my graduation. I am particularly grateful to Dr. Matthew Paul Gleeson for his comments and discussion. Moreover, I am sincerely thankful to Assistant Professor Dr. Chak Sangma for his helpful criticism and suggestion.

I am deeply thankful to Thailand Research Fund (RTA5380010) and the National Research University research fund of Kasetsart University. The National Center of Excellence for Petroleum, Petrochemical, and Advanced Materials (NCE-PPAM) and Faculty of Science, Kasetsart University is grateful for my financial support. I am also thankful to Graduate School of Kasetsart University for a partial research support. The authors are grateful to Yan Zhao and Donald G. Truhlar for their support with the M05 functional. The computational resource at Laboratory for Computational & Applied Chemistry (LCAC), Kasetsart University is also acknowledged.

Finally, I am specially appreciated my parents, sisters, friends and colleagues at LCAC for everything given to me throughout my entire study.

Tuanjai Somboon

May 2011

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

2-PGA	=	2-Phosephoglycerate
3-Methyl Asp	=	3-methylaspartate
Å	=	Angstrom
AE Epim	=	L-Ala-D/L-Glu epimerases
AltD	=	D-altronate dehydratase
ASN	=	Asparagine
ASP	=	Asparagine
B3LYP	=	Becke's three parameter hybrid functional using the LYP correlation functional
C	=	Carbon
CDK	=	Cyclin dependent kinase
CDK2	=	Cyclin dependent kinase 2
Cpd	=	Compound
DFT	=	Density Functional Theory
E	=	Energy
EE	=	Electrical Embedding
fs	=	Femtosecond
GalD	=	D-galactonate dehydratase
GB	=	Giga Bytes
GHz	=	Giga Hertz
GlcD	=	D-gluconate dehydratase
GLN	=	Glutamine
GLU	=	Glutamic acid
GlucD	=	D-glucarate/L-idarate dehydratase
H	=	Hydrogen
HIE	=	Histidine
ILE	=	Isoleucine
IPNS	=	Isopenicilin N synthase
K	=	Kelvin

LIST OF ABBREVIATIONS (Continued)

kcal/mol	=	Kilocalorie per mole
LCAC	=	Laboratory for Computational and Applied Chemistry
LEU	=	Leucine
LYS	=	Lysine
MAL	=	3-methylaspartate ammonia lyase
ManD	=	D-mannonate dehydratase
MD	=	Molecular Dynamics
Mg	=	Magnesium
Mg^{2+}	=	Magnesium ion
MLE	=	Muconate Lactonizing Enzyme
MM	=	Molecular Mechanics
MR	=	Mandelate racemase
N	=	Nitrogen
NVT	=	Number of particle, volume and temperature
O	=	Oxygen
ONIOM	=	Our own N-layered Integrated molecular Orbital and molecular Mechanics
OSB	=	*o*-succinylbenzoate
OSBS	=	*o*-succinylbenzoate synthases
PDB	=	Protein Data Bank
PEP	=	P-enolpyruvate
PHE	=	Phenylalanine
ps	=	Picosecond
QM	=	Quantum Mechanics
QM/MM	=	Quantum Mechanical/Molecular Mechanical
RH	=	Hydrocarbon
RhamD	=	L-rhammonate dehydratase

LIST OF ABBREVIATIONS (Continued)

RMS	=	Root Mean Square
RMSD	=	Root Mean Square Deviation
SER	=	Serine
SHCHC	=	2-succinyl-6-hydroxyl-2,4-cyclohexadiene-1-carboxylate
sub	=	Substrate
T	=	Temperature
THR	=	Threonine
TIM	=	Triose phosphate isomerase
TS	=	Transition State
UFF	=	Universal Force Field
WAT	=	Water
ΔE	=	Energy difference

HYBRID QUANTUM MECHANICAL/MOLECULAR MECHANICAL STUDIES ON THE REACTION MECHANISM OF *CIS,CIS*-MUCONATE LACTONIZING ENZYMES

INTRODUCTION

Muconate lactonizing enzymes (MLEs) are an interesting member of this superfamily which catalyze the conversion of *cis,cis*-muconates to muconolactones. Sakai *et al.* (Sakai *et al.*, 2009a; b; Sakai *et al.*, 2006) have recently reported data on two different MLEs derived from *Mycobacterium smegmatis* and *Pseudomonas fluorescens* that share ~76% identity. These proteins catalyze the same chemical reaction (Figure 1), but involve a stereochemically distinct reaction mechanism even though the product is achiral. Studies in deuterated solvent have established that cyclo-isomerization catalyzed by *M. smegmatis* (*anti*-MLE) proceeds along an *anti*-stereochemical course, whereas *P. fluorescens* (*syn*-MLE) catalyzes the *syn*-stereochemical course.

MLEs are Mg^{2+} containing metallo-proteins consisted of ~370 amino acids arranged into a TIM barrel-like α/β protein fold. The Mg^{2+} is found deep within a non-solvent exposed cavity, coordinated by 2 aspartate and 1 glutamic acid residues, a single water molecule and the substrate, in a distorted octahedral form. The carboxylate of *cis,cis* muconate binds across the Mg^{2+} ion in conformations that display mirror-like symmetry in the two different proteins (Figure 2). The 2^{nd} carboxylate group of *cis,cis* muconate interacts with residues towards the rear of the pocket, glutamine and threonine in *anti*-MLE and histidine and threonine residues in *syn*-MLE. Two Lys residues, located on the 2^{th} and 6^{th} strands, are sufficiently closed to the substrate alpha carbon to act as the general base in the reaction.

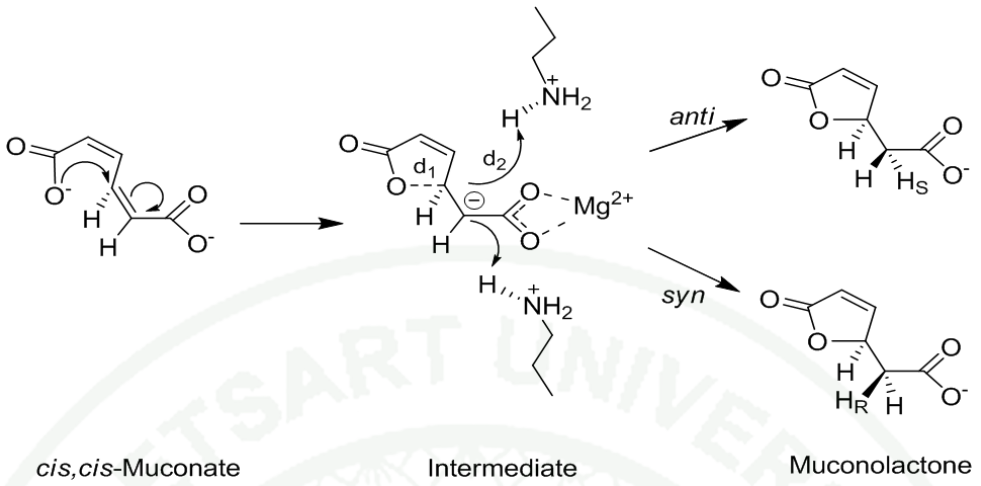


Figure 1 Reaction mechanism proposed by Sakai *et al.* for *anti*- and *syn*-MLEs. Two proximal Lys residues exist in both active sites, which could potentially lead to the formation of either the *syn*- or *anti*- products. In fact, while both proteins catalyze the formation of the same natural precursor, they lead to the formation of different chiral products based on experiments performed in deuterated solvent.

Structural information has proved crucial to understand of the sequence of events that lead to the chiral products in deuterated solvent. Sakai *et al.* have determined X-ray crystal structures of the product state (muconate lactone), for both *syn*- and *anti*-proteins, and have used these to identify the most probable base in the catalytic reaction. They propose that the identity of the base in both proteins is the Lys residue at the end of the 2nd strand that explains the stereochemical aspects of the reaction. This result is interesting given members of the structurally related epimerase sub-class achieve their stereo-specific outcomes by relying on either the base from the 2nd or 6th strand. Furthermore, analysis of the interaction distances between the nitrogen atoms of the two possible bases and the product α-carbon in both MLE X-ray structures (Table 1) simplistically suggests that Lys on the 6th strand is more likely to be the base in *syn*-MLE (3.03 Å and 3.57 Å). This however would not explain the stereo-chemical differences in the reaction. Additionally, it is not known if a stable enolate anion does in fact exist as a meta-stable intermediate in the reaction.

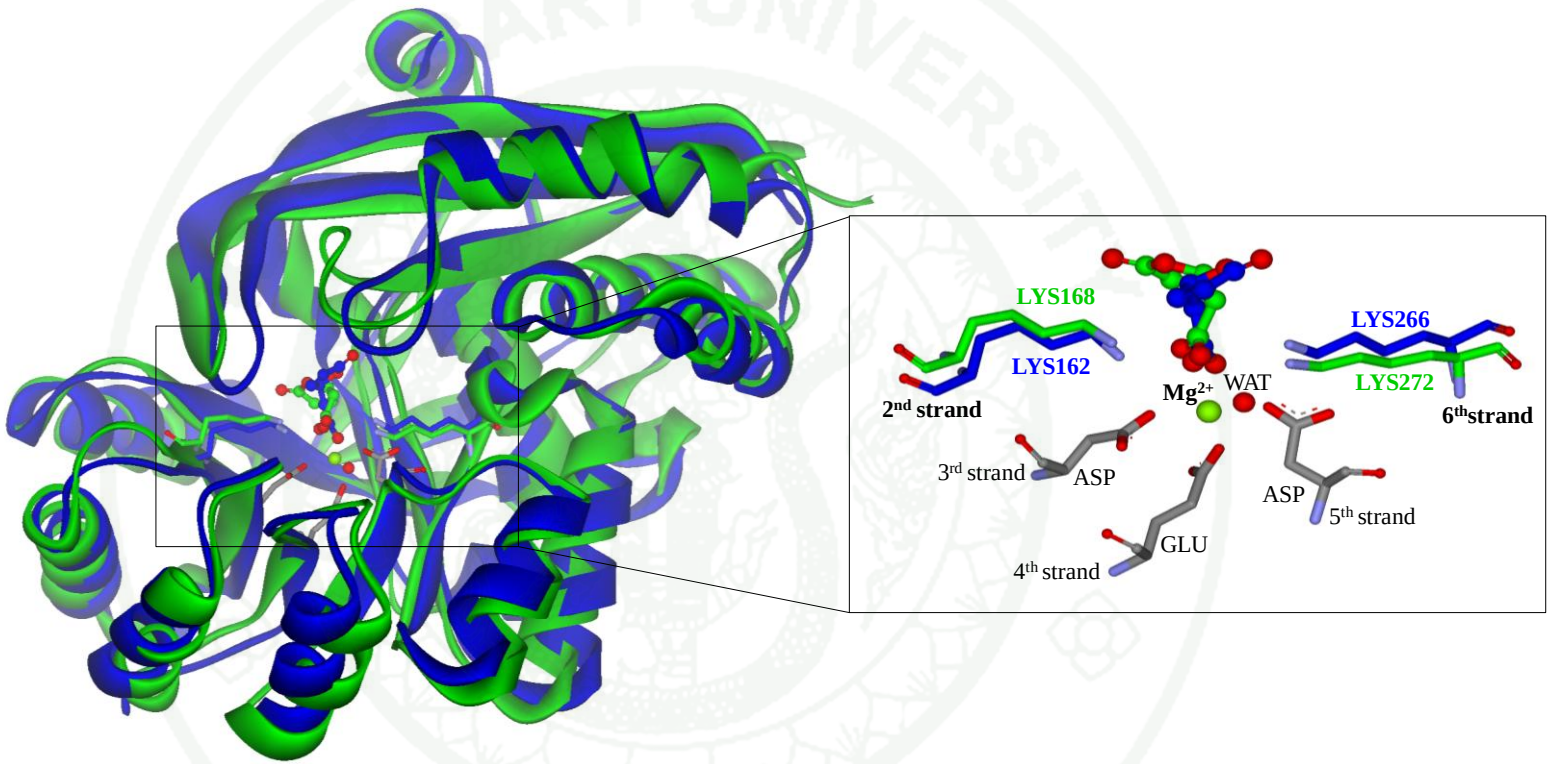


Figure 2 A comparison of the *anti*- (blue) and *syn*-MLE (green) proteins and active site region.

Table 1 Experimental distances observed for Muconolactone product. Distances are reported in Å.

Distances	X-ray structure
***Anti*-MLE**	
Mg−O(ASP191)	2.08
Mg−O(GLU217)	1.96
Mg−O(ASP242)	2.19
Mg−O(Sub)	2.16
Mg−O(Sub)	2.29
C(Sub)−N(LYS162)	3.03
C(Sub)−N(LYS266)	3.57
***Syn*-MLE**	
Mg−O(ASP197)	2.34
Mg−O(GLU223)	2.33
Mg−O(ASP248)	2.24
Mg−O(Sub)	2.42
Mg−O(Sub)	2.96
C(Sub)−N(LYS168)	4.22
C(Sub)−N(LYS272)	3.45

OBJECTIVES

1. To understand the origin of differences in stereochemical courses for *anti*- and *syn*-MLEs

2. To prove the identity of the base in the *anti*- and *syn*-MLE proteins given the contrasting way related enolase family member achieve stereo-specific catalysis

LITERATURE REVIEW

The enolase superfamily has received a considerable amount of attention from a biochemical perspective recently as it has helped to illustrate the complexity and redundancy in enzyme evolution. The enolase superfamily has been identified by the presence of the conserved amino acid residues in the enzymes that catalyze the same reaction, which guides the recognition of an enzyme family (Silverman, 2000). The fundamental common chemistry of this diverse superfamily is the proton abstraction α to a carboxylate to form metal-ion stabilized enolate intermediates as the first step of their overall reaction (Babbitt *et al.*, 1996; Glasner *et al.*, 2006). The word "mechanistically diverse" was used to express the functional relationships with in the enolase superfamily (Gerlt and Babbitt, 1998; Neidhart *et al.*, 1990).

This superfamily can be partitioned into four subgroups on the basis of active sites residues and phylogenetic analysis (Gerlt *et al.*, 2005): (1) enolase, (2) mandelate racemase (MR), (3) muconate lactonizing enzyme (MLE), and (4) 3-methylaspartate ammonia lyase (MAL). The active site of this superfamily contains the conserved residues (Glu or Asp) coordinated with essential Mg^{2+} ion located at the ends of third, fourth, and fifth β-stands (Gerlt and Babbitt, 2001). The general acid/base residues are positioned in the $(\beta,\alpha)_8$-barell domain. In MLE subgroup, the acid/base catalysts are Lys residues at the end of both the second and sixth β-stands, whereas His and Asp residues are located at the end of the seventh and sixth β-stands of MR subgroup, respectively. From an analysis of sequence and structural similarities of the enzymes of this superfamily, it is found that their chemical reactions are mediated by similar rather active sites (Schmidt *et al.*, 2003; Vick and Gerlt, 2007; Vick *et al.*, 2005).

Eight different overall reactions have been identified for the members of this large superfamily: (1) enolase (Larsen *et al.*, 1996), (2) mandelate racemase (Neidhart *et al.*, 1991), (3) muconate lactonizing enzymes (Hasson *et al.*, 1998; Helin *et al.*, 1995; Kajander *et al.*, 2003), (4) D-glucarate dehydratase (Gulick *et al.*, 2000; Gulick *et al.*, 1998; Gulick *et al.*, 2001), (5) D-galactonate dehydratase (Wieczorek *et al.*, 1999), (6) *o*-succinylbenzoate synthases (Klenchin *et al.*, 2003; Thoden *et al.*, 2004; Thompson *et al.*, 2000), (7) L-Ala-D/L-Glu epimerases (Gulick *et al.*, 2001), and (8) 3-methylaspartate ammonia lyase (Asuncion *et al.*, 2002). It was suggested that the evolution of the enzymes was recruited by modification of an existing enzymes to catalyze similar chemistry with different substrate specification (Petsko *et al.*, 1993).

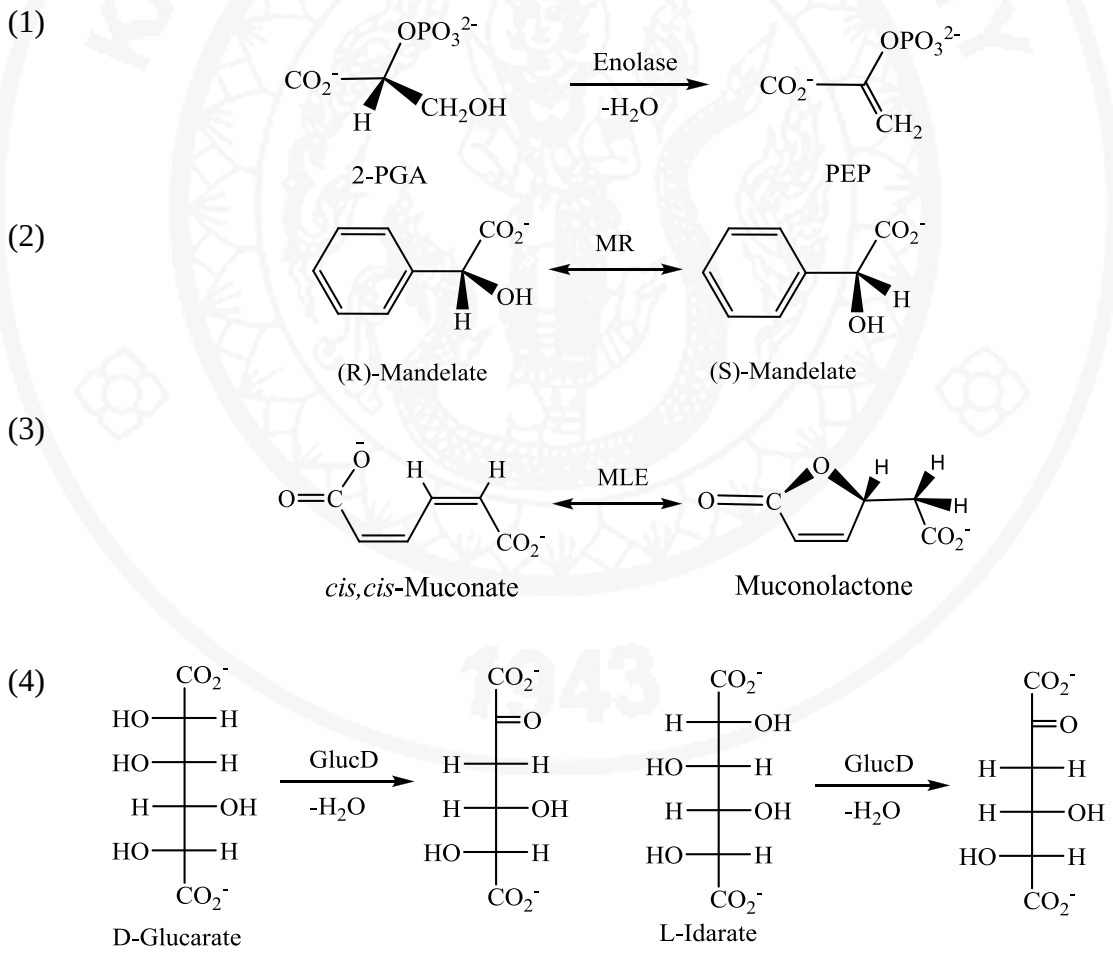


Figure 3 The reactions catalyzed by members of the enolase superfamily.

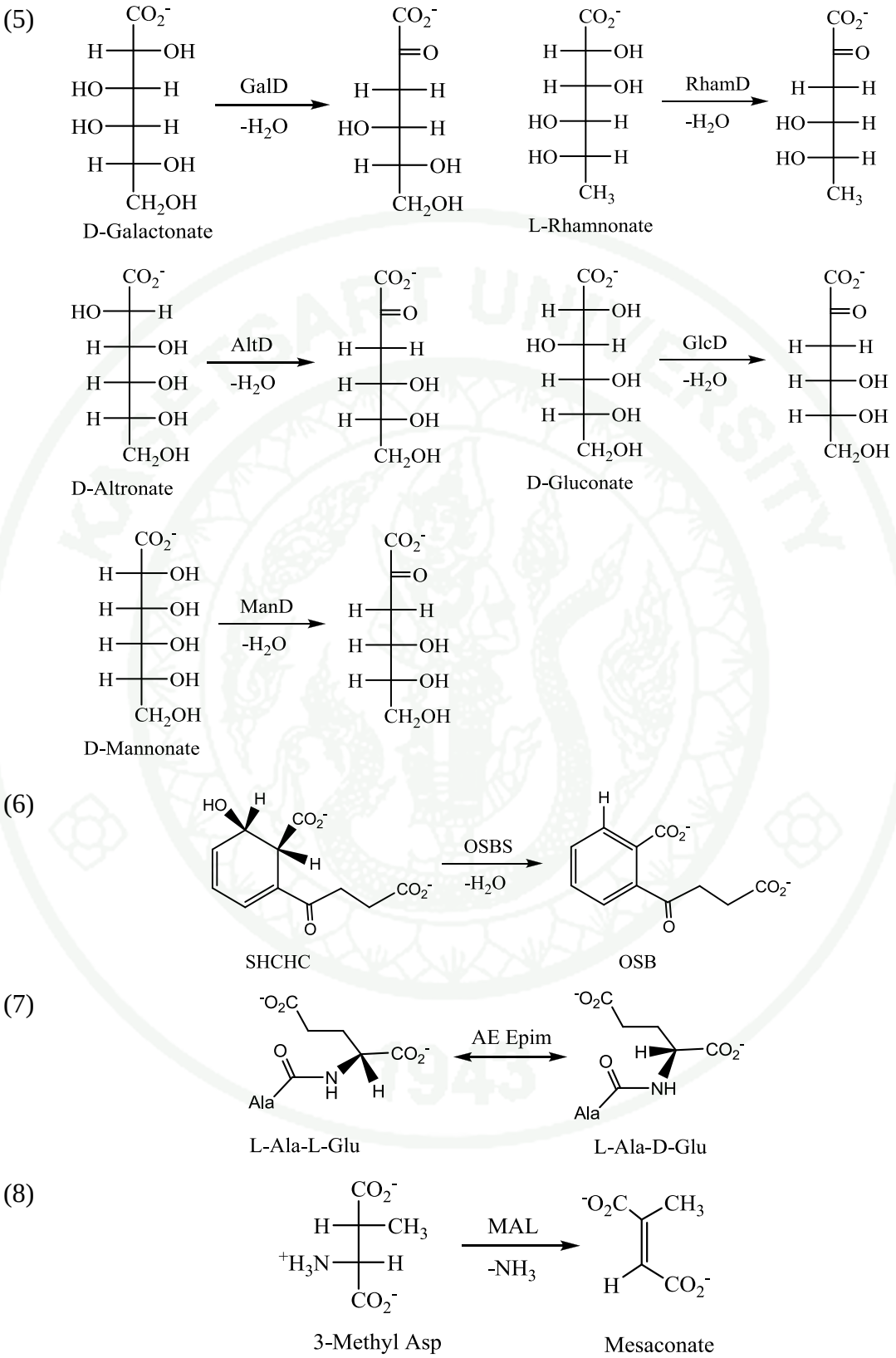


Figure 3 (continued).

Computational chemistry can play an important role in understanding protein function since it enables us to simulate events at an atomic level. The preferred high level QM methods cannot however be employed on protein sized systems so more approximated methods such as the Hybrid Quantum Mechanical/Molecular Mechanical (QM/MM) methods must be used. In the QM/MM technique the active site residues that undergo chemical change, or directly influence the sequence of events in the catalytic reaction, are treated using more accurate QM methods, while the remainder of the protein is treated using less rigorous, but more computationally efficient MM methods. For more extensive reviews of the QM/MM technique see the following references (Bruice, 2006; Lin and Truhlar, 2007; Senn and Thiel, 2007; Senn and Thiel, 2009).

QM/MM calculations have been applied to elucidate different phenomena in biomolecular systems, such as chemical processes, as it is able to show the structure and energetic of enzyme reactions. Such simulations have been used to model the mechanisms of action of numerous proteins.

QM/MM calculations have been used to investigate different phenomena of heme-containing proteins. The cytochrome P450 superfamily is one of the enzymes that perform a variety of essential functions in all life forms and play a big role in drug metabolism. The key step of substrate oxidation involved in the reaction mechanism of compound I in human isoforms of cytochrome P450 has been studied (Bathelt *et al.*, 2005). The results remarkably reveal that compound I is similar in all isoforms suggesting that the presence of substrate has no effect. QM/MM studies on the mechanisms of compound I formation in the catalytic cycle of cytochrome P450cam and chloroperoxidase has been elucidated (Chen *et al.*, 2008; Zheng *et al.*, 2006). The results of these studies show that cpd I is 1.5 kcal/mol more stable than cpd 0 in chloroperoxidase and 8.0 kcal/mol in P450cam. The effects of substrate, protein environment, and proximal ligand mutation on compound I and compound 0 of chloroperoxidase were then determined in order to explore the root cause of compound I stability of P450cam (Lai *et al.*, 2009).

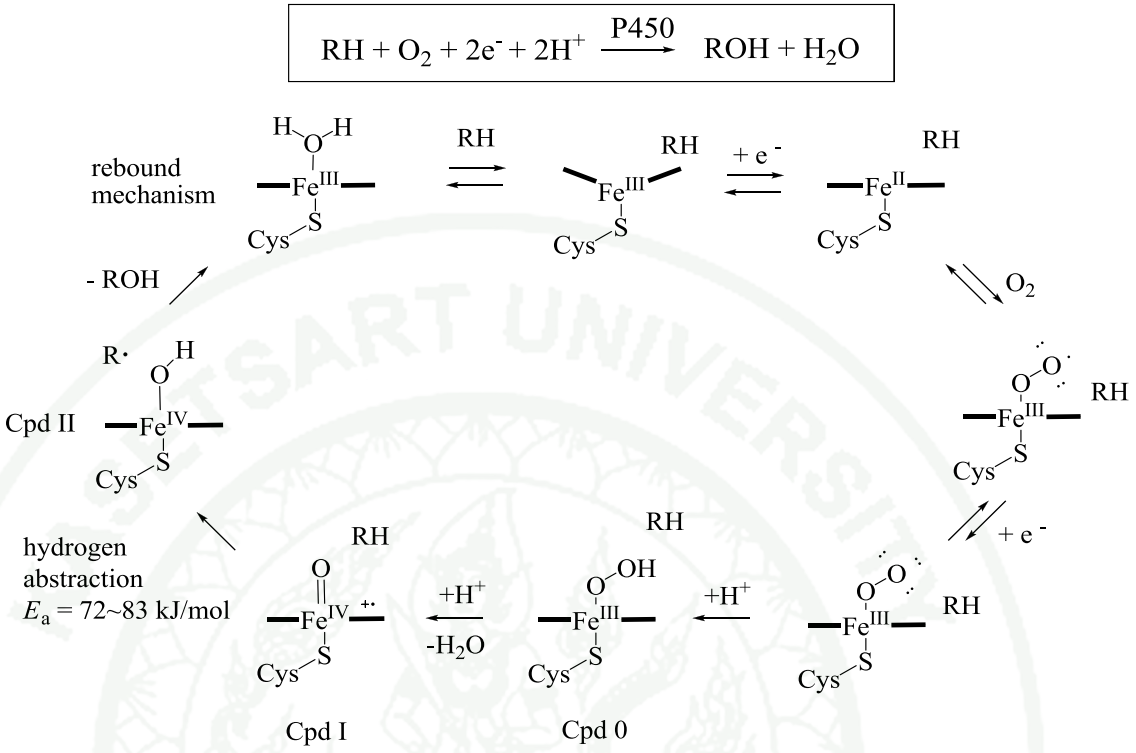


Figure 4 Catalytic cycle for the hydroxylation of an aliphatic hydrocarbon substrate RH in typical cytochrome P450 (Montellano, 2005).

Non-heme iron enzymes are also modeled using the ONIOM combined QM/MM method. A binding of dioxygen in isopenicilin N synthase (IPNS) is calculated showing that the effect of binding energies is found to be consistently depending on the coordination mode of O_2 (side-on or end-on) and the stabilization of Van der Waals interactions between dioxygen and the surrounding enzyme (Lundberg and Morokuma, 2007). Transition state of IPNS has been also determined and it was highlighted that the main catalytic effect of metal enzyme catalysis has come from the metal center while the reactivity to achieve high product specificity has controlled by the protein (Lundberg *et al.*, 2009; Lundberg *et al.*, 2008; Vreven *et al.*, 2006b).

ONIOM method was also applied to study cyclin dependent kinases (CDKs). The CDKs are a class of enzyme, which play a fundamental role in cell cycle regulation (Malumbres and Barbacid, 2005; Morgan, 1995; 1997; Norbury and Nurse, 1992). Prediction of binding modes and potency of the inhibitors on guanine derivatives and cyclin dependent kinase 2 (CDK2) have been studied by docking and ONIOM calculations (Alzate-Morales *et al.*, 2009). The interaction obtained from ONIOM calculations shows clear advantages; (a) the most active compounds are distinguished in a proper way from non-active ones and (b) the results show the most important contributions, the H-bond interactions, in this protein-ligand interaction. Additionally, the QM/MM interaction energy is used as a predictor of the biological activity of protein-ligand interactions in CDK2 inhibitors (Alzate-Morales *et al.*, 2007). The computational results reveal that the QM/MM interaction energy is strongly correlated to the biological activity and can be used as a predictor, at least within a family of substrates.

QM/MM calculations are also applied to calculate spectroscopic and investigate excited states properties. Mössbauer spectroscopic properties have been used to elucidate the oxidation state of metal enzymes as it has helped to analyze their spectra. For example, this technique is used to elucidate the oxidation state intermediates of iron-containing enzymes in cytochrome P450 (Schöneboom *et al.*, 2005) and non-heme iron dioxygenae (Sinnecker *et al.*, 2007). Vertical electronic excitation energies are also performed using QM/MM methods to study in groups of photoactive proteins as it has shown that the protein environment normally affects the spectroscopic properties. For instance, a study of protonation states of chromophore in photoswitching fluorescent protein dronpa (Li *et al.*, 2009) and the mechanism of color turning in rhodopsins (Hoffmann *et al.*, 2006) can be found in the applications of this category.

QM/MM methods have been applied to study the enolase family member phosphoenolpyruvate. This protein contains two Mg^{2+} ions in the active site and the mechanism involves the two separate CH proton abstraction steps. Liu et al. (Liu *et al.*, 2000) reported QM/MM free energy perturbation barrier heights of ~13.1 kcal/mol for the initial proton abstraction step, resulting in a stable enolate intermediate of ~5.0 kcal/mol. Decomposition of the intermediate by abstraction of a further proton from the β-carbon was found to require ~9.0 kcal/mol. The stability of enolate intermediates has also been investigated in the unrelated protein citrate synthase (Kamp *et al.*, 2008). They find that proton abstraction from the α-carbon of oxaloacetate requires ~10.2 kcal/mol and results in a stable intermediate ~8.0 kcal/mol higher in energy than the reactants. While the mechanisms differ considerably from MLE, the results discussed above serve as useful benchmarks for studies on MLEs.

MATERIALS AND METHODS

Materials

1. Hardware

1. LCAC super computer cluster (Chemistry Department, Faculty of Science, Kasetsart University, Bangkok)

2. High performance workstations; Intel core I7, 4 CPU, 2.80 GHz, RAM 8.0 GB (Chemistry Department, Faculty of Science, Kasetsart University, Bangkok)

2. Software

1. Gaussian 03
2. Gauss View 03
3. Discovery Studio 2.5
4. Discovery Studio Visualizer 2.0
5. OpenBabel
6. Gvim 7.2
7. Statistica 10
8. ChemBioDraw Ultra 11.0

Methods

1. Set-up of the system

The crystal structures of *anti*-MLE (PDB code: 3DG6) and *syn*-MLE (PDB code: 3DGB) were downloaded from RCSB protein data bank and prepared as follows. Cofactors, ions, and water molecules beyond 15 Å of the active site were deleted. Missing side chain data from both PDB structures as well as hydrogen atoms were added using Discovery Studio 2.5 according to the CHARMM 22 forcefield. The

protonation states of ionizable residues were determined by visual analysis. Ligand charges were determined using the AM1BCC method and parameters according the Accelrys CHARMm forcefield. Both proteins were solvated in a box of TIP3P water with a minimum distance of 7 Å between the protein and box edge (i.e. 14 Å between proteins in a periodic box). Counterions were added to neutralize the system. Default non-bonded cut-offs of 12 Å were used in all MM simulations.

Due to the difficulty in accurately simulating metalo-proteins using MM methods, the Mg^{2+} ion, its three chelating carboxylate groups and one water molecule were harmonically restrained to their X-ray positions during all of the MM preparation steps. MM optimization was achieved in three distinct steps. All optimizations were performed in Discovery Studio 2.5 (2010a) using the smart optimizer conditions and an RMS gradient below 0.1 kcal/mol. These were; (1) optimization of hydrogen atoms only followed by (2) optimization of all amino acid side chain atoms and solvent molecules and finally (3) optimization of all atomic coordinates. This setup is equivalent to the default protein preparation procedure to prepare protein X-ray structures for Docking, Molecular Dynamics or QM/MM in the modeling package Maestro (2010b).

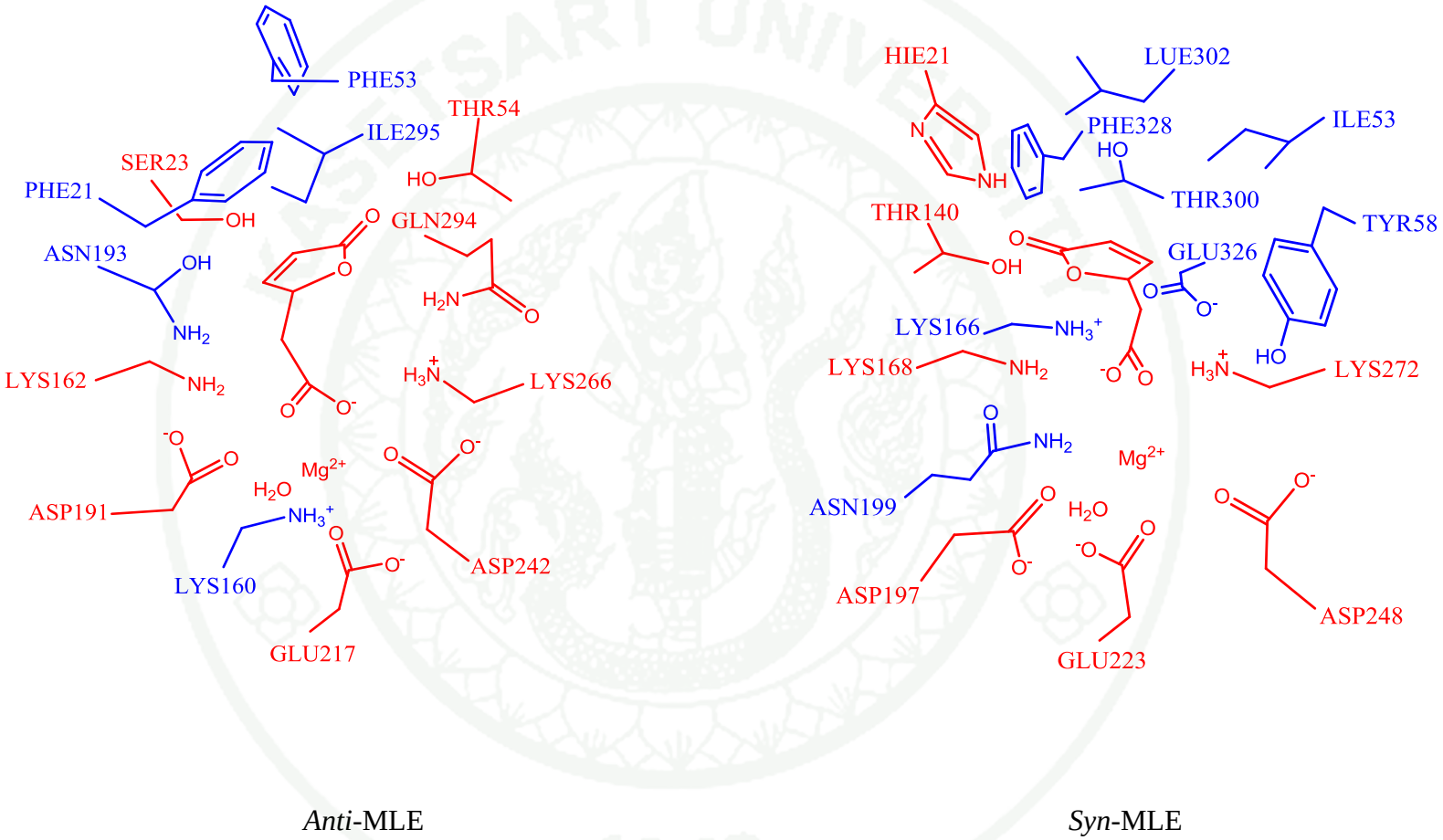
The MM optimized coordinates were then used in a subsequent short molecular dynamics (MD) step to help minimize any high energy contacts that are often present in X-ray protein structures. Atoms beyond 10 Å of the active site were harmonically restrained. MD was performed in two stages; (a) heating from T=0 to 300 K over 200 ps. (b) equilibration for 800 ps. Simulations were performed using the default CHARMM settings in Discovery Studio 2.5. These stages include a time step of 1.0 fs., NVT conditions, 12 Å non-bonded cut-offs and particle mesh Ewald (Darden *et al.*, 1993). The flexible atoms from the final MD step were subsequently re-optimized and used as input for QM/MM calculations.

2. QM/MM calculations

In this study the hybrid quantum mechanics/molecular mechanics (QM/MM) calculations were applied to probe aspects of MLE function and gain further insight into the catalytic events in both *anti*- and *syn*-MLEs that lead to the observed stereochemical differences. The QM/MM calculations performed here rely on the ONIOM methodology (Dapprich *et al.*, 1999; Vreven *et al.*, 2006a) using the electrical embedding scheme. Here the total QM/MM energy of the system is computed in a subtractive fashion as given in the following equation; QM energy of the active site region, or "model", plus the MM energy of the "real" or whole protein system, minus the MM energy of the model region.

$$E_{QM/MM} = E_{QM\text{-}EE\ (model)} + E_{MM(real)} - E_{MM(model)} \quad (1)$$

All QM/MM calculations were performed using the ONIOM methodology developed by Morokuma and co-workers as implemented in Gaussian 03 (Frisch *et al.*, 2004). A QM region has been selected so that key polar residues that directly interact with the substrate over the course of the reaction are included explicitly. For *anti*-MLE 70 atoms are treated QM consisting of the side chains of; SER23, THR54, LYS162, ASP191, GLU217, ASP242, LYS266, GLN294, the Mg^{2+} ion and the substrate. For *syn*-MLE 76 atoms are treated QM consisting of the side chains of; HIE21, THR140, LYS168, ASP197, GLU223, ASP248, LYS272, the Mg^{2+} ion and the substrate. The side chains of the following active site residues were treated flexibly with the rest of the system being fixed (*anti*-MLE: PHE21, PHE53, LYS160, ASN193, ILE295; *syn*-MLE: ILE53, THR58, LYS166, ASN199, THR300, LEU302, GLU326, PHE328), shown in figure 5. All water molecules were removed for computational efficiency except the water molecule that chelates Mg^{2+}.



Anti-MLE *Syn*-MLE

Figure 5 QM model (orange) and flexible residues in low region (blue) of *anti*- and *syn*-MLE proteins in the QM/MM calculations.

The M05 functional developed by Zhao et al. (Zhao and Truhlar, 2006; 2008) has been used in these calculations as it has been shown to be more effective than the popular B3LYP method for describing aspects of non-bonded interactions. The 6-31G(d) basis set was employed for geometrical optimizations. Single point energies on stationary points being characterized using the 6-311+G(d,p) basis set. The MM region was treated using universal Force field (UFF) in conjunction with CHARMM partial charges. The reaction coordinates of both possible bases, in both proteins, have been estimated by the stepwise variation of the C-O and C-H bonds between their reactant and product configurations. Due to the large memory requirements of Gaussian 03 with MM regions of this size (>5000 atoms) transitions states were characterized as saddle points by doing a frequency calculations of the optimized QM region coordinates only. ONIOM optimization has been performed using default settings; fixed link atom positions and involves the electrical embedding of MM charges into the QM calculation.

RESULTS AND DISCUSSION

In attempt to confirm the identity of the catalytic base in the *anti*- and *syn*-MLE proteins, the minimum energy structures for the reactants and both the *anti*- and *syn*-products in both proteins have been determined. As the products may be dictated by kinetic factors, the potential energy surface between reactant and products have also determined to estimate the reaction barriers and also to assess the possibility of an explicit enolate anion existing during the course of the reaction.

1. Energetic Features

The QM/MM calculations performed here on the two possible products formed by both *anti*- and *syn*-MLEs appear to confirm the identity of the catalytic bases in the reaction from both a structural and energetic perspective. However, it is also desirable to assess the reaction barriers associated with the two possible bases in each protein to rule out any differences that might affect the reaction products from a kinetic perspective.

1.1. Potential energy surface (PES)

The potential energy surface between reactants and both possible products in *anti*- and *syn*-MLE proteins were therefore summarized (Figure 6 and Figure 7). The reaction barriers of each process have been estimated and these results are summarized in figure 8. The PES scans were obtained by scanning the C-O and C-H distances corresponding to the distance d_1 and d_2 in figure 1. The reactants and TS species were obtained by scanning from the structures of products as the reaction is reversible. The reaction mechanism will be discussed later (Section 1.3).

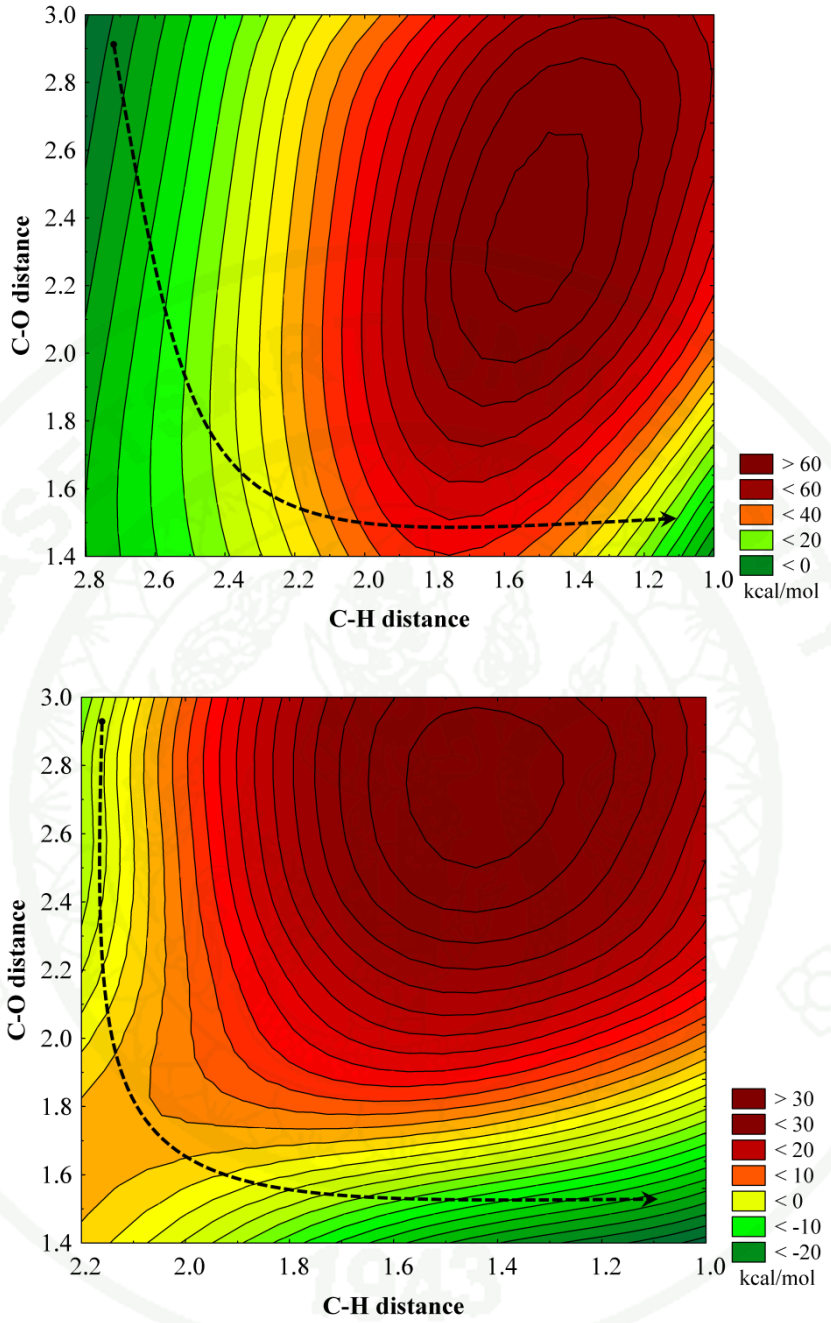


Figure 6 Potential energy surfaces for the *anti-syn* (top) and *anti-anti* (bottom) catalysed reactions. In the former, LYS266 acts as the catalytic acid and in the latter, LYS162. C-O and C-H distances correspond to d_1 and d_2 in figure 1.

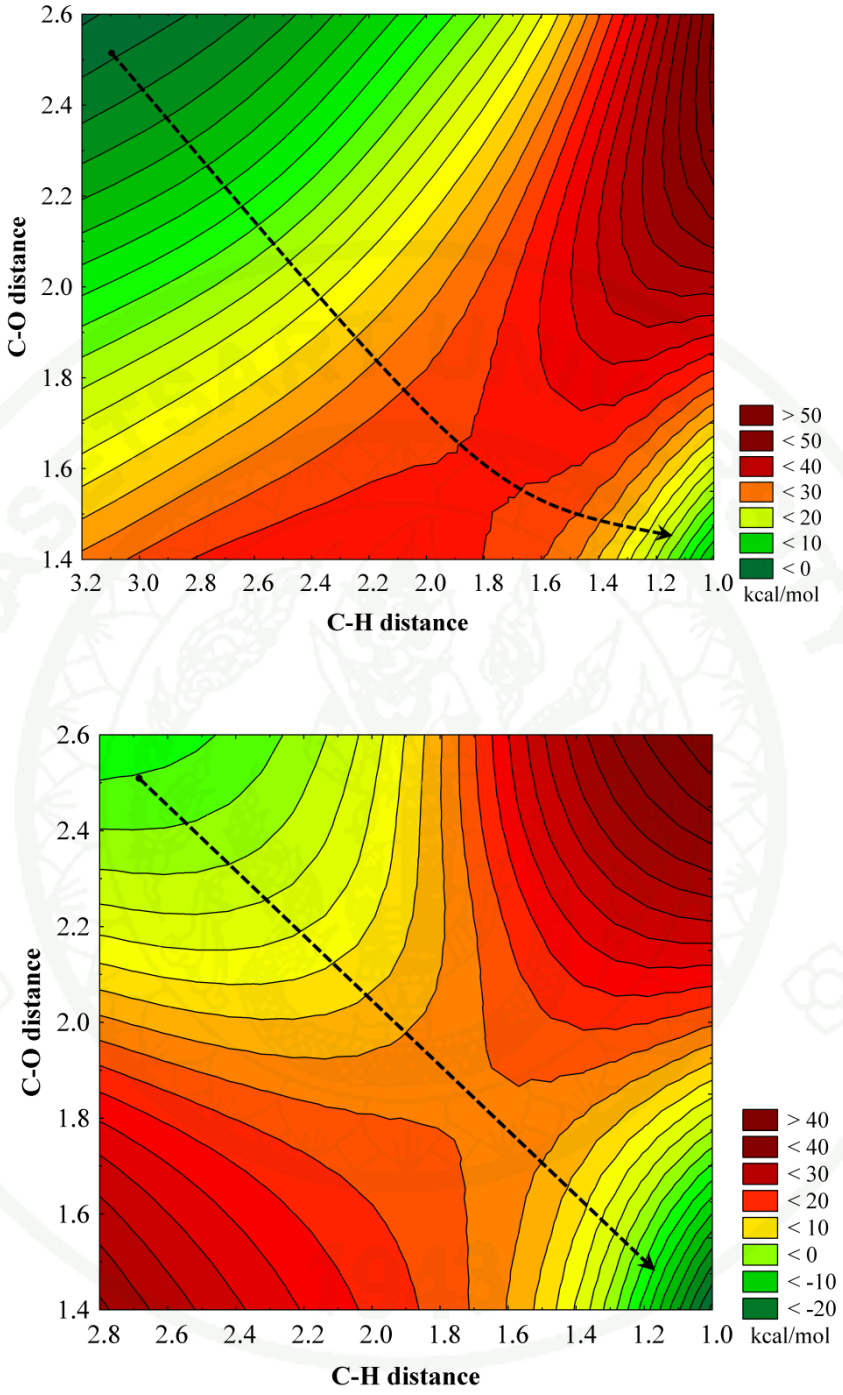


Figure 7 Potential energy surfaces for the *syn-anti* (top) and *syn-syn* catalysed reactions. In the former, LYS272 acts as the catalytic acid and in the latter, LYS168. C-O and C-H distances correspond to d_1 and d_2 in figure 1.

1.2. Minimum energy pathway

The optimized and single point QM/MM energies are explored and summarized in figure 8. The optimized energies show that the barrier to reaction associated with the *anti*-MLE *anti*-Product *(anti-anti* Product) is 11.6 kcal/mol lower than the corresponding *anti*-MLE *syn*-Product *(anti-syn* Product). Similarly, a 13.0 kcal/mol difference between the *syn-syn* and *syn-anti* products of *syn*-MLE was observed. Likewise, the single point energies show that the barrier of *anti-anti* Product is 16.4 kcal/mol lower than *anti-syn* Product barrier. Correspondingly, a 20.1 kcal/mol difference between the *syn-syn* and *syn-anti* products of *syn*-MLE was examined. These results clearly suggest that the basic residues involved in the reaction are located at the end of the 2nd strand of both *anti*-MLE and *syn*-MLE, corresponding to Lys-162 and Lys-168 respectively.

Assessment of the relative QM/MM optimized energies shows that the *anti-anti* product is lower than the corresponding *anti-syn* product by 26.1 kcal/mol. In addition, the energetic differences between the *syn-syn* product and *syn-anti* product show that the former is favored by 19.8 kcal/mol. These results are also correlated well with single point calculations, in which a 26.4 and 27.6 kcal/mol are observed to confirm that *anti-anti* and *syn-syn* Products are lower in energy than the corresponding products, respectively. Thus, from a purely thermodynamic perspective it would be expected that Lys-162 from *anti*-MLE and Lys-168 from *syn*-MLE acting as the catalytic base since it leads to the lowest energy product and will lead to the known stereochemical outcome.

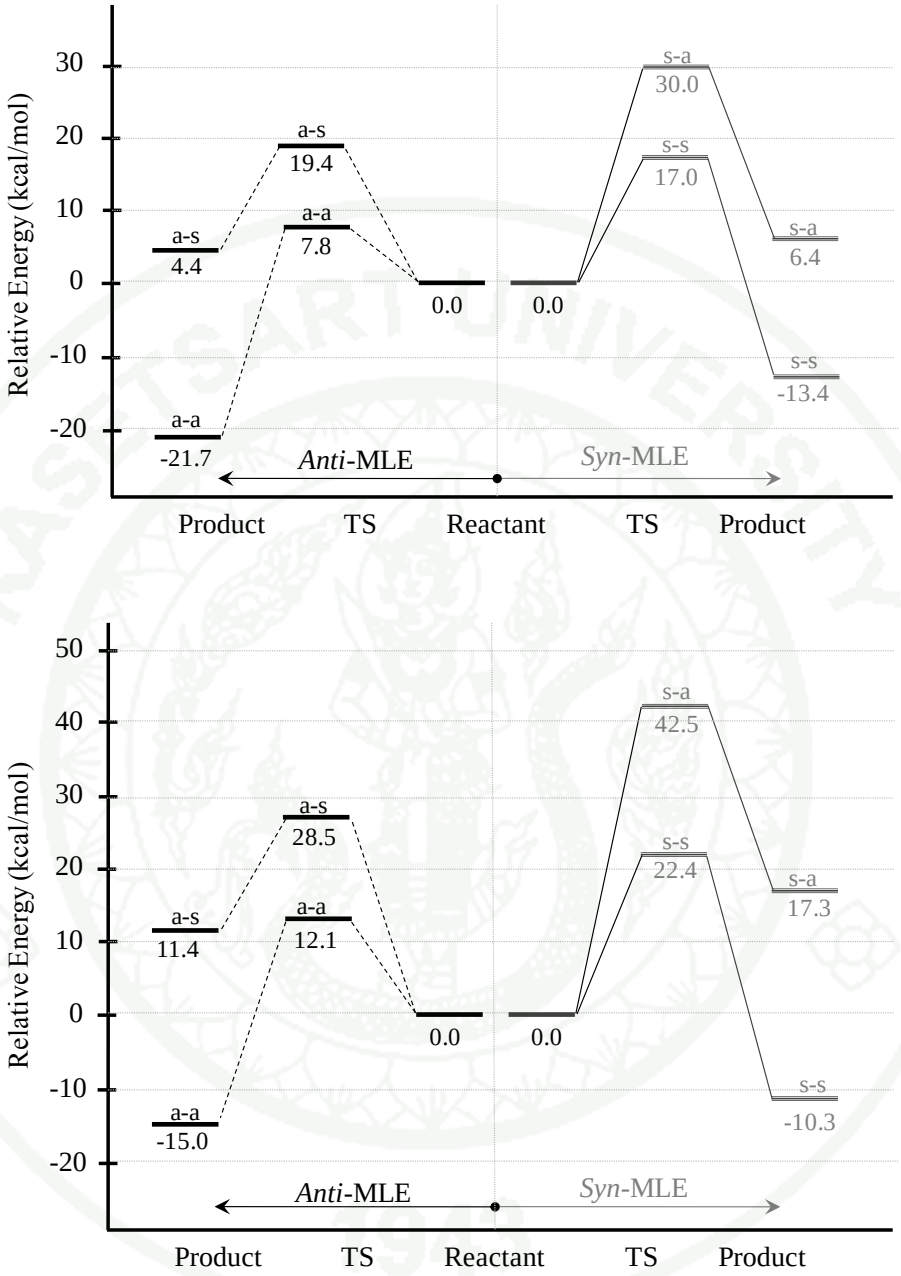


Figure 8 Relative M05/6-31G(d):UFF optimized energies (top) and M05/6-311+G(d,p) single point energies (bottom) for the *anti*- and *syn*-products in both *anti*- and *syn*-MLEs. a-a refers to *anti*-product and a-s refers to *syn*-product of *anti*-MLE., For *syn*-MLE, s-s refers to *syn*-product and s-a refers to *anti*-product.

1.3. Reaction mechanism

The potential energy surfaces of *anti-anti* and *syn-syn* catalyzed reactions are used to discuss. From an analysis of the potential energy surface (Figure 6 and Figure 7) it is also apparent that the reaction proceeds along a more stepwise pathway with C-O bond breaking occurring before proton transfer to Lys-162 regarding to the conversion of muconolactone to *cis,cis*-muconate. In contrast the process in *syn*-MLE appears to proceed along a concerted route with both C-O bond breaking and proton transfer to Lys-168.

1.4. Enolate anion

The theoretical DFT QM/MM models employed have helped to confirm the identity of the base in the reaction. Therefore, an understanding whether an explicit enolate anion exists in the active site of either protein were also interested. The relevant structure obtained from the QM/MM potential energy scans (i.e. the structure having a C-O bond formed but a long C-H bond) was taken to determine whether the enolate was a stable stationary point within the protein. This constrained structure was subsequently fully optimized QM/MM using the conditions as used for the other stationary points obtained in this study. In the case of both *syn*- and *anti*-MLE this enolate-like structure decomposed to the muconolactone by accepting a proton from the active site Lys residue which leads to the forming of the C-O bond. While this result would appear to suggest that an explicit enolate anion does not exist in these proteins further work is needed to prove whether this high energy structure is truly a stable stationary point or not. These calculations would necessitate additional polarization and diffuse functions as well as full frequency analyses given that the structure is likely to occupy a shallow energy minimum.

2. Geometrical features

2.1. Reaction species

The QM/MM optimized active site regions of the two possible products of both MLE proteins are obtained. The transition state (TS) species are collected from the PES scans. The reaction species of *anti*-MLE are shown in figure 9 and figure 10. Similarly, figure 10 and figure 11 reveal the reaction species of *syn*-MLE. Listed in Table 2 and 3 are the key geometrical parameters associated with the stationary points.

Assessment of the key geometrical parameters of *anti*-MLE is firstly discussed. Table 2 reveals that *anti*-Product is coordinated with Mg^{2+} and Gln-294 stronger than *syn*-Product (Mg-O(sub) 2.36 and 2.49 Å and O(sub)-H(GLN294) 1.88 and 2.01 Å, respectively. The coordination of *anti*-TS is also found to be better than *syn*-TS. The water molecule which coordinates with Mg^{2+} is found to be consistent at around 2.0 Å. The Asp-191, Glu-217 and Asp-242 are chelated with Mg^{2+} with the distances 2.3, 2.0 and 2.0 Å, respectively.

Similarly, the key geometrical parameters of *syn*-MLE (Table 3) are then discussed. The Mg-O distances of *syn*-Product are 2.16 and 2.49 Å while it is found to be 2.03 and 3.37 Å for *anti*-Product. It can be noticed that the carboxylate of *syn*-Product are both coordinated with Mg^{2+}, whereas only single Mg-O coordination found in *anti*-Product (Figure 11 and Figure 12). The distance of O(sub)-H(HIE21) is found to be 1.78 Å for reactant specie and around 1.86 Å for the rest. The hydrogen bond of reactant is stronger as because five members ring is opened to form *cis,cis* muconate, which produces the other carboxylate group. The conserved residues (Glu and Asp) coordinated with essential Mg^{2+} ion are found to be comparable to those coordination found in *anti*-MLE.

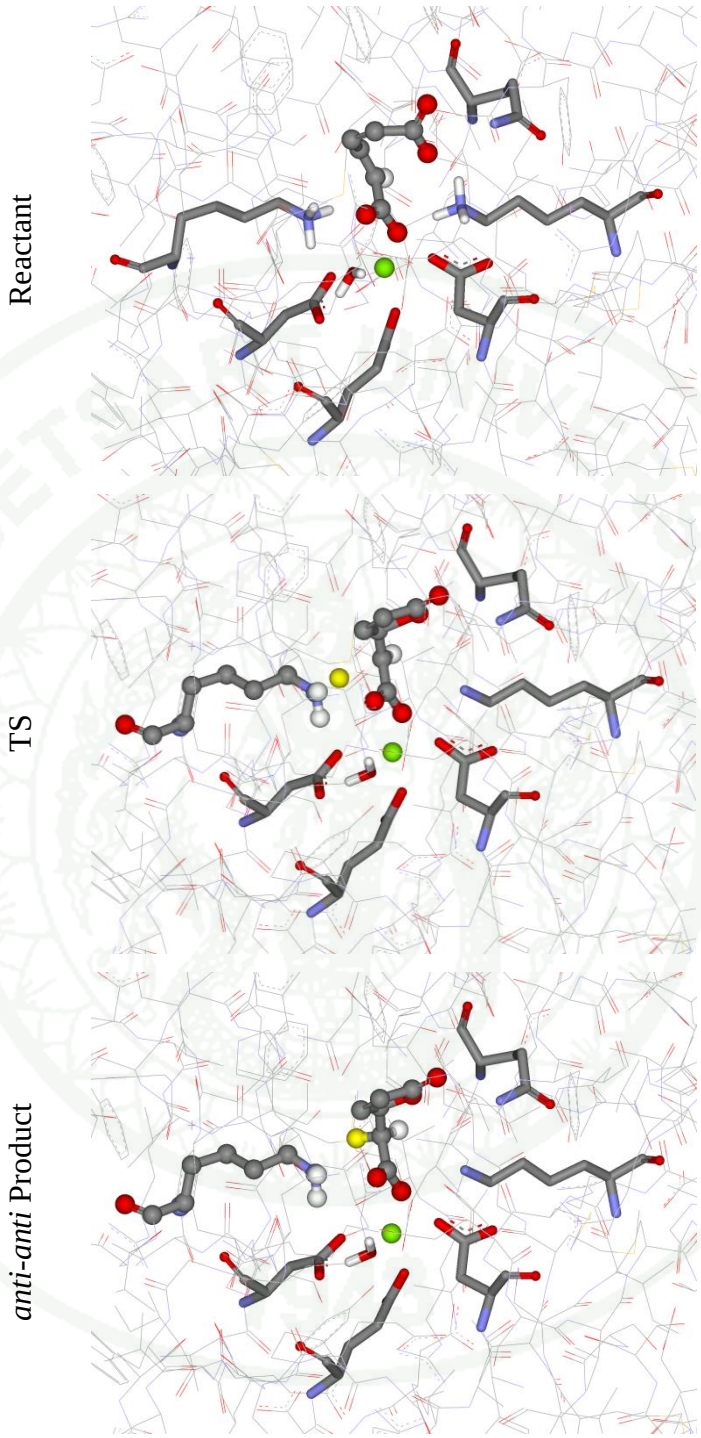


Figure 9 QM/MM optimized active site regions of *anti*-Product of *anti*-MLE (a-a pathway).

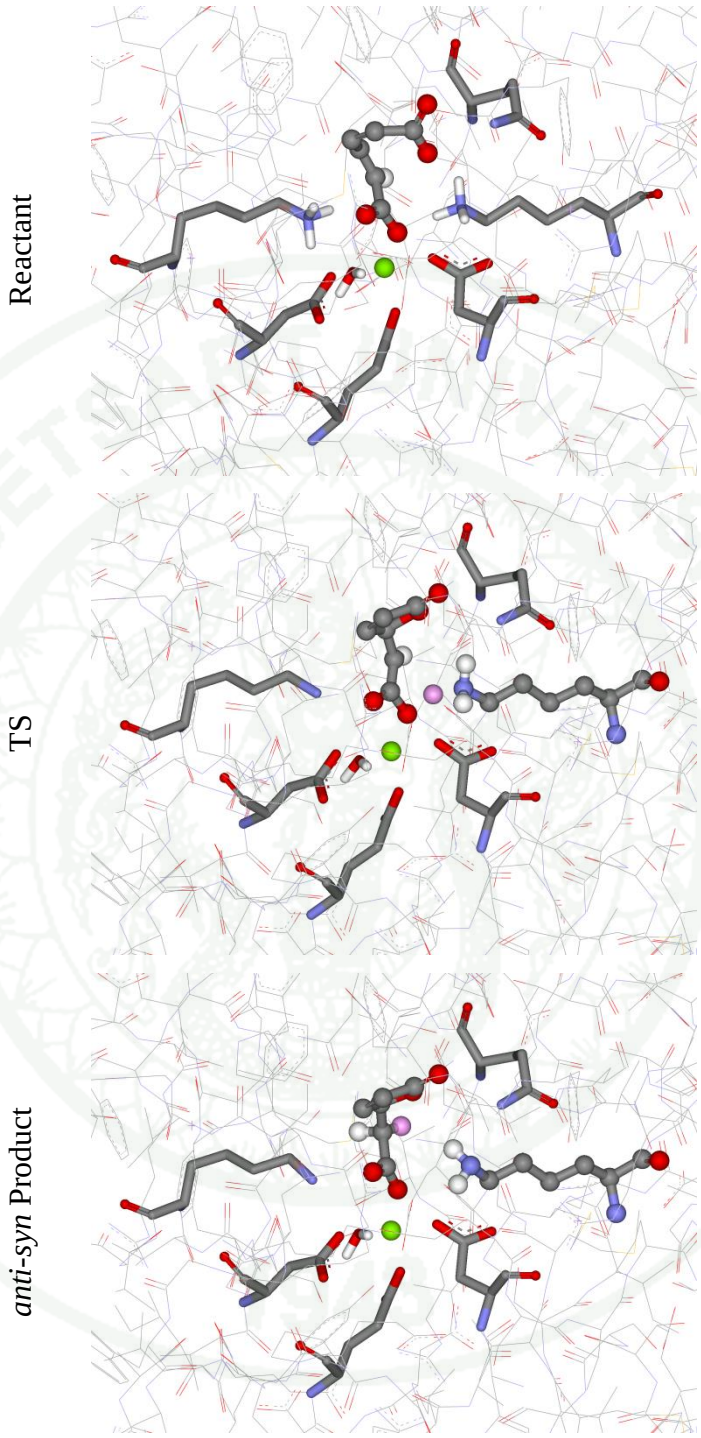


Figure 10 QM/MM optimized active site regions of *syn*-Product of *anti*-MLE (a-s pathway).

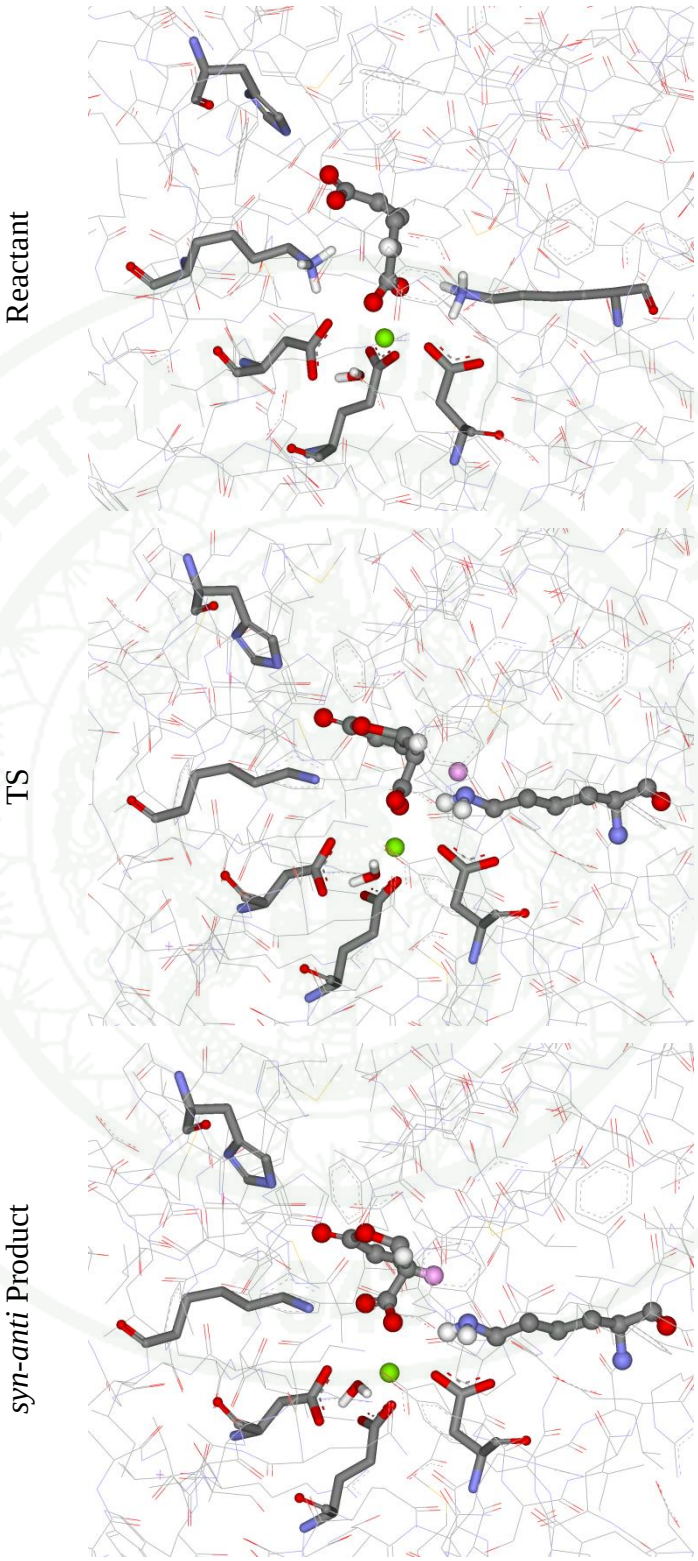


Figure 11 QM/MM optimized active site regions of *anti*-Product of *syn*-MLE (s-a pathway).

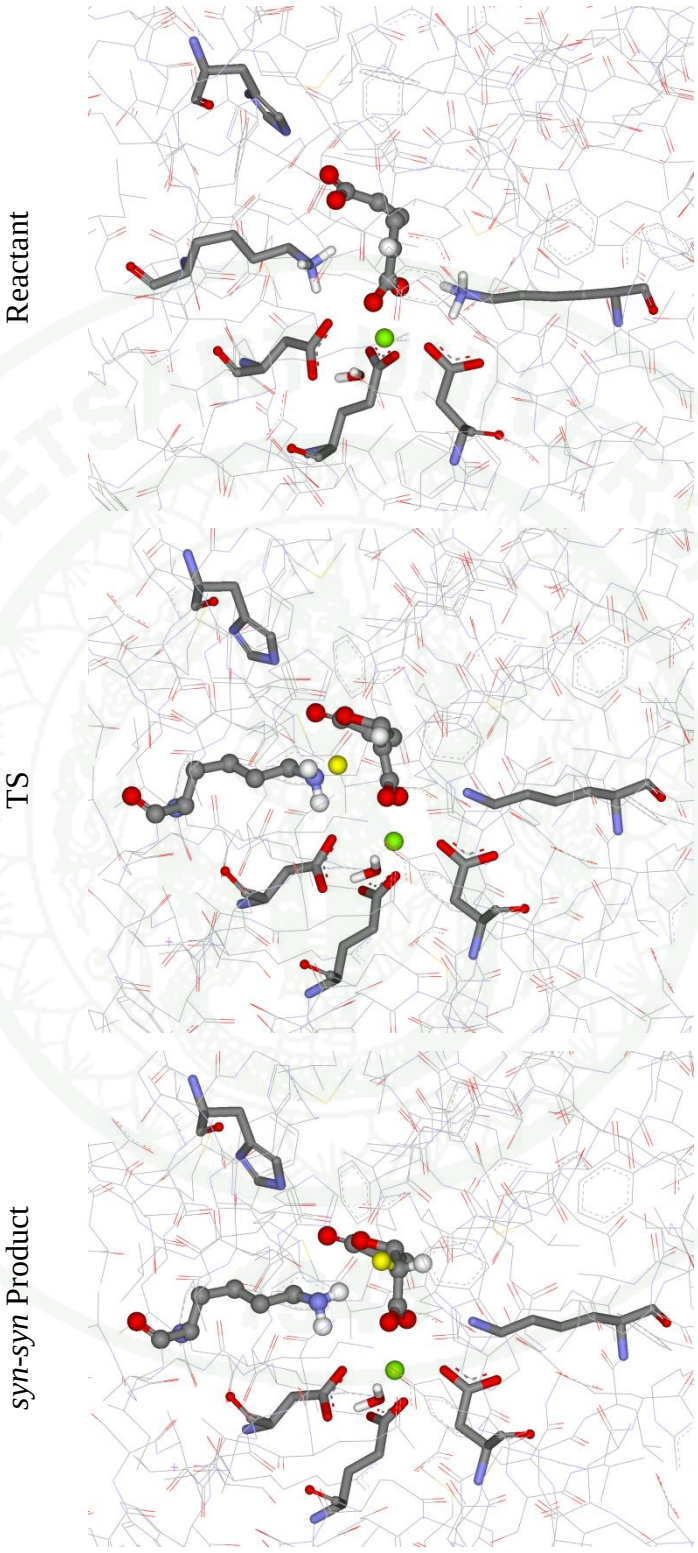


Figure 12 QM/MM optimized active site regions of *syn*-Product of *syn*-MLE (s-s pathway).

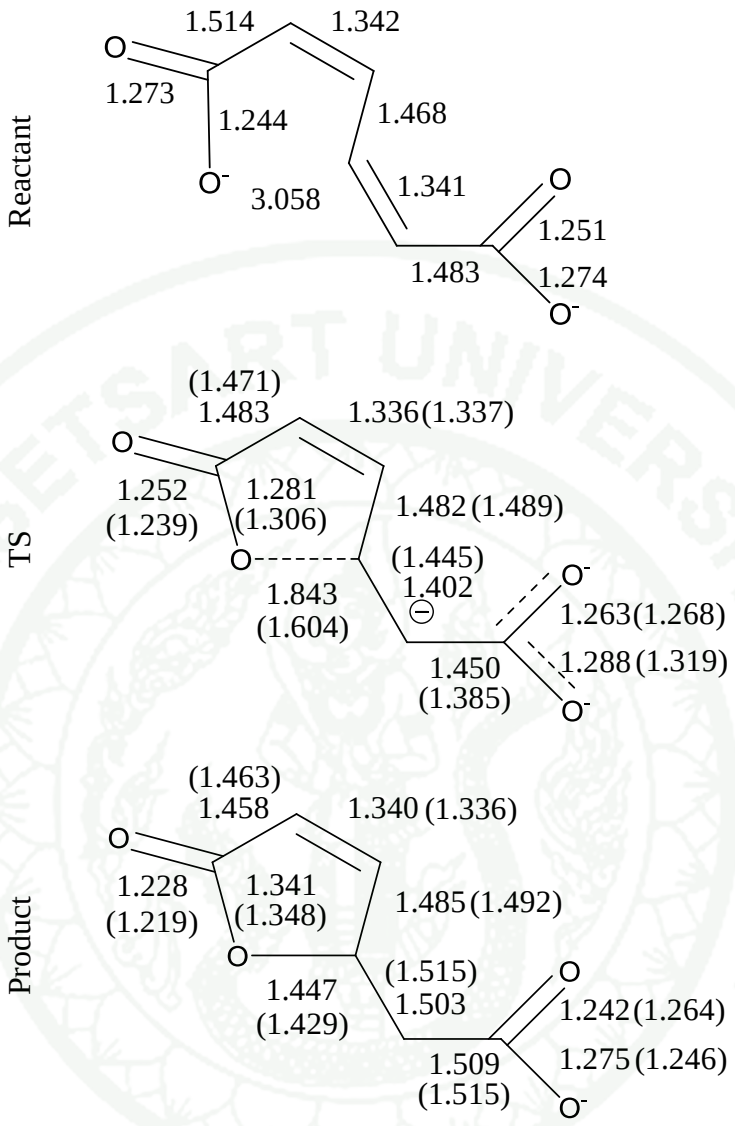


Figure 13 Geometrical parameter changes in heavy atoms of substrate of a-a and a-s (in parenthesis) pathways of *anti*-MLE.

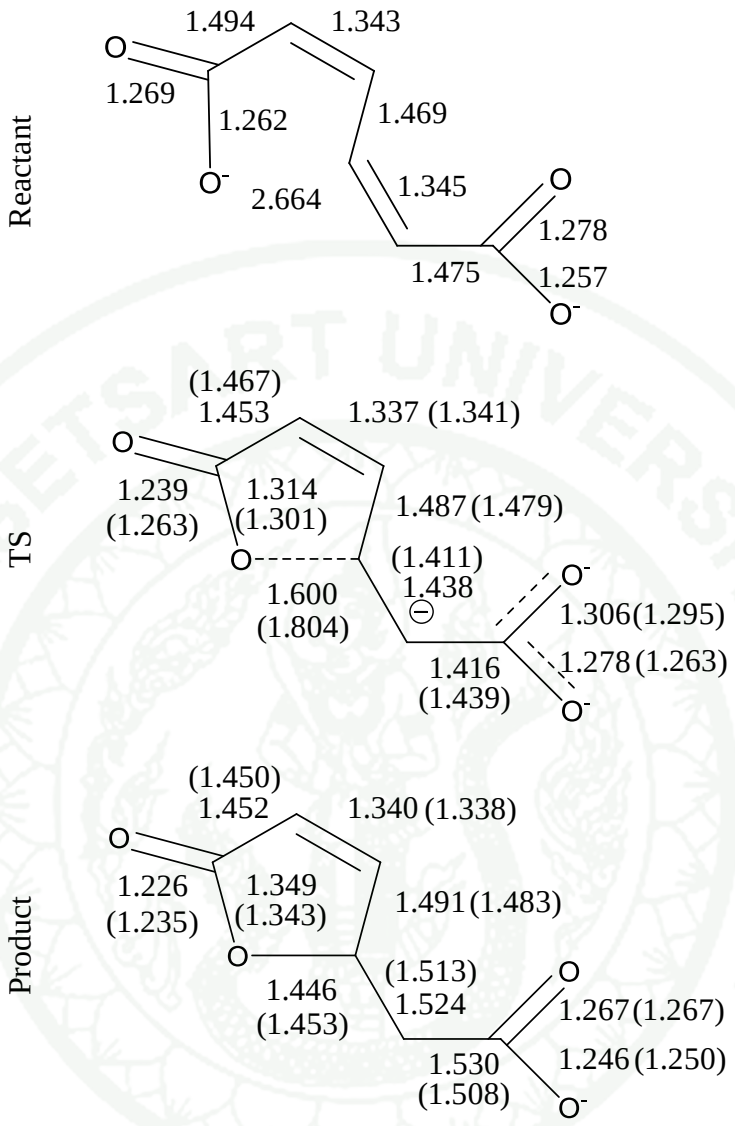


Figure 14 Geometrical parameter changes in heavy atoms of substrate of s-s and s-a (in parenthesis) pathways of *syn*-MLE.

Table 2 Predicted and experimental distances observed for the conversion of *cis,cis*-Muconate to Muconolactone during the *anti*-MLE catalyzed reactions. Distances are reported in Å.

Distances	Structures					
***Anti*-MLE**	Reactant	*syn*-TS	*syn*-Product	*anti*-TS	*anti*-Product	*X-ray (Product)*
Mg−O(ASP191)	2.35	2.47	2.30	2.41	2.22	2.08
Mg−O(GLU217)	2.02	2.02	2.00	2.01	2.02	1.96
Mg−O(ASP242)	2.00	2.01	1.95	2.02	2.05	2.19
Mg−O(Sub)	2.31	2.29	2.23	2.17	2.23	2.16
Mg−O(Sub)	2.21	2.12	2.49	2.22	2.36	2.29
Mg−O(WAT)	2.06	2.06	2.10	1.99	1.99	2.20
C(Sub)−N(LYS162)	3.11	3.39	3.45	2.99	3.04	3.03
C(Sub)−N(LYS266)	3.66	3.03	3.87	3.56	3.66	3.57
C(Sub)−HN(LYS162)	2.25	2.47	2.67	1.90	1.10	-
C(Sub)−HN(LYS266)	2.92	2.10	1.09	2.81	2.90	-
O(Sub)−H(GLN294)	1.86	1.90	2.01	1.84	1.88	-

Table 3 Predicted and experimental distances observed for the conversion of *cis,cis*-Muconate to Muconolactone during the *syn*-MLE catalyzed reactions. Distances are reported in Å.

Distances	Structures					
***Syn*-MLE**	Reactant	*syn*-TS	*syn*-Product	*anti*-TS	*anti*-Product	*X-ray (Product)*
Mg−O(ASP197)	2.59	2.62	2.27	2.72	2.65	2.34
Mg−O(GLU223)	1.99	2.02	2.01	2.02	2.00	2.33
Mg−O(ASP248)	2.04	2.03	2.10	2.02	1.95	2.24
Mg−O(Sub)	2.15	2.07	2.16	2.16	2.03	2.42
Mg−O(Sub)	2.22	2.22	2.49	2.11	3.37	2.96
Mg−O(WAT)	1.99	2.00	1.99	1.98	1.97	1.79
C(Sub)−N(LYS168)	3.66	3.07	3.60	4.24	4.26	4.22
C(Sub)−N(LYS272)	3.55	4.02	3.81	2.72	3.31	3.45
C(Sub)−HN(LYS168)	2.80	2.00	1.09	3.33	3.31	-
C(Sub)−HN(LYS272)	3.13	3.77	3.42	1.93	1.10	-
O(Sub)−H(HIE21)	1.78	1.86	1.88	1.89	1.82	-

2.2. X-ray versus QM/MM structures

The differences between the two different product structures for each protein were then analyzed, which differ only in terms of the position of a single proton, and compare these to the corresponding heavy atom coordinates reported in the respective X-ray structure PDB files. The expectation is that the most thermodynamically favorable QM/MM coordinates will be more similar to the original X-ray structure and that the corresponding base will give rise to the expected stereochemical outcome.

Analysis of the RMSDs of the QM/MM optimized active site regions of the two possible products could potentially be used to identify which base is involved in the catalytic reaction. One would expect that the optimized product that displays the lowest RMSD to the experimental X-ray structure would identify which base is involved in the reaction. While the optimized QM/MM structures of both products for both *anti*- and *syn*-MLE proteins, display C_α RMSDs of < 0.05 Å to the corresponding X-ray structure, the flexible QM and MM heavy atoms in the QM/MM system understandably display larger differences in structure. Figure 15 and figure 16 display the overlay of (1) the original X-ray coordinates, (2) the MD output structure and (3) the QM/MM optimized geometries, highlighting the rather small differences between them. For *syn*-MLE, it was found that the product structure involving the base which leads to the *syn*-product (Lys-168 located on the 2nd strand) has a lower RMSD than that formed with Lys-272, which results in the *anti*-product (RMSDs of 0.36 and 0.56 Å). Furthermore, it was also found that the *anti*-product formed within the *anti*-MLE active site has a lower overall RMSD that that of the *syn*-product (0.58 and 0.63 Å) suggesting that the 2nd strand lysine, Lys-162, is the base in this reaction and not Lys-266 located on the 6th strand.

Looking in more detail at the product structures (Table 2) was allowed to observe that the key interacts made in the *anti*-*anti* product are clearly in better agreement with the X-ray coordinates than the *anti*-*syn* product. However, for *syn*-MLE (Table 3) both product show distances considerably more variable from the X-ray structure. In fact, the configuration of the Mg^{2+} ion and its ligands in the *syn*-MLE QM/MM models is more similar to that observed for the *anti*-MLE products. It is possible that this artifact may have arisen from the fitting of the atomic solution of *syn*-MLE to its electron density. If fact, the Mg-O distances observed in the *syn*-MLE X-ray structure appear considerably longer than the experimentally expected values of ~2.1 Å (Plyasova *et al.*, 1998). Contrasting the two experimental *anti*- and *syn*-MLE X-ray structures (Table 2, Table 3, and Figure 2) it can be seen that Mg^{2+} coordination in the *syn*-MLE structure is much weaker than in the *anti*-MLE. Interestingly, the QM/MM optimized *syn*-structures of *syn*-MLE displays interactions that are more comparable to the *anti*-structures.

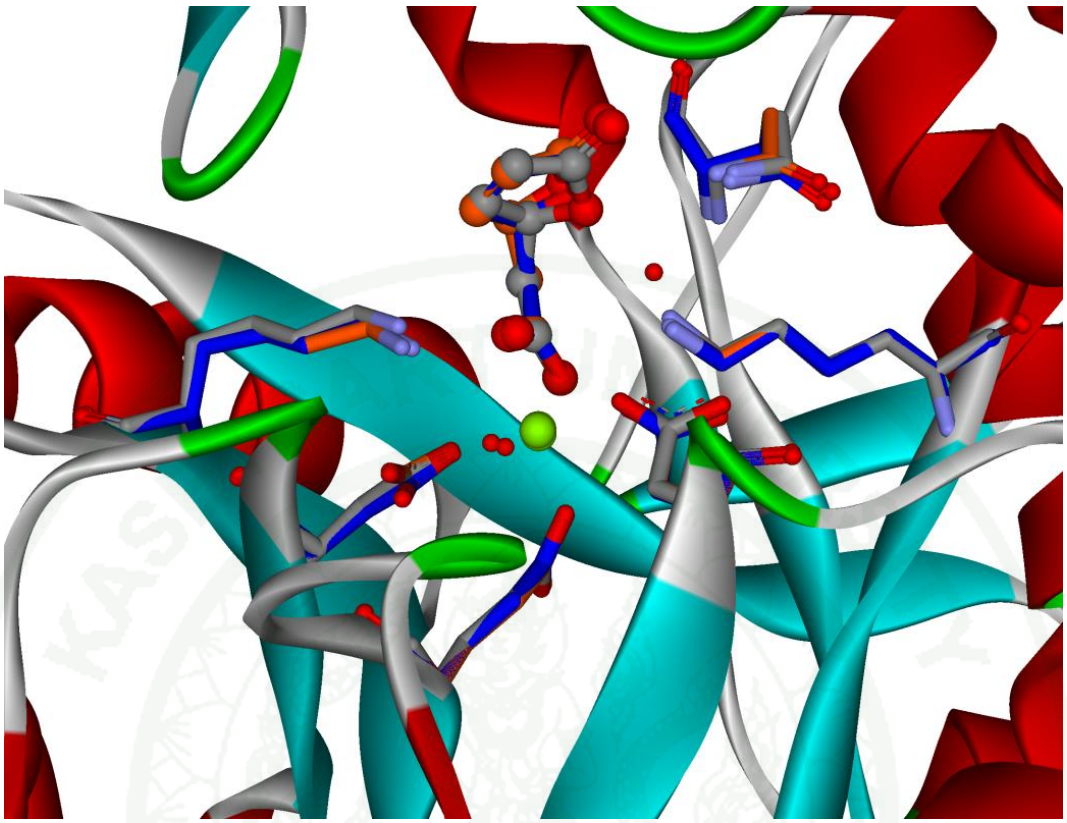


Figure 15 Superposition of the X-ray coordinates (grey), the MD output (orange) and the QM/MM optimized geometries (blue) of *anti*-MLE.

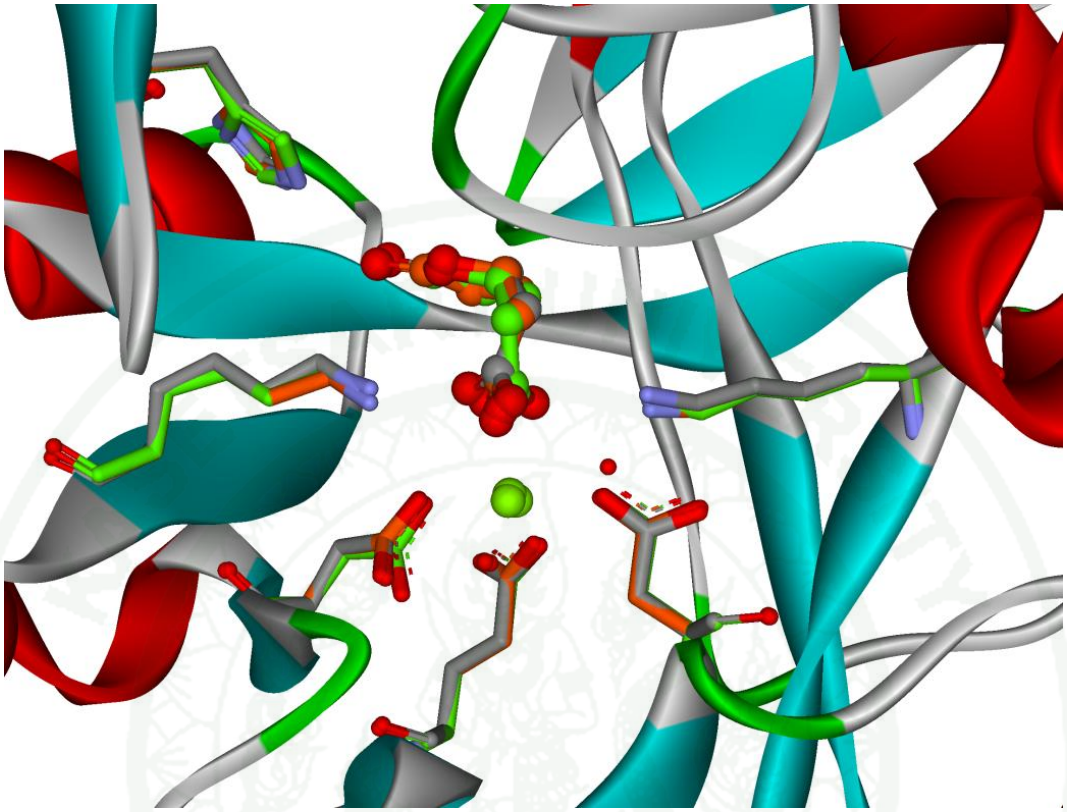


Figure 16 Superposition of the X-ray coordinates (grey), the MD output (orange) and the QM/MM optimized geometries (green) of *syn*-MLE (bottom).

To try and decipher the contribution electrostatics and van der Waals interactions, as well as protein pre-organization, gas phase single point QM calculations on each of the QM/MM optimized product geometries were subsequently performed. Analysis of the contributing terms to the QM/MM energy reveals that the Van der Waals term is essentially constant for structures obtained in the two different protein models. The gas phase QM single point energies show that both the *syn-syn* and *anti-anti* products are still preferred in the gas phase suggesting active site pre-organization is important (ΔE between *syn-syn* and *syn-anti* products are -26 and -33 kcal/mol at QM/MM and QM levels respectively; ΔE between *anti-anti* and *anti-syn* products are -20 and -6 kcal/mol at QM/MM and QM levels respectively). Interestingly, the inclusion of the protein electrostatic term has a much greater effect in stabilizing the *anti-anti* product over the *anti-syn* product since the difference in QM energy is just 6 kcal/mol in the gas phase calculations but 20 kcal/mol in the protein. In contrast, the difference in energy between the *syn-syn* and *syn-anti* product is somewhat larger in the gas phase than in the protein calculations. This might suggests that the two proteins achieve their product selectivities in subtly different ways, *anti*-MLE relying more on organizing its active site to favor the *anti*-product conformation and *syn*-MLE by preferentially stabilizing the *syn*-product as a result of its particular electrostatic characteristics.

CONCLUSION AND RECOMMENDATION

Conclusion

The application of DFT QM/MM calculations on *cis,cis*-Muconate lactonizing enzymes is discussed to determine the reaction energetics associated with the conversion of *cis,cis*-Muconate to Muconolactone in both *anti*- and *syn*-MLEs. The proposal have been investigated that the base involved in the catalytic reaction is the Lys residue found at the end of the 2^{nd} strand, rather than a Lys residue 6^{th} which is almost equally well positioned.

The QM/MM results show that the expected Muconolactone *anti*-product derived from *anti*-MLE is (1) closer to the X-ray structure in terms of RMSD, (2) lower in energy and (3) has a lower barrier to reaction than the corresponding *syn*-product. These results also show that the *syn*-product derived from *syn*-MLE also displays a lower RMSD to the original X-ray coordinates, is lower in energy, and has a lower barrier to reaction than the corresponding *anti*-product. Thus, Lys-162 is found to results in the lowest energy reaction and Lys-168 for *syn*-MLE.

The theoretical calculation performed here have helped to confirm the identity of the basic residues involved in the MLE reaction originally proposed by Sakai et al through indirect means from an analysis of the products produced in experiments performed in deuterated solvents. The results show that although the basic Lys residues located on the 6^{th} strand in both MLEs are almost equally well positioned to accept a proton, however, it is the base located on the 2^{nd} strand that is thermodynamically and kinetically more favorable, as well as giving the optimized QM/MM product closest to the original X-ray structure.

Recommendation

The reactant geometries were performed using the protein conformation of product geometry of the substrate. The MD simulations might be needed to allow protein movement to obtain the suitable protein conformation of the reactant substrate.

LITERATURE CITED

Alzate-Morales, J.H., J. Caballero, A. Vergara Jague and F.D. GonzÃ¡lez Nilo. 2009. Insights into the structural basis of N2 and O6 substituted guanine derivatives as cyclin-dependent kinase 2 (CDK2) Inhibitors: prediction of the binding modes and potency of the inhibitors by docking and ONIOM calculations. **J. Chem. Inf. Model.** 49(4):886-899.

Alzate-Morales, J.H., R. Contreras, A. Soriano, I. Tuñon and E. Silla. 2007. A computational study of the protein-ligand interactions in CDK2 inhibitors: using quantum mechanics/molecular mechanics interaction energy as a predictor of the biological activity. **Biophys. J.** 92(2):430-439.

Asuncion, M., W. Blankenfeldt, J.N. Barlow, D. Gani and J.H. Naismith. 2002. The structure of 3-methylaspartase from clostridium tetanomorphum functions via the common enolase chemical step. **J. Biol. Chem.** 277(10):8306-8311.

Babbitt, P.C., M.S. Hasson, J.E. Wedekind, D.R.J. Palmer, W.C. Barrett, G.H. Reed, I. Rayment, D. Ringe, G.L. Kenyon and J.A. Gerlt. 1996. The enolase superfamily: a general strategy for enzyme-catalyzed abstraction of the α-protons of carboxylic acids. **Biochemistry** 35(51):16489-16501.

Bathelt, C.M., J. Zurek, A.J. Mulholland and J.N. Harvey. 2005. Electronic structure of compound I in human isoforms of cytochrome P450 from QM/MM modeling. **J. Am. Chem. Soc.** 127(37):12900-12908.

Bruice, T.C. 2006. Computational approaches: reaction trajectories, structures, and atomic motions. enzyme reactions and proficiency. **Chemical. Rev.** 106(8):3119-3139.

Chen, H., H. Hirao, E. Derat, I. Schlichting and S. Shaik. 2008. Quantum mechanical/molecular mechanical study on the mechanisms of compound I formation in the catalytic cycle of chloroperoxidase: an overview on heme enzymes. **J. Phys. Chem. B.** 112(31):9490-9500.

Dapprich, S., I. Komáromi, K.S. Byun, K. Morokuma and M.J. Frisch. 1999. A new ONIOM implementation in Gaussian98. part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. **J. Mol. Struct. (Theochem)** 462: 1-21.

Darden , T., D. York and L. Pedersen. 1993. Particle mesh ewald: an N•log(N) method for ewald sums in large systems. **J. Chem. Phys.** 98(12):10089-10092.

Drug Discovery Studio 2.5. 2010. **Accelrys, Inc, San Diego, CA, USA**

Frisch, M. J. T.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; 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. E.; 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 B.05. **Gaussian, Inc., Wallingford CT**

Gerlt, J.A. and P.C. Babbitt. 1998. Mechanistically diverse enzyme superfamilies: the importance of chemistry in the evolution of catalysis. **Curr. Opin. Chem. Biol.** 2(5):607-612.

Gerlt, J.A. and P.C. Babbitt. 2001. Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. **Annu. Rev. Biochem.** 70: 209-246.

Gerlt, J.A., P.C. Babbitt and I. Rayment. 2005. Divergent evolution in the enolase superfamily: the interplay of mechanism and specificity. **Arch. Biochem. Biophys.** 433(1):59-70.

Glasner, M.E., J.A. Gerlt and P.C. Babbitt. 2006. Evolution of enzyme superfamilies. **Curr. Opin. Chem. Biol.** 10(5):492-497.

Gulick, A.M., B.K. Hubbard, J.A. Gerlt and I. Rayment. 2000. Evolution of enzymatic activities in the enolase superfamily: crystallographic and mutagenesis studies of the reaction catalyzed by D-glucarate dehydratase from Escherichia coli. **Biochemistry Easton** 39(16):4590-4602.

Gulick, A.M., D.R.J. Palmer, P.C. Babbitt, J.A. Gerlt and I. Rayment. 1998. Evolution of enzymatic activities in the enolase superfamily: crystal structure of (D)-glucarate dehydratase from Pseudomonas putida. **Biochemistry** 37(41):14358-14368.

Gulick, A.M., D.M. Schmidt, J.A. Gerlt and I. Rayment. 2001. Evolution of enzymatic activities in the enolase superfamily: crystal structures of the L-Ala-D/L-Glu epimerases from Escherichia coli and Bacillus subtilis. **Biochemistry Easton** 40(51):15716-15724.

Hasson, M.S., I. Schlichting, J. Moulai, K. Taylor, W. Barrett, G.L. Kenyon, P.C. Babbitt, J.A. Gerlt, G.A. Petsko and D. Ringe. 1998. Evolution of an enzyme active site: the structure of a new crystal form of muconate lactonizing enzyme compared with mandelate racemase and enolase. **Proc. Natl. Acad. Sci.** 95(18):10396-10401.

Helin, S., P.C. Kahn, B.L. Guha, D.G. Mallows and A. Goldman. 1995. The refined X-ray structure of muconate lactonizing enzyme from Pseudomonas putida PRS2000 at 1.85 A resolution. **J. Mol. Biol.** 254(5):918-941.

Hoffmann, M., M. Wanko, P. Strodel, P.H. König, T. Frauenheim, K. Schulten, W. Thiel, E. Tajkhorshid and M. Elstner. 2006. Color tuning in rhodopsins: the mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II. **J. Am. Chem. Soc.** 128(33):10808-10818.

Kajander, T., L. Lehtiö, M. Schlömann and A. Goldman. 2003. The structure of Pseudomonas P51 Cl-muconate lactonizing enzyme: co-evolution of structure and dynamics with the dehalogenation function. **Protein. Sci.** 12(9):1855-1864.

Kamp, M.W.v.d., F. Perruccio and A.J. Mulholland. 2008. High-level QM/MM modelling predicts an arginine as the acid in the condensation reaction catalysed by citrate synthase. **Chem. Commun.** (16):1874-1876.

Klenchin, V.A., E.A. Taylor Ringia, J.A. Gerlt and I. Rayment. 2003. Evolution of enzymatic activity in the enolase superfamily: structural and mutagenic studies of the mechanism of the reaction catalyzed by o-succinylbenzoate synthase from Escherichia coli. **Biochemistry Easton** 42(49):14427-14433.

Lai, W., H. Chen, K.B. Cho and S. Shaik. 2009. Effects of substrate, protein environment, and proximal ligand mutation on compound I and compound 0 of chloroperoxidase (dagger). **J. Phys. Chem. A** 113(43)**:** 11764-11771

Larsen, T.M., J.E. Wedekind, I. Rayment and G.H. Reed. 1996. A carboxylate oxygen of the substrate bridges the magnesium ions at the active site of enolase: structure of the yeast enzyme complexed with the equilibrium mixture of 2-phosphoglycerate and phosphoenolpyruvate at 1.8 Å resolution. **Biochemistry** 35(14):4349-4358.

Li, X., L.W. Chung, H. Mizuno, A. Miyawaki and K. Morokuma. 2009. A theoretical study on the nature of on- and off-states of reversibly photoswitching fluorescent protein dronpa: absorption, emission, protonation, and raman. **J. Phys. Chem. B** 114(2):1114-1126.

Lin, H. and D.G. Truhlar. 2007. QM/MM: What have we learned, where are we, and where do we go from here? **Theor. Chem. Acc.** 117(2):185-199.

Liu, H., Y. Zhang and W. Yang. 2000. How is the active site of enolase organized to catalyze two different reaction steps? **J. Am. Chem. Soc.** 122(28):6560-6570.

Lundberg, M., T. Kawatsu, T. Vreven, J.M. Frisch and K. Morokuma. 2009. Transition states in a protein environment − ONIOM QM:MM modeling of isopenicillin N synthesis. **J. Chem. Theory. Comput.** 5(1):222-234.

Lundberg, M. and K. Morokuma. 2007. Protein environment facilitates O2 binding in non-heme iron enzyme. An insight from ONIOM calculations on isopenicillin N synthase (IPNS). **J. Phys. Chem. B** 111(31):9380-9389.

Lundberg, M., P.E. Siegbahn and K. Morokuma. 2008. The mechanism for isopenicillin N synthase from density-functional modeling highlights the similarities with other enzymes in the 2-His-1-carboxylate family. **Biochemistry Easton** 47(3):1031-1042.

Maestro, version 9.1. 2010. **SchrÖdinger, LCC, New York, NY**

Malumbres, M. and M. Barbacid. 2005. Mammalian cyclin-dependent kinases. **Trends. Biochem. Sci.** 30(11):630-641.

Morgan, D.O. 1995. Principles of CDK regulation. **Nature** 374(6518):131-134.

Morgan, D.O. 1997. Cyclin-dependent kinases: engines, clocks, and microprocessors. **Annu. Rev. Cell. Dev. Biol.** 13261-291.

Neidhart, D.J., P.L. Howell, G.A. Petsko, V.M. Powers, R.S. Li, G.L. Kenyon and J.A. Gerlt. 1991. Mechanism of the reaction catalyzed by mandelate racemase. 2. Crystal structure of mandelate racemase at 2.5 Å resolution: identification of the active site and possible catalytic residues. **Biochemistry Easton** 30(38):9264-9273.

Neidhart, D.J., G.L. Kenyon, J.A. Gerlt and G.A. Petsko. 1990. Mandelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologous. **Nature** 347(6294):692-694.

Norbury, C. and P. Nurse. 1992. Animal cell cycles and their control. **Annu. Rev. Biochem.** 61(1):441-468.

Petsko, G.A., G.L. Kenyon, J.A. Gerlt, D. Ringe and J.W. Kozarich. 1993. On the origin of enzymatic species. **Trends. Biochem. Sci.** 18(10):372-376.

Plyasova, L.M., N.A. Vasilieva, S.V. Cherepanova, A.N. Shmakov and A.L. Chuvilin. 1998. Structure investigation of defect MgO - high temperature process catalyst. **Nucl. Instrum. Meth. A** 405(2-3):473-475.

Sakai, A., A.A. Fedorov, E.V. Fedorov, A.M. Schnoes, M.E. Glasner, S. Brown, M.E. Rutter, K. Bain, S. Chang, T. Gheyi, J.M. Sauder, S.K. Burley, P.C. Babbitt, S.C. Almo and J.A. Gerlt. 2009a. Evolution of enzymatic activities in the enolase superfamily: stereochemically distinct mechanisms in two families of cis,cis-muconate lactonizing enzymes. **Biochemistry** 48(11):2569-2570.

Sakai, A., A.A. Fedorov, E.V. Fedorov, A.M. Schnoes, M.E. Glasner, S. Brown, M.E. Rutter, K. Bain, S. Chang, T. Gheyi, J.M. Sauder, S.K. Burley, P.C. Babbitt, S.C. Almo and J.A. Gerlt. 2009b. Evolution of enzymatic activities in the enolase superfamily: stereochemically distinct mechanisms in two families of cis,cis-muconate lactonizing enzymes. **Biochemistry** 48(7):1445-1453.

Sakai, A., D.F. Xiang, C. Xu, L. Song, W.S. Yew, F.M. Raushel and J.A. Gerlt. 2006. Evolution of enzymatic activities in the enolase superfamily: N-succinylamino acid racemase and a sew pathway for the irreversible conversion of d- to l-amino acids. **Biochemistry** 45(14):4455-4462.

Schmidt, D.M., E.C. Mundorff, M. Dojka, E. Bermudez, J.E. Ness, S. Govindarajan, P.C. Babbitt, J. Minshull and J.A. Gerlt. 2003. Evolutionary potential of (beta/alpha)8-barrels: functional promiscuity produced by single substitutions in the enolase superfamily. **Biochemistry Easton** 42(28):8387-8393.

Schöneboom, J.C., F. Neese and W. Thiel. 2005. Toward identification of the compound I reactive intermediate in cytochrome P450 chemistry: a QM/MM study of its EPR and Mössbauer parameters. **J. Am. Chem. Soc.** 127(16):5840-5853.

Senn, H. M. and W. Thiel. 2009. QM/MM methods for biomolecular systems. **Angew. Chem. Int. Ed.** 48(7):1198-1229.

Sinnecker, S., N. Svensen, E.W. Barr, S. Ye, J.M. Bollinger, F. Neese and C. Krebs. 2007. Spectroscopic and computational evaluation of the structure of the high-spin Fe(IV)-oxo intermediates in taurine: alpha-ketoglutarate dioxygenase from Escherichia coli and its His99Ala ligand variant. **J. Am. Chem. Soc.** 129(19):6168-6179.

Thoden, J.B., E.A. Taylor Ringia, J.B. Garrett, J.A. Gerlt, H.M. Holden and I. Rayment. 2004. Evolution of enzymatic activity in the enolase superfamily: structural studies of the promiscuous o-succinylbenzoate synthase from Amycolatopsis. **Biochemistry Easton** 43(19):5716-5727.

Thompson, T.B., J.B. Garrett, E.A. Taylor, R. Meganathan, J.A. Gerlt and I. Rayment. 2000. Evolution of enzymatic activity in the enolase superfamily: structure of o-succinylbenzoate synthase from Escherichia coli in complex with Mg2+ and o-succinylbenzoate. **Biochemistry** 39(35):10662-10676.

Vick, J.E. and J.A. Gerlt. 2007. Evolutionary potential of (beta/alpha)(8)-barrels: stepwise evolution of a "new" reaction in the enolase superfamily. **Biochemistry Easton** 46(50):14589-14597.

Vick, J.E., D.M.Z. Schmidt and J.A. Gerlt. 2005. Evolutionary potential of (β/α)8-barrels: in vitro enhancement of a “new” reaction in the enolase superfamily. **Biochemistry** 44(35):11722-11729.

Vreven, T., K.S. Byun, I. Komáromi, S. Dapprich, J.A. Montgomery, K. Morokuma and M.J. Frisch. 2006a. Combining quantum mechanics methods with molecular mechanics methods in ONIOM. **J. Chem. Theory. Comput.** 2(3):815-826.

Vreven, T., M.J. Frisch, K.N. Kudin, H.B. Schlegel and K. Morokuma. 2006b. Geometry optimization with QM/MM methods II: explicit quadratic coupling. **Molec. Phys.** 104(5):701-714.

Wieczorek, S.J., K.A. Kalivoda, J.G. Clifton, D. Ringe, G.A. Petsko and J.A. Gerlt. 1999. Evolution of enzymatic activities in the enolase superfamily: identification of a “new” general acid catalyst in the active site of d-galactonate dehydratase from Escherichia coli. **J. Am. Chem. Soc.** 121(18):4540-4541.

Zhao, Y. and D.G. Truhlar. 2006. Density functionals for noncovalent interaction energies of biological importance. **J. Chem. Theory. Comput.** 3(1):289-300.

Zhao, Y. and D.G. Truhlar. 2008. Density functionals with broad applicability in chemistry. **Acc. Chem. Res.** 41(2):157-167.

Zheng, J., D. Wang, W. Thiel and S. Shaik. 2006. QM/MM study of mechanisms for compound I formation in the catalytic cycle of cytochrome P450cam. **J. Am. Chem. Soc.** 128(40):13204-13215.

APPENDICES

Appendix A

Coordinates of the QM regions

***Syn*-MLE: Reactant**

F	-8.341	-8.351	-9.616	C	-9.468	2.869	-0.116
C	-7.141	-8.048	-8.708	O	-9.826	1.810	0.429
H	-6.634	-8.989	-8.469	O	-8.341	3.032	-0.702
H	-6.430	-7.416	-9.224	F	-6.419	7.482	-0.759
C	-7.504	-7.337	-7.435	C	-6.889	6.079	-1.147
N	-8.215	-7.938	-6.430	H	-7.287	5.576	-0.258
C	-8.192	-7.080	-5.431	H	-7.718	6.101	-1.858
H	-8.665	-7.227	-4.473	C	-5.734	5.218	-1.660
N	-7.499	-5.963	-5.723	O	-4.546	5.484	-1.407
H	-7.353	-5.190	-5.047	O	-6.068	4.164	-2.317
C	-7.043	-6.112	-7.002	F	-1.987	3.435	-0.805
H	-6.409	-5.375	-7.473	C	-3.372	2.854	-0.559
F	-10.209	-7.236	-1.214	H	-3.311	1.948	0.024
C	-9.353	-6.276	-2.047	H	-3.995	3.590	-0.057
H	-8.580	-6.870	-2.560	N	-4.044	2.530	-1.840
O	-8.760	-5.331	-1.199	H	-4.765	1.802	-1.711
H	-8.135	-4.845	-1.770	H	-4.610	3.345	-2.183
C	-10.199	-5.543	-3.100	Mg	-7.696	2.969	-2.587
H	-10.942	-4.917	-2.606	O	-6.870	2.119	-4.377
H	-9.535	-4.892	-3.673	C	-6.210	1.285	-3.708
H	-10.719	-6.194	-3.814	O	-6.370	1.198	-2.444
F	-10.887	-1.731	-5.415	C	-5.261	0.397	-4.408
C	-9.551	-1.748	-4.678	H	-8.717	-0.273	-5.859
H	-9.723	-1.884	-3.615	H	-4.976	0.709	-5.409
H	-8.909	-2.557	-5.028	C	-4.932	-0.832	-3.973
N	-8.777	-0.506	-4.875	O	-6.784	-2.560	-4.800
H	-9.136	0.312	-4.313	H	-4.376	-1.468	-4.651
H	-3.448	2.197	-2.596	C	-5.241	-1.467	-2.685
H	-7.807	-0.743	-4.609	H	-4.803	-1.031	-1.790
F	-13.360	1.547	-2.580	C	-5.985	-2.578	-2.563
C	-11.967	0.925	-2.506	H	-6.085	-3.048	-1.591
H	-11.748	0.702	-1.454	C	-6.720	-3.192	-3.709
H	-11.944	-0.026	-3.039	O	-7.274	-4.314	-3.497
C	-10.856	1.887	-2.930	O	-8.942	4.287	-3.396
O	-11.136	3.098	-2.987	H	-9.864	3.924	-3.353
O	-9.684	1.389	-3.086	H	-8.741	4.457	-4.323
F	-11.808	3.731	0.214				
C	-10.368	4.114	-0.101				
H	-10.316	4.607	-1.071				
H	-9.993	4.827	0.643				

***Syn*-MLE: *Syn*-TS**

F	3.424	-10.736	10.498
C	3.449	-9.505	9.586
H	4.150	-8.772	10.000
H	3.806	-9.777	8.602
C	2.108	-8.865	9.390
N	1.458	-8.155	10.369
C	0.412	-7.627	9.771
H	-0.331	-7.007	10.248
N	0.352	-7.941	8.460
H	-0.364	-7.622	7.802
C	1.438	-8.729	8.197
H	1.662	-9.090	7.204
F	-3.029	-5.059	11.706
C	-2.628	-5.621	10.338
H	-1.540	-5.483	10.223
O	-3.319	-4.913	9.337
H	-3.016	-5.280	8.494
C	-2.961	-7.113	10.221
H	-4.038	-7.263	10.299
H	-2.654	-7.474	9.236
H	-2.471	-7.751	10.965
F	-4.300	-10.548	7.625
C	-3.937	-9.324	6.791
H	-4.500	-8.458	7.189
H	-2.881	-9.076	6.946
N	-4.116	-9.592	5.389
H	-5.100	-9.611	5.129
H	-3.870	-5.947	0.016
H	-3.717	-8.823	4.873
F	-9.257	-10.411	7.071
C	-8.138	-9.492	6.584
H	-8.440	-8.454	6.773
H	-7.224	-9.666	7.148
C	-7.923	-9.566	5.064
O	-8.749	-10.209	4.382
O	-6.953	-8.885	4.589
F	-11.194	-8.052	4.690
C	-10.406	-7.733	3.425
H	-10.110	-8.657	2.930
H	-11.033	-7.176	2.718
C	-9.155	-6.900	3.755
O	-9.063	-6.277	4.828
O	-8.266	-6.884	2.832
F	-9.694	-7.452	-1.748
C	-8.941	-7.573	-0.416
H	-9.382	-6.880	0.310
H	-9.051	-8.561	0.036
C	-7.466	-7.185	-0.541
O	-7.034	-6.527	-1.506
O	-6.706	-7.512	0.446
F	-4.701	-4.123	-1.633
C	-5.320	-4.405	-0.269
H	-5.082	-3.617	0.430
H	-6.397	-4.515	-0.371
N	-4.815	-5.685	0.291
H	-4.828	-5.679	1.322
H	-5.472	-6.486	0.073
Mg	-6.814	-8.220	2.425
O	-4.763	-8.899	2.320
C	-4.205	-7.829	2.630
O	-4.843	-6.757	2.854
C	-2.676	-7.782	2.644
H	-2.282	-8.779	2.855
H	-2.368	-7.557	1.615
C	-1.993	-6.766	3.553
O	-1.696	-7.322	4.854
H	-1.019	-6.531	3.116
C	-2.647	-5.463	3.864
H	-3.056	-4.796	3.118
C	-2.554	-5.229	5.180
H	-2.852	-4.352	5.732
C	-1.964	-6.401	5.803
O	-1.718	-6.600	6.988
O	-7.866	-9.862	2.040
H	-8.258	-10.153	2.908
H	-7.328	-10.590	1.710

***Syn*-MLE: *Anti*-TS**

F	3.424	-10.736	10.498
C	3.441	-9.495	9.585
H	4.123	-8.755	10.017
H	3.823	-9.757	8.609
C	2.095	-8.863	9.355
N	1.424	-8.148	10.315
C	0.396	-7.613	9.688
H	-0.347	-6.985	10.151
N	0.362	-7.923	8.373
H	-0.345	-7.611	7.684
C	1.453	-8.720	8.143
H	1.702	-9.076	7.154
F	-3.029	-5.059	11.707
C	-2.628	-5.607	10.328
H	-1.542	-5.462	10.214
O	-3.327	-4.895	9.336
H	-3.022	-5.240	8.483
C	-2.955	-7.099	10.190
H	-4.031	-7.253	10.271
H	-2.644	-7.435	9.192
H	-2.459	-7.747	10.920
F	-4.300	-10.548	7.625
C	-3.899	-9.308	6.810
H	-4.520	-8.467	7.102
H	-2.860	-9.028	6.990
N	-4.058	-9.512	5.352
H	-5.093	-9.484	5.068
H	-3.630	-5.956	0.621
H	-3.553	-8.765	4.848
F	-9.257	-10.411	7.071
C	-8.105	-9.498	6.637
H	-8.440	-8.458	6.743
H	-7.248	-9.624	7.299
C	-7.749	-9.632	5.154
O	-8.597	-10.157	4.409
O	-6.648	-9.084	4.771
F	-11.194	-8.052	4.690
C	-10.419	-7.750	3.425
H	-10.131	-8.676	2.928
H	-11.039	-7.183	2.720
C	-9.159	-6.948	3.761
O	-9.064	-6.329	4.840
O	-8.271	-6.957	2.851
F	-9.694	-7.452	-1.748
C	-8.953	-7.537	-0.428
H	-9.415	-6.854	0.294
H	-9.019	-8.527	0.028
C	-7.504	-7.090	-0.601
O	-7.146	-6.363	-1.546
O	-6.708	-7.478	0.314
F	-4.701	-4.123	-1.633
C	-5.192	-4.431	-0.205
H	-4.963	-3.615	0.461
H	-6.266	-4.592	-0.246
N	-4.637	-5.686	0.411
H	-4.996	-5.716	1.376
H	-5.116	-6.512	0.002
Mg	-6.730	-8.111	2.231
O	-4.705	-8.751	1.836
C	-4.307	-7.631	2.261
O	-5.154	-6.869	2.876
C	-2.979	-7.140	2.002
H	-3.609	-10.349	5.035
H	-2.363	-7.822	1.419
C	-2.303	-6.434	3.020
O	-1.936	-7.402	4.497
H	-1.251	-6.267	2.796
C	-2.840	-5.288	3.786
H	-3.323	-4.478	3.258
C	-2.617	-5.310	5.109
H	-2.852	-4.517	5.806
C	-2.059	-6.604	5.516
O	-1.801	-6.907	6.715
O	-7.735	-9.808	2.045
H	-8.091	-10.110	2.916
H	-7.210	-10.522	1.669

***Syn*-MLE: *Syn*-Product**

F	3.424	-10.736	10.498
C	3.449	-9.505	9.586
H	4.150	-8.772	10.000
H	3.806	-9.777	8.602
C	2.108	-8.865	9.390
N	1.458	-8.155	10.369
C	0.412	-7.627	9.771
H	-0.331	-7.007	10.248
N	0.352	-7.941	8.460
H	-0.364	-7.622	7.802
C	1.438	-8.729	8.197
H	1.662	-9.090	7.204
F	-3.029	-5.059	11.706
C	-2.628	-5.621	10.338
H	-1.540	-5.483	10.223
O	-3.319	-4.913	9.337
H	-3.016	-5.280	8.494
C	-2.961	-7.113	10.221
H	-4.038	-7.263	10.299
H	-2.654	-7.474	9.236
H	-2.471	-7.751	10.965
F	-4.300	-10.548	7.625
C	-3.937	-9.324	6.791
H	-4.500	-8.458	7.189
H	-2.881	-9.076	6.946
N	-4.116	-9.592	5.389
H	-5.100	-9.611	5.129
H	-3.870	-5.947	0.016
H	-3.717	-8.823	4.873
F	-9.257	-10.411	7.071
C	-8.138	-9.492	6.584
H	-8.440	-8.454	6.773
H	-7.224	-9.666	7.148
C	-7.923	-9.566	5.064
O	-8.749	-10.209	4.382
O	-6.953	-8.885	4.589
F	-11.194	-8.052	4.690
C	-10.406	-7.733	3.425
H	-10.110	-8.657	2.930
H	-11.033	-7.176	2.718
C	-9.155	-6.900	3.755
O	-9.063	-6.277	4.828
O	-8.266	-6.884	2.832
F	-9.694	-7.452	-1.748
C	-8.941	-7.573	-0.416
H	-9.382	-6.880	0.310
H	-9.051	-8.561	0.036
C	-7.466	-7.185	-0.541
O	-7.034	-6.527	-1.506
O	-6.706	-7.512	0.446
F	-4.701	-4.123	-1.633
C	-5.320	-4.405	-0.269
H	-5.082	-3.617	0.430
H	-6.397	-4.515	-0.371
N	-4.815	-5.685	0.291
H	-4.828	-5.679	1.322
H	-5.472	-6.486	0.073
Mg	-6.814	-8.220	2.425
O	-4.763	-8.899	2.320
C	-4.205	-7.829	2.630
O	-4.843	-6.757	2.854
C	-2.676	-7.782	2.644
H	-2.282	-8.779	2.855
H	-2.368	-7.557	1.615
C	-1.993	-6.766	3.553
O	-1.696	-7.322	4.854
H	-1.019	-6.531	3.116
C	-2.647	-5.463	3.864
H	-3.056	-4.796	3.118
C	-2.554	-5.229	5.180
H	-2.852	-4.352	5.732
C	-1.964	-6.401	5.803
O	-1.718	-6.600	6.988
O	-7.866	-9.862	2.040
H	-8.258	-10.153	2.908
H	-7.328	-10.590	1.710

***Syn*-MLE: *Anti*-Product**

F	3.424	-10.736	10.498
C	3.452	-9.504	9.589
H	4.148	-8.770	10.009
H	3.817	-9.773	8.608
C	2.114	-8.864	9.380
N	1.452	-8.148	10.347
C	0.433	-7.594	9.722
H	-0.303	-6.955	10.183
N	0.400	-7.903	8.408
H	-0.268	-7.556	7.711
C	1.476	-8.711	8.171
H	1.718	-9.067	7.181
F	-3.029	-5.058	11.705
C	-2.630	-5.611	10.332
H	-1.546	-5.460	10.211
O	-3.342	-4.909	9.341
H	-3.030	-5.239	8.487
C	-2.943	-7.107	10.207
H	-4.016	-7.272	10.296
H	-2.631	-7.449	9.214
H	-2.433	-7.744	10.939
F	-4.300	-10.548	7.625
C	-3.943	-9.336	6.780
H	-4.523	-8.480	7.114
H	-2.886	-9.078	6.875
N	-4.254	-9.577	5.375
H	-5.287	-9.676	5.186
H	-4.067	-8.688	4.833
H	-3.707	-10.337	4.986
F	-9.257	-10.411	7.071
C	-8.135	-9.486	6.637
H	-8.465	-8.450	6.788
H	-7.253	-9.634	7.260
C	-7.842	-9.580	5.150
O	-8.696	-10.089	4.396
O	-6.754	-9.031	4.774
F	-11.194	-8.052	4.690
C	-10.400	-7.762	3.419
H	-10.166	-8.693	2.906
H	-10.996	-7.151	2.730
C	-9.087	-7.034	3.748
O	-8.983	-6.355	4.783
O	-8.162	-7.168	2.871
F	-9.694	-7.452	-1.748
C	-8.936	-7.575	-0.427
H	-9.297	-6.807	0.267
H	-9.140	-8.532	0.059
C	-7.417	-7.369	-0.609
O	-6.958	-6.773	-1.593
O	-6.705	-7.834	0.343
F	-4.701	-4.123	-1.633
C	-5.339	-4.413	-0.278
H	-5.130	-3.593	0.405
H	-6.431	-4.469	-0.407
N	-4.771	-5.623	0.305
H	-5.187	-5.709	1.226
H	-5.146	-6.428	-0.189
Mg	-6.703	-8.354	2.217
O	-4.676	-8.249	2.185
C	-3.904	-7.591	2.916
O	-4.134	-7.324	4.133
C	-2.609	-7.093	2.326
H	-2.016	-7.940	1.970
H	-2.863	-6.495	1.436
C	-1.819	-6.233	3.288
O	-1.354	-6.963	4.456
H	-0.934	-5.862	2.766
C	-2.549	-5.079	3.866
H	-3.118	-4.400	3.243
C	-2.391	-5.061	5.194
H	-2.775	-4.347	5.907
C	-1.727	-6.297	5.561
O	-1.542	-6.755	6.693
O	-7.082	-10.272	2.466
H	-7.835	-10.333	3.109
H	-7.305	-10.776	1.688

***Anti*-MLE: Reactant**

F	-2.249	-16.593	2.455
C	-2.745	-15.588	1.417
H	-2.442	-15.946	0.414
H	-3.845	-15.587	1.434
O	-2.206	-14.327	1.686
H	-2.730	-13.647	1.250
F	-4.831	-10.066	7.499
C	-4.173	-8.988	6.645
H	-4.270	-8.020	7.138
H	-3.114	-9.212	6.507
N	-4.805	-8.853	5.323
H	-5.869	-8.802	5.200
H	-3.370	-6.110	0.038
H	-4.509	-9.507	4.595
F	-9.696	-10.004	6.626
C	-8.619	-9.032	6.181
H	-9.029	-8.021	6.237
H	-7.771	-9.055	6.865
C	-8.236	-9.231	4.724
O	-8.966	-9.898	3.980
O	-7.177	-8.602	4.353
F	-11.665	-7.856	3.954
C	-10.785	-7.734	2.705
H	-10.403	-8.711	2.419
H	-11.374	-7.364	1.864
C	-9.604	-6.807	2.923
O	-9.712	-5.818	3.658
O	-8.539	-7.130	2.272
F	-9.170	-7.300	-2.379
C	-8.637	-7.403	-0.940
H	-9.203	-6.715	-0.301
H	-8.824	-8.387	-0.528
C	-7.185	-7.056	-0.751
O	-6.672	-6.142	-1.431
O	-6.546	-7.655	0.164
F	-3.980	-3.827	-1.251
C	-4.807	-4.469	-0.160
H	-4.765	-3.881	0.746
H	-5.830	-4.585	-0.492
N	-4.362	-5.853	0.089
H	-4.683	-6.116	1.031
H	-4.920	-6.474	-0.512
F	0.300	-4.043	-0.823
C	0.491	-5.171	-1.820
H	1.070	-4.787	-2.663
H	1.035	-5.993	-1.365
C	-0.859	-5.544	-2.329
O	-1.556	-4.678	-2.859
N	-1.251	-6.835	-2.145
H	-2.248	-6.993	-2.120
H	-0.789	-7.434	-1.457
Mg	-6.885	-8.268	2.041
O	-5.324	-6.651	2.562
C	-4.479	-7.590	2.724
O	-4.812	-8.789	2.593
C	-3.152	-7.209	3.266
H	-3.090	-6.174	3.575
H	-4.464	-7.976	4.930
C	-2.115	-8.037	3.463
O	-2.386	-7.732	0.432
H	-1.307	-7.654	4.080
C	-1.893	-9.382	2.921
H	-1.763	-10.182	3.651
C	-1.649	-9.658	1.630
H	-1.372	-10.679	1.372
C	-1.563	-8.664	0.492
O	-0.599	-8.882	-0.310
O	-7.233	-10.301	2.147
H	-7.385	-10.852	1.375
H	-8.035	-10.327	2.715

***Anti*-MLE: *Syn*-TS**

F	-2.249	-16.593	2.455
C	-2.747	-15.590	1.416
H	-2.447	-15.950	0.414
H	-3.846	-15.589	1.436
O	-2.209	-14.325	1.676
H	-2.769	-13.650	1.284
F	-4.831	-10.066	7.499
C	-4.166	-8.988	6.636
H	-4.247	-8.019	7.131
H	-3.109	-9.224	6.497
N	-4.795	-8.828	5.308
H	-5.853	-8.678	5.208
H	-4.358	-8.021	4.832
H	-4.633	-9.549	4.607
F	-9.696	-10.004	6.626
C	-8.620	-9.029	6.193
H	-9.032	-8.018	6.251
H	-7.775	-9.054	6.881
C	-8.230	-9.221	4.740
O	-8.955	-9.901	3.997
O	-7.185	-8.582	4.371
F	-11.665	-7.856	3.954
C	-10.789	-7.741	2.701
H	-10.408	-8.719	2.418
H	-11.381	-7.376	1.859
C	-9.604	-6.816	2.903
O	-9.699	-5.830	3.644
O	-8.554	-7.138	2.228
F	-9.170	-7.300	-2.379
C	-8.623	-7.399	-0.947
H	-9.178	-6.703	-0.308
H	-8.812	-8.381	-0.530
C	-7.164	-7.048	-0.794
O	-6.694	-6.090	-1.434
O	-6.483	-7.708	0.053
F	-3.980	-3.827	-1.251
C	-4.792	-4.464	-0.120
H	-4.795	-3.813	0.744
H	-5.805	-4.639	-0.456
N	-4.285	-5.812	0.267
H	-3.299	-5.987	0.103
H	-4.854	-6.514	-0.220
F	0.300	-4.043	-0.823
C	0.452	-5.183	-1.812
H	1.042	-4.825	-2.659
H	0.963	-6.021	-1.348
C	-0.917	-5.500	-2.309
O	-1.540	-4.639	-2.925
N	-1.429	-6.722	-1.970
H	-2.432	-6.835	-2.006
H	-0.950	-7.325	-1.307
Mg	-6.919	-8.278	1.931
O	-5.427	-6.602	2.378
C	-4.560	-7.568	2.608
O	-4.963	-8.770	2.598
C	-3.272	-7.105	2.813
H	-4.591	-6.135	1.493
H	-3.140	-6.076	3.104
C	-2.084	-7.925	2.756
O	-1.259	-7.655	1.407
H	-1.349	-7.617	3.498
C	-2.155	-9.411	2.694
H	-2.602	-9.999	3.486
C	-1.573	-9.908	1.597
H	-1.489	-10.945	1.296
C	-1.049	-8.791	0.797
O	-0.470	-8.857	-0.296
O	-7.303	-10.296	2.104
H	-7.519	-10.859	1.356
H	-8.064	-10.298	2.733

***Anti*-MLE: *Anti*-TS**

F	-2.249	-16.593	2.455
C	-2.743	-15.586	1.418
H	-2.448	-15.948	0.415
H	-3.843	-15.577	1.441
O	-2.195	-14.327	1.681
H	-2.711	-13.649	1.235
F	-4.831	-10.066	7.499
C	-4.170	-8.988	6.644
H	-4.259	-8.017	7.133
H	-3.111	-9.218	6.509
N	-4.791	-8.853	5.325
H	-5.798	-8.569	5.241
H	-3.360	-6.029	0.101
H	-4.736	-9.645	4.696
F	-9.696	-10.004	6.626
C	-8.614	-9.030	6.196
H	-9.026	-8.019	6.238
H	-7.779	-9.044	6.898
C	-8.193	-9.244	4.754
O	-8.832	-10.048	4.057
O	-7.225	-8.515	4.346
F	-11.665	-7.856	3.954
C	-10.768	-7.719	2.720
H	-10.318	-8.681	2.477
H	-11.364	-7.427	1.851
C	-9.663	-6.690	2.923
O	-9.816	-5.737	3.694
O	-8.602	-6.893	2.214
F	-9.170	-7.300	-2.379
C	-8.618	-7.399	-0.947
H	-9.196	-6.731	-0.299
H	-8.767	-8.392	-0.537
C	-7.177	-6.993	-0.788
O	-6.708	-6.046	-1.449
O	-6.479	-7.584	0.094
F	-3.980	-3.827	-1.251
C	-4.812	-4.448	-0.145
H	-4.781	-3.832	0.742
H	-5.829	-4.592	-0.485
N	-4.355	-5.822	0.180
H	-4.646	-6.024	1.158
H	-4.941	-6.503	-0.336
F	0.300	-4.043	-0.823
C	0.442	-5.186	-1.810
H	1.027	-4.833	-2.662
H	0.953	-6.025	-1.345
C	-0.931	-5.501	-2.298
O	-1.556	-4.639	-2.913
N	-1.444	-6.713	-1.943
H	-2.434	-6.871	-2.057
H	-0.964	-7.352	-1.309
Mg	-6.990	-8.081	1.988
O	-5.433	-6.682	2.559
C	-4.595	-7.649	2.709
O	-4.967	-8.838	2.499
C	-3.280	-7.337	3.234
H	-3.102	-6.290	3.451
H	-4.167	-8.095	4.736
C	-2.145	-8.119	2.983
O	-1.362	-7.677	1.375
H	-1.282	-7.796	3.553
C	-2.135	-9.576	2.711
H	-2.526	-10.268	3.449
C	-1.570	-9.953	1.561
H	-1.458	-10.971	1.205
C	-1.119	-8.781	0.773
O	-0.553	-8.863	-0.341
O	-7.826	-9.878	1.761
H	-7.370	-10.645	1.404
H	-8.226	-10.114	2.633

***Anti*-MLE: *Syn*-Product**

F	-2.249	-16.593	2.455
C	-2.740	-15.586	1.418
H	-2.463	-15.960	0.414
H	-3.840	-15.557	1.455
O	-2.166	-14.334	1.665
H	-2.637	-13.663	1.164
F	-4.831	-10.066	7.499
C	-4.300	-9.093	6.467
H	-4.405	-8.070	6.827
H	-3.248	-9.305	6.267
N	-5.095	-9.245	5.214
H	-5.930	-8.604	5.080
H	-4.602	-9.173	4.299
H	-5.586	-10.131	5.105
F	-9.696	-10.004	6.626
C	-8.628	-9.025	6.172
H	-9.026	-8.013	6.274
H	-7.758	-9.070	6.827
C	-8.288	-9.178	4.695
O	-8.956	-9.924	3.968
O	-7.307	-8.446	4.298
F	-11.665	-7.856	3.954
C	-10.791	-7.721	2.706
H	-10.370	-8.688	2.438
H	-11.393	-7.394	1.856
C	-9.656	-6.735	2.912
O	-9.792	-5.760	3.659
O	-8.592	-7.006	2.238
F	-9.170	-7.300	-2.379
C	-8.654	-7.419	-0.937
H	-9.235	-6.746	-0.297
H	-8.832	-8.409	-0.537
C	-7.218	-7.033	-0.751
O	-6.751	-6.140	-1.460
O	-6.576	-7.602	0.199
F	-3.980	-3.827	-1.251
C	-4.738	-4.631	-0.220
H	-4.760	-4.103	0.734
H	-5.762	-4.741	-0.553
N	-4.173	-5.997	-0.192
H	-3.560	-6.134	0.597
H	-4.942	-6.649	-0.069
F	0.300	-4.043	-0.823
C	0.412	-5.186	-1.818
H	1.034	-4.854	-2.652
H	0.875	-6.051	-1.346
C	-0.978	-5.428	-2.331
O	-1.494	-4.561	-3.031
N	-1.633	-6.549	-1.884
H	-2.624	-6.431	-1.632
H	-1.140	-7.091	-1.189
Mg	-6.938	-8.112	2.051
O	-5.311	-6.759	2.743
C	-4.415	-7.623	2.781
O	-4.627	-8.854	2.588
C	-3.032	-7.264	3.284
H	-2.781	-6.226	3.094
H	-3.096	-7.353	4.375
C	-1.886	-8.162	2.864
O	-1.313	-7.759	1.618
H	-1.130	-8.020	3.635
C	-2.089	-9.632	2.710
H	-2.481	-10.265	3.494
C	-1.594	-10.026	1.534
H	-1.546	-11.019	1.108
C	-1.098	-8.838	0.840
O	-0.532	-8.747	-0.236
O	-7.097	-10.202	2.220
H	-7.123	-10.730	1.414
H	-7.965	-10.287	2.670

***Anti*-MLE: *Anti*-Product**

F	-2.249	-16.593	2.455
C	-2.746	-15.590	1.416
H	-2.480	-15.969	0.411
H	-3.844	-15.558	1.462
O	-2.165	-14.336	1.641
H	-2.758	-13.648	1.326
F	-4.831	-10.066	7.499
C	-4.175	-9.000	6.624
H	-4.166	-8.043	7.156
H	-3.128	-9.277	6.441
N	-4.860	-8.798	5.375
H	-5.822	-8.494	5.438
H	-3.397	-6.031	0.076
H	-4.886	-9.594	4.761
F	-9.696	-10.004	6.626
C	-8.620	-9.036	6.158
H	-9.011	-8.020	6.237
H	-7.750	-9.080	6.812
C	-8.261	-9.242	4.687
O	-8.865	-10.117	4.036
O	-7.381	-8.456	4.215
F	-11.665	-7.856	3.954
C	-10.766	-7.715	2.721
H	-10.290	-8.670	2.499
H	-11.366	-7.457	1.844
C	-9.693	-6.642	2.900
O	-9.865	-5.690	3.666
O	-8.636	-6.809	2.170
F	-9.170	-7.300	-2.379
C	-8.621	-7.401	-0.945
H	-9.206	-6.742	-0.292
H	-8.766	-8.395	-0.536
C	-7.186	-6.990	-0.767
O	-6.693	-6.065	-1.449
O	-6.508	-7.547	0.152
F	-3.980	-3.827	-1.251
C	-4.823	-4.436	-0.146
H	-4.790	-3.819	0.740
H	-5.839	-4.576	-0.493
N	-4.386	-5.815	0.183
H	-4.662	-6.009	1.166
H	-5.000	-6.487	-0.322
F	0.300	-4.043	-0.823
C	0.435	-5.187	-1.809
H	1.024	-4.838	-2.660
H	0.940	-6.030	-1.344
C	-0.942	-5.486	-2.298
O	-1.547	-4.628	-2.933
N	-1.489	-6.678	-1.909
H	-2.480	-6.821	-2.033
H	-1.031	-7.298	-1.249
Mg	-7.046	-8.057	2.060
O	-5.380	-6.684	2.608
C	-4.553	-7.649	2.715
O	-4.841	-8.816	2.405
C	-3.203	-7.317	3.300
H	-2.954	-6.264	3.182
H	-3.375	-7.495	4.366
C	-2.030	-8.166	2.897
O	-1.482	-7.790	1.612
H	-1.253	-7.962	3.636
C	-2.182	-9.640	2.799
H	-2.584	-10.233	3.608
C	-1.666	-10.067	1.638
H	-1.587	-11.075	1.255
C	-1.230	-8.893	0.892
O	-0.689	-8.820	-0.208
O	-7.827	-9.865	1.765
H	-7.375	-10.631	1.401
H	-8.245	-10.119	2.632

Appendix B

Poster and oral presentations

Poster Presentation

1. Tuanjai Somboon, Matthew Paul Gleeson, and Supa Hannongbua. **Insight into the Reaction Mechanisms of *cis,cis*-Muconate Lactonizing Enzymes from Theoretical Calculations.** The Regional Meeting of Physical Chemistry, Hitachi, Ibaraki Prefecture, Japan, 5th November, 2010.

Insight into the Reaction Mechanism of *cis,cis*-Muconate Lactonizing Enzymes from Theoretical Calculations

Tuanjai Somboon[ϕ], M. Paul Gleeson[ϕ*], and Supa Hannongbua[ϕ,δ*]

[ϕ] Department of Chemistry, Faculty of Science, Kasetsart University, 50 Phaholyothin Road, Chatuchak, Bangkok 10900, Thailand.

[δ] Center of Nanotechnology KU, Kasetsart University, 50 Phaholyothin Road, Chatuchak, Bangkok 10900, Thailand.

Corresponding author : E-mail: fscisph@ku.ac.th; Tel. +66-2-5625555 ext. 2140 Fax.+66-2-562-5555

Introduction:

Muconate lactonizing enzymes (MLEs) are members of the enolase superfamily which catalyze the conversion of *cis,cis*-muconates to muconolactones. The members of this superfamily are much studied in order to illustrate the complexity and redundancy in enzyme evolution. It was recently reported by Sakai et al that two different MLEs from *Mycobacterium smegmatis* and *Pseudomonas fluorescens*, derived from divergent families and displaying around 30% sequence identity, both catalyze the same chemical reaction, but involve stereochemically distinct mechanisms (*anti*- and *syn*-cycloisomerization). This is particularly interesting from a fundamental evolutionary perspective in that nature has evolved two distinct proteins, which bare striking similarity at the active site level, but achieve the same product in distinctly different ways.

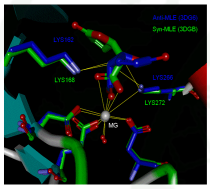


Figure 1. Comparison of the *Anti*- (blue) and *Syn* (green) MLE active site regions.

The nature of reactions of enzymes in this diverse superfamily is abstraction the α-protons of carboxylic acids that bind an essential Mg^{2+} ion. The stereochemically distinct mechanism of MLEs is shown in Figure 2. The experimental observation established that cycloisomerization catalyzed by *M. smegmatis* goes along *anti*-stereochemical course, whereas *syn*-stereochemical course catalyzed by *P. fluorescens*. The fundamental understanding of this much experimentally studied biological process is important for biological understanding.

In this case there are no theoretical studies have been performed and also nature of the base is unknown in the reaction.

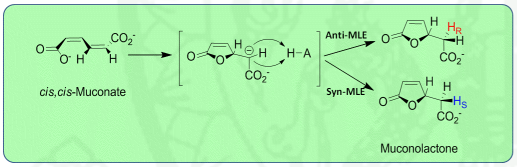


Figure 2. MLEs mechanism proposed by Sakai et al for Syn and Anti-MLEs. Two proximal Lys residues exist in both active sites, which will lead to either syn- or anti products. While they catalyse the formation of the natural precourser, remarkably, each protein catalyzes the formation of different chiral product in experiments performed in deuterated solvent. It is also unclear to whether the intermediate proposed by Sakai et al (square brackets above) is a true intermediate on the reaction potential energy surface.

Computational details:

Protein preparation: The MLEs crystal structures (3DGB and 3DG6) were downloaded from RCSB protein databank and prepared as follows. Cofactors, ions, and water molecules were removed from the protein-ligand complexes. Hydrogen atoms were added into the system and ionizable amino acids side chains were protonated assuming a pH of 7.4. The system underwent restrained minimization using the IMPREF utility to optimize hydrogen atoms and remove any high energy contacts or distorted bonds, angles, and dihedrals. These coordinates were then used for subsequent QM/MM calculations.

QM/MM calculations: All QM/MM calculations were obtained by using the ONIOM methodology developed by Morokuma and co-workers as implemented in Gaussian 03. The inner QM region was treated by 6-31G* basis set for optimization and 6-311+G** for single point calculations. The low region is treated by MM using universal Force field (UFF). The point charges of outer MM region were directly embedded in SCF to account for polarisation. The energy of the entire QM/MM system is obtained as following,

$$E_{QM/MM} = E_{QM\text{-}EE\ (model)} + E_{MM(model)} - E_{MM(model)} \qquad \text{Equation 1}$$

Results:

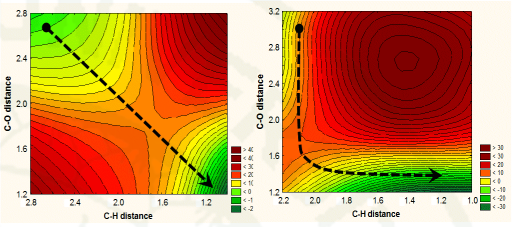


Figure 3. Potential energy surfaces for the Anti-anti (right) and Syn-syn catalysed reactions. In the former, P162 acts as the catalytic acid and in the former, P168. Energies are at the M05/6-31G*:UFF level of theory.

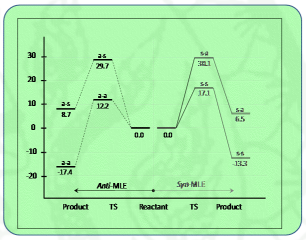


Figure 4. QM/MM predicted mechanism energetics for the syn and anti products in both *Syn*- and *Anti*-MLE. Relative M05/6-31G*:UFF energies reported in kcal/mol.

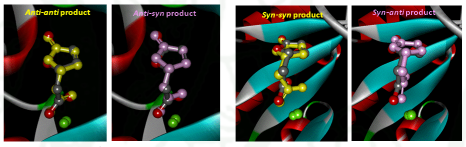


Figure 5. Structures *anti*- and *syn*- products. X-ray structures are shown in atom color. QM/MM structures are shown in pink and yellow color

Conclusions:

In this study we were interested in explaining the origin of the very interesting stereochemical differences found in two members of the enolase superfamily which catalyze the same chemical process. Our QM/MM results confirm that the Lys adjacent to the ligands must act as the base in the catalytic reaction. The calculations also show that the base leading to the formation of the lowest energy product for both enzyme leads to the expected experimental stereochemistry. Finally, these results have further confirmation of the value of theoretical simulations to help understand results derived from complex biochemical experiments.

Acknowledgments:

We would like to thank Department of Chemistry, Kasetsart University, the Thailand Research Fund, The National Center of Excellence for Petroleum, Petrochemicals and Advance Materials (NCE-PPAM) and National Nanotechnology Center for supporting this work.

References:

1. Sakai, A., Fedorov, A. A., Fedorov, E. V., Schnoes, A. M., Glasner, M. E., Brown, S., Rutter, M. E., Bain, K., Chang, S., Gheyi, T., Sauder, J. M., Burley, S. K., Babbitt, P. C., Almo, S. C., and Gerlt, J. A., *Biochemistry*, **2009**, 48, 1445-1453.
2. Babbitt, P. C.; Hasson, M. S.; Wedekind, J. E.; Palmer, D. R.; Barrett, W. C.; Reed, G. H.; Rayment, I.; Ringe, D.; Kenyon, G. L.; Gerlt, J. A. *Biochemistry* **1996**, 35, 16489–16501.
3. Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct.* ***1999***, 1-21, 461–462.
4. Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* ***2006***, 2, 815–826.

Oral Presentation

1. Tuanjai Somboon, Matthew Paul Gleeson, and Supa Hannongbua. **Hybrid Quantum Mechanical/Molecular Mechanical studies on Two Families of *cis,cis*-Muconate Lactonizing Enzymes.** The 14th International Annual Symposium on Computational Science and Engineering, Mae Fah Luang University, Chiang Rai, Thailand, 23rd – 26th March, 2010.

2. Tuanjai Somboon, Matthew Paul Gleeson, and Supa Hannongbua. **Insight into the Reaction Mechanisms of *cis,cis*-Muconate Lactonizing Enzymes from Theoretical Calculations**. The Great Tsukuba Seminar for Physical Chemistry Chiba, Ibaraki Prefecture, Japan, 12th -13th November, 2010.

Appendix C

Publications

PUBLICATIONS

1. Tuanjai Somboon, Matthew Paul Gleeson, and Supa Hannongbua. **Hybrid Quantum Mechanical/Molecular Mechanical studies on Two Families of *cis,cis*-Muconate Lactonizing Enzymes**. The abstract of 14th International Annual Symposium on Computational Science and Engineering, Mae Fah Luang University, Chiang Rai, Thailand, 23rd – 26th March, 2010.

B00003 119

Hybrid Quantum Mechanical/Molecular Mechanical studies on Two Families of *cis,cis*-Muconate Lactonizing Enzymes

T. Somboon[1], **M. P. Gleeson**[1], **and S. Hannongbua** [1,C]
[1] *Faculty of Science, Kasetsart University, 50, Phahon Yothin Rd, Chatuchak, Bangkok, 10900, Thailand*
[C] **E-mail**: fscisph@ku.ac.th; **Fax**: 02-562-5555; **Tel.** 02-5625555

ABSTRACT

Muconate lactonizing enzymes (MLEs) are members of the enolase family which catalyse the conversion of *cis,cis*-muconates to muconolactones. It was recently reported by Sakai et al that two different MLEs, derived from divergent families and displaying only 26% sequence identity, both catalyze the same chemical reaction, but involve stereochemically distinct mechanisms (*anti*- and *syn*-cycloisomerization). This is particularly interesting from a fundamental evolutionary perspective, in that nature has evolved two distinct proteins, which bare striking similarity at the active site level, but achieve the same product in distinctly different ways.

This example represents an ideal case study for a computational analysis, in an effort to understand the reasons for the distinct reactivity differences observed. Computational Chemistry can play a significant role in the elucidation of physical processes by allowing us to simulate many diverse processes at the atomic level, including biochemical reactions. The preferred QM methods cannot however be employed on protein sized systems so more approximate methods such as the Hybrid Quantum Mechanical-Molecular Mechanical (QM/MM) methods must be used. In this method the key portion of the system is treated QM and the remaining environment modeled using the less computationally demanding MM method. Such simulations have been used to model the mechanisms of action of numerous proteins as a method to model the reaction mechanism of enzymes.

The aim of this research is to gain insight into the mechanism of MLEs using QM/MM methods. Particular emphasis will be placed on: (a) understanding the origin of differences in stereochemical courses for *anti* and *syn*-MLE, (b) identify the basic residue involved in the reaction and (c) to estimate the energy profiles along the reaction coordinate between two possibilities of *anti* and *syn*-MLE. From this research we add insight into this important area of biochemistry, and help to shed the light on the mechanism for these two enzymes and the reason for the stereochemical differences. Particularly in this case where evolution has resulted in diverse, mechanistically distinct ways of producing the same substrate.

Keywords: Muconate lactonizing enzymes, enolase family, distinct mechanisms, Hybrid QM/MM methods.

REFERENCES

1. Sakai, A., Fedorov, A. A., Fedorov, E. V., Schnoes, A. M., Glasner, M. E., Brown, S., Rutter, M. E., Bain, K., Chang, S., Gheyi, T., Sauder, J. M., Burley, S. K., Babbitt, P. C., Almo, S. C., and Gerlt, J. A., *Biochemistry*, 2009, **48**, 1445-1453.

2. Tuanjai Somboon, Matthew Paul Gleeson, and Supa Hannongbua**. Insight into the Reaction Mechanisms of *cis,cis*-Muconate Lactonizing Enzymes from Theoretical Calculations**. The abstract of Regional Meeting of Physical Chemistry, Hitachi, Ibaraki Prefecture, Japan, 5th November, 2010.

30A Insight into the Reaction Mechanism of *cis,cis*-Muconate Lactonizing Enzymes from Theoretical Calculations

(理論計算によるムコン酸ラクトン化酵素の反応機構解明)

(Kasetsart University, Bangkok, 10900, Thailand)
○Tuanjai Somboon,# Matthew Paul Gleeson, and Supa Hannongbua

The enolase superfamily has received considerable attention from a biochemical perspective recently as it has helped to illustrate the complexity and redundancy in enzyme evolution. *Cis,cis*-Muconate Lactonizing Enzymes (MLEs) represent an interesting subclass of this family. MLEs derived from *Mycobacterium smegmatis (Anti-MLE)* and *Pseudomonas fluorescens* (*Syn*-MLE) share ~76% identity and have a very similar arrangement of catalytic residues in their active sites, however, while they catalyze the conversion of *cis,cis*-muconate to the same achiral product, muconolactone, studies in deuterated solvent surprisingly show that the cyclo-isomerization proceeds with the formation of a chiral product.

The hybrid quantum mechanics/molecular mechanics (QM/MM) calculations is applied to probe aspects of MLE function and gain further insight into the catalytic events in both *Anti*- and *Syn*-MLEs that lead to the observed stereochemical differences.

In this work we discuss the application of DFT QM/MM calculations on both MLEs, to our knowledge the first reported in the literature on this protein. We investigate the proposal that the base involved in the catalytic reaction is the lysine residue found at the end of the 2nd strand given: (a) that the lysine residue at the end of the 6th strand is in an apparently equally effective position to catalyze reaction and (b) that the structural related epimerases in-fact achieve their stereo-specific outcomes by relying on either the base from the 2nd or 6th strand.

Keywords: Muconate lactonizing enzymes, enolase family, Hybrid QM/MM methods.

Visiting student to Prof. Mori's group: College of Science, Ibaraki University

3. Tuanjai Somboon, Matthew Paul Gleeson, and Supa Hannongbua. **Insight into the Reaction Mechanisms of *cis,cis*-Muconate Lactonizing Enzymes:** A DFT QM/MM Study. Journal of Molecular Modeling. Accepted: 7 April, 2011.

J Mol Model
DOI 10.1007/s00894-011-1088-2

ORIGINAL PAPER

Insight into the reaction mechanism of *cis,cis*-muconate lactonizing enzymes: a DFT QM/MM study

Tuanjai Somboon · Matthew Paul Gleeson · Supa Hannongbua

Received: 30 January 2011 / Accepted: 7 April 2011

Abstract MLEs derived from *mycobacterium smegmatis* and *seudomonas fluorescens* share ~76% identity and have a very similar arrangement of catalytic residues in their active site configuration. However, while they catalyze the conversion of *cis,cis*-muconate to the same achiral product, muconolactone, studies in deuterated solvent surprisingly show that the cyclo-isomerization proceeds with the formation of a chiral product. In this paper we discuss the application of DFT QM/MM calculations on both MLEs, to our knowledge the first reported in the literature on this protein. We investigate the proposal that the base involved in the catalytic reaction is the lysine residue found at the end of the 2nd strand given: (a) that the lysine residue at the end of the 6th strand is in an apparently equally effective position to catalyze reaction and (b) that the structural related epimerase in-fact achieve their stereo-specific outcomes by relying on either the base from the 2nd or 6th strand.

Keywords Enolase family · Hybrid QM/MM methods · Muconate lactonizing enzymes

Introduction

The enolase superfamily has received a considerable amount of attention from a biochemical perspective recently as it has helped to illustrate the complexity and redundancy in enzyme evolution [1–3]. Muconate lactonizing enzymes (MLEs) are interesting members of this superfamily which catalyze the conversion of *cis,cis*-muconates to muconolactones. Sakai et al. [4–6] have recently reported data on two different MLEs derived from *mycobacterium smegmatis* and *seudomonas fluorescens*, that share ~76% identity. These proteins catalyze the same chemical reaction (Scheme 1), but involve a stereochemically distinct reaction mechanism even though the product is achiral. Studies in deuterated solvent have established that cyclo-isomerization catalyzed by *mycobacterium smegmatis* (*anti*-MLE) proceeds along an *anti*-stereochemical course, whereas *seudomonas fluorescens* (*syn*-MLE) catalyzes the *syn*-stereochemical course.

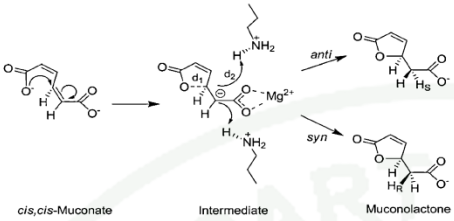
MLEs are Mg^{2+} containing metallo-proteins consisted of ~370 amino acids arranged into a TIM barrel-like α/β protein fold. The Mg^{2+} is found deep within a non-solvent exposed cavity, coordinated by 2 aspartate and 1 glutamic acid residues, a single water molecule and the substrate, in a distorted octahedral form. The carboxylate of *cis,cis*-muconate binds across the Mg^{2+} ion in conformations that display mirror-like symmetry in the two different proteins (Fig. 1). The 2nd carboxylate group of *cis,cis*-muconate interacts with residues toward the rear of the pocket, glutamine and threonine in *anti*-MLE and histidine and threonine residues in *syn*-MLE. Two Lys residues, located on the 2nd and 6th

Electronic supplementary material The online version of this article (doi:10.1007/s00894-011-1088-2) contains supplementary material, which is available to authorized users.

T. Somboon · M. P. Gleeson (✉) · S. Hannongbua
Department of Chemistry, Faculty of Science,
Kasetsart University,
50 Phaholyothin Road, Chatuchak,
Bangkok 10900, Thailand
e-mail: paul.gleeson@ku.ac.th

S. Hannongbua (✉)
Center of Nanotechnology KU, Kasetsart University,
50 Phaholyothin Road, Chatuchak,
Bangkok 10900, Thailand
e-mail: fscisph@ku.ac.th

Published online: 04 May 2011



Scheme 1 Reaction mechanism proposed by Sakai et al. for *anti*- and *syn*-MLEs. Two proximal Lys residues exist in both active sites, which could potentially lead to the formation of either the *syn*- or *anti*-products. In fact, while both proteins catalyze the formation of the same natural precursor, they lead to the formation of different chiral products based on experiments performed in deuterated solvent. It is also unclear whether the enolate intermediate proposed by Sakai et al. is a stable intermediate during the reaction

strands, are sufficiently closed to the substrate alpha carbon to act as the general base in the reaction.

Structural information has proved crucial to understand the sequence of events that lead to the chiral products in deuterated solvent. Sakai et al. have determined X-ray crystal structures of the product state (muconate lactone), for both *syn*- and *anti*-proteins, and have used these to identify the most probable base in the catalytic reaction. They propose that the identity of the base in both proteins is the Lys residue at the end of the 2nd strand that explains the stereochemical aspects of the reaction. This result is interesting given members of the structurally related epimerase sub-class achieve their stereo-specific outcomes by relying on either the base from the 2nd or 6th strand. Furthermore, analysis of the interaction distances between the nitrogen atoms of the two possible bases and the product α-carbon in both MLE PDBs (Table 1) simplistically suggests that Lys on the 6th strand is more likely to be the base in *syn*-MLE (3.03 Å vs 3.57 Å). This however would not explain the stereo-chemical differences in the reaction. Additionally, it is not known if a stable enolate anion does in fact exist as a meta-stable intermediate in the reaction.

Computational chemistry can play an important role in understanding protein function since it enables us to simulate events at an atomic level. In this study we apply hybrid quantum mechanics/molecular mechanics (QM/MM) calculations to probe aspects of MLE function and gain further insight into the catalytic events in both *anti*- and *syn*-MLEs that lead to the observed stereochemical differences. In the QM/MM technique the active site residues that undergo chemical change, or directly influence the sequence of events in the catalytic reaction, are treated using more accurate QM methods, while the remainder of the protein is treated using less rigorous, but more computationally efficient MM methods. The QM/MM calculations performed here rely on the ONIOM methodology [7, 8] using the electrical embedding scheme. Here the total QM/MM energy of the system is computed in a subtractive fashion as given in the following equation; QM energy of the active site region, or "model", plus the MM energy of the "real" or whole protein system, minus the MM energy of the model region. For more extensive reviews of the QM/MM technique see the following references [9–12].

$$E_{QM/MM} = E_{QM-EE(model)} + E_{MM(real)} - E_{MM(model)} \quad (1)$$

QM/MM methods have been applied to study the enolase family member phosphoenolpyruvate. This protein contains two Mg^{2+} ions in the active site and the mechanism involves the two separate CH proton abstraction steps. Liu et al. [13] reported QM/MM free energy perturbation barrier heights of ~13.1 kcal mol^{-1} for the initial proton abstraction step, resulting in a stable enolate intermediate of ~5 kcal mol^{-1}. Decomposition of the intermediate by abstraction of a further proton from the β-carbon was found to require ~9 kcal mol^{-1}. The stability of enolate intermediates has also been investigated by Van Der Kamp et al. [14] in the unrelated protein citrate synthase. They find that proton abstraction from the α-carbon of oxaloacetate requires ~10.2 kcal mol^{-1} and results in a stable intermediate ~8 kcal mol^{-1} higher in energy than

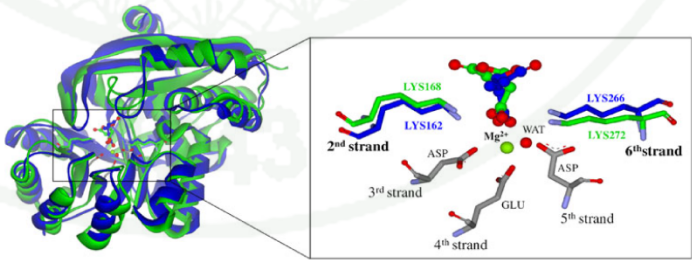


Fig. 1 A comparison of the *anti*- (blue) and *syn*-MLE (green) proteins (left) and active site regions

Table 1 Predicted and experimental distances observed for the conversion of *cis,cis*-muconate to muconolactone during the MLE catalyzed reactions. Distances are reported in Å

Distances	Structures					
Anti-MLE	Reactant	*Syn*-TS	*Syn*-Product	*Anti*-TS	*Anti*-Product	X-ray (product)
Mg−O(ASP191)	2.35	2.47	2.30	2.41	2.22	2.08
Mg−O(GLU217)	2.02	2.02	2.00	2.01	2.02	1.96
Mg−O(ASP242)	2.00	2.01	1.95	2.02	2.05	2.19
Mg−O(Sub)	2.31	2.29	2.23	2.17	2.23	2.16
Mg−O(Sub)	2.21	2.12	2.49	2.22	2.36	2.29
Mg−O(WAT)	2.06	2.06	2.10	1.99	1.99	2.20
C(Sub)−N(LYS162)	3.11	3.39	3.45	2.99	3.04	3.03
C(Sub)−N(LYS266)	3.66	3.03	3.87	3.56	3.66	3.57
C(Sub)−HN(LYS162)	2.25	2.47	2.67	1.90	1.10	-
C(Sub)−HN(LYS266)	2.92	2.10	1.09	2.81	2.90	-
O(Sub)−H(GLN294)	1.86	1.90	2.01	1.84	1.88	-
Syn-MLE	Reactant	*Syn*-TS	*Syn*-Product	*Anti*-TS	*Anti*-Product	X-ray (Product)
Mg−O(ASP197)	2.59	2.62	2.27	2.72	2.65	2.34
Mg−O(GLU223)	1.99	2.02	2.01	2.02	2.00	2.33
Mg−O(ASP248)	2.04	2.03	2.10	2.02	1.95	2.24
Mg−O(Sub)	2.15	2.07	2.16	2.16	2.03	2.42
Mg−O(Sub)	2.22	2.22	2.49	2.11	3.37	2.96
Mg−O(WAT)	1.99	2.00	1.99	1.98	1.97	1.79
C(Sub)−N(LYS168)	3.66	3.07	3.60	4.24	4.26	4.22
C(Sub)−N(LYS272)	3.55	4.02	3.81	2.72	3.31	3.45
C(Sub)−HN(LYS168)	2.80	2.00	1.09	3.33	3.31	-
C(Sub)−HN(LYS272)	3.13	3.77	3.42	1.93	1.10	-
O(Sub)−H(HIE21)	1.78	1.86	1.88	1.89	1.82	-

the reactants. While the mechanisms differ considerably from MLE, the results discussed above serve as useful benchmarks for studies on MLEs.

In this paper we discuss the application of QM/MM calculations to MLEs, to our knowledge the first reported in the literature on this system. The goal of this study is to help conclusively prove the identity of the base in the *anti*- and *syn*-MLE proteins given the contrasting way related enolase family member achieve stereo-specific catalysis and the apparent ambiguities from an analysis of the X-ray interactions distances between possible bases. We are also interested in assessing whether an explicit enolate anion exists as a reaction intermediate as found in other enolases, or as a transition state.

Computational details

The crystal structures of *anti*-MLE (pdb accession code: 3DG6) and *syn*-MLE (pdb accession code: 3DGB) were downloaded from RCSB protein data bank and prepared as follows. Cofactors, ions, and water molecules beyond 15 Å of the active site were deleted. Missing side chain data from both PDB structures as well as hydrogen atoms were added using Discovery Studio 2.5 according to the CHARMM 22 forcefield. The protonation states of ionizable residues were determined by visual analysis. Ligand charges were determined using the AM1BCC method and parameters according to the Accelrys CHARMm forcefield. Both proteins were solvated in a box of TIP3P water with a minimum distance of 7 Å between the protein and box edge (i.e, 14 Å between proteins in a periodic box). Counterions were added to neutralize the system. Default non-bonded cut-offs of 12 Å were used in all MM simulations.

Due to the difficulty in accurately simulating metallo-proteins using MM methods, the Mg^{2+} ion, its three chelating carboxylate groups and one water molecule were harmonically restrained to their X-ray positions during all of the MM preparation steps. MM optimization was achieved in three distinct steps. All optimizations were performed in Discovery Studio 2.5 [15] using the smart optimizer conditions and an RMS gradient below 0.1 kcal mol^{-1}. These were; (1) optimization of hydrogen atoms only followed by (2) optimization of all amino acid side chain

atoms and solvent molecules and finally (3) optimization of all atomic coordinates. This setup is equivalent to the default protein preparation procedure to prepare protein X-ray structures for docking, molecular dynamics or QM/MM in the modeling package Maestro [16].

The MM optimized coordinates were then used in a subsequent short molecular dynamics (MD) step to help minimize any high energy contacts that are often present in X-ray protein structures. Atoms beyond 10 Å of the active site were harmonically restrained. MD was performed in two stages; (a) heating from T=0 to 300K over 200 ps. (b) equilibration for 800 ps. Simulations were performed using the default CHARMM settings in Discovery Studio 2.5. These stages include a time step of 0.001 ps., NVT conditions, 12 Å non-bonded cut-offs and particle mesh Ewald [17]. The flexible atoms from the final MD step were subsequently re-optimized and used as input for QM/MM calculations.

All QM/MM calculations were performed using the ONIOM methodology developed by Morokuma and co-workers as implemented in Gaussian 03 [18]. A QM region has been selected so that key polar residues that directly interact with the substrate over the course of the reaction are included explicitly. For *anti*-MLE 70 atoms are treated QM consisting of the side chains of; SER23, THR54, LYS162, ASP191, GLU217, ASP242, LYS266, GLN294, the Mg^{2+} ion and the substrate. For *syn*-MLE 76 atoms are treated QM consisting of the side chains of; HIE21, THR140, LYS168, ASP197, GLU223, ASP248, LYS272, the Mg^{2+} ion and the substrate. The side chains of the following active site residues were treated flexibly with the rest of the system being fixed (*anti*-MLE: PHE21, PHE53, LYS160, ASN193, ILE295; *syn*-MLE: ILE53, THR58, LYS166, ASN199, THR300, LEU302, GLU326, PHE328). All water molecules were removed for computational efficiency except the water molecule that chelates Mg^{2+}.

The M05 functional developed by Zhao et al. [19, 20] has been used in these calculations as it has been shown to be more effective than the popular B3LYP method for describing aspects of non-bonded interactions. The 6-31G(d) basis set was employed for geometry optimizations. Single point energies on stationary points being characterized using the 6-311+G(d,p) basis set. The MM region was treated using universal Force field (UFF) in conjunction with CHARMM partial charges. The reaction coordinates of both possible bases, in both proteins, have been estimated by the stepwise variation of the C-O and C-H bonds between their reactant and product configurations. Due to the large memory requirements of Gaussian 03 with MM regions of this size (>5000 atoms) transitions states were characterized as saddle points by doing a frequency calculations of the optimized QM region coordinates only. ONIOM optimization has been performed using default settings; fixed link atom positions and involves the electrical embedding of MM charges into the QM calculation.

Results and discussion

In our attempt to confirm the identity of the catalytic base in the *anti*- and *syn*-MLE proteins, we have determined the minimum energy structures for the reactants and both the *anti*- and *syn*-products in both proteins. As the products may be dictated by kinetic factors, we have also determined the potential energy surface between reactant and products to estimate the reaction barriers and also to assess the possibility of an explicit enolate anion existing during the course of the reaction.

X-ray versus QM/MM structures

Firstly we analyze the differences between the two different product structures for each protein, which differ only in terms of the position of a single proton, and compare these to the corresponding heavy atom coordinates reported in the respective X-ray structure PDB files. The expectation is that the most thermodynamically favorable QM/MM coordinates will be more similar to the original X-ray structure and that the corresponding base will give rise to the expected stereochemical outcome.

Analysis of the RMSDs of the QM/MM optimized active site regions of the two possible products could potentially be used to identify which base is involved in the catalytic reaction. One would expect that the optimized product that displays the lowest RMSD to the experimental X-ray structure would identify which base is involved in the reaction. Listed in Table 1 are the key geometrical parameters associated with the stationary points. While the optimized QM/MM structures of both products for both *anti*- and *syn*-MLE proteins, display C_α RMSDs of<0.05 Å to the corresponding X-ray structure, the flexible QM and MM heavy atoms in the QM/MM system understandably display larger differences in structure. Note, supplementary information Fig. S1 displays the overlay of (a) the original X-ray coordinates, (b) the MD output structure and (c) the QM/MM optimized geometries, highlighting the rather small differences between them. For *syn*-MLE we find that the product structure involving the base which leads to the *syn*-product (Lys-168 located on the 2nd strand) has a lower RMSD than that formed with Lys-272, which results in the *anti*-product (RMSDs of 0.36 vs 0.56 Å). Furthermore, we also find that the *anti*-product formed within the *anti*-MLE active site has a lower overall RMSD that that of the *syn*-

product (0.58 vs 0.63 Å) suggesting that the 2nd strand lysine, Lys-162, is the base in this reaction and not Lys-266 located on the 6th strand.

Looking in more detail at the product structures (Table 1) we observe that the key interacts made in the *anti-anti* product are clearly in better agreement with the X-ray coordinates than the *anti-syn* product. However, for *syn*-MLE both products show distances considerably more variable from the X-ray structure. In fact, the configuration of the Mg^{2+} ion and its ligands in the *syn*-MLE QM/MM models is more similar to that observed for the *anti*-MLE products. It is possible that this artifact may have arisen from the fitting of the atomic solution of *syn*-MLE to its electron density. In fact, the Mg-O distances observed in the *syn*-MLE X-ray structure appear considerably longer than the experimentally expected values of ~2.1 Å [21]. Contrasting the two experimental *anti*- and *syn*-MLE X-ray structures (Table 1 and Fig. 1) we can see that Mg^{2+} coordination in the *syn*-MLE structure is much weaker than in the *anti*-MLE. Interestingly, the QM/MM optimized *syn*-product of *syn*-MLE display interactions that are more comparable to the experimental *anti*-product.

Assessment of the relative QM/MM energies shows that the *anti*-MLE *anti*-product (*anti-anti* product) is lower than the corresponding *anti*-MLE *syn*-product (*anti-syn* product) by 26.1 kcal mol^{-1}. Thus, from a purely thermodynamic perspective we would expect Lys-162 to act as the catalytic base since it leads to the lowest energy product and will lead to the known stereochemical outcome. In addition, the energetic differences between the *syn-syn* product and *syn-anti* product shows that the former is favored by 19.8 kcal mol^{-1}. Thus, from a purely thermodynamic perspective, we would expect Lys-168 to act as the catalytic base since it leads to the lowest energy product and the known stereochemical product.

To try and decipher the contribution electrostatics and van der Waals interactions, as well as protein pre-organization, we subsequently performed gas phase single point QM calculations on each of the QM/MM optimized product geometries. Analysis of the contributing terms to the QM/MM energy reveals that the Van der Waals term is essentially constant for structures obtained in the two different protein models. The gas phase QM single point energies show that both the *syn-syn* and *anti-anti* products are still preferred in the gas phase suggesting active site pre-organization is important (ΔE between *syn-syn* and *syn-anti* products are −26 vs −33 kcal mol^{-1} at QM/MM and QM levels respectively; ΔE between *anti-anti* and *anti-syn* products are −20 vs −6 kcal mol^{-1} at QM/MM and QM levels respectively). Interestingly, the inclusion of the protein electrostatic term has a much greater effect in stabilizing the *anti-anti* product over the *anti-syn* product since the difference in QM energy is just 6 kcal mol^{-1} in the gas phase calculation but 20 kcal mol^{-1} in the protein. In contrast, the difference in energy between the *syn-syn* and *syn-anti* product is somewhat larger in the gas phase than in the protein calculations. This might suggests that the two proteins achieve their product selectivities in subtly different ways, *anti*-MLE relying more on organizing its active site to favor the *anti*-product conformation and *syn*-MLE by preferentially stabilizing the *syn*-product as a result of its particular electrostatic characteristics.

QM/MM reaction energetics

The QM/MM calculations performed here on the two possible products formed by both *anti*- and *syn*-MLEs appear to confirm the identity of the catalytic bases in the reaction from both a structural and energetic perspective. However, it is also desirable to assess the reaction barriers associated with the two possible bases in each protein to rule out any differences that might affect the reaction products from a kinetic perspective. We have therefore mapped out the potential energy surface between reactants and both possible products in *anti*- and *syn*-MLE proteins. From this we have estimated the reaction barriers for each process and these results are summarized in Fig. 2. The full potential energy surfaces for the *syn-syn* and *anti-anti* processes are summarized in Fig. 3.

The overall QM/MM results summarized in Fig. 2 show that the barriers to reaction associated with the *anti-anti* product is 11.6 kcal mol^{-1} lower than the corresponding *anti-syn* product. Similarly, we observe a 13.0 kcal mol^{-1} difference between the *syn-syn* and *syn-anti* products of *syn*-MLE. These results clearly suggest that the basic residues involved in the reaction are located at the end of the 2nd strand of both *anti*-MLE and *syn*-MLE, corresponding to Lys-162 and Lys-168 respectively. From

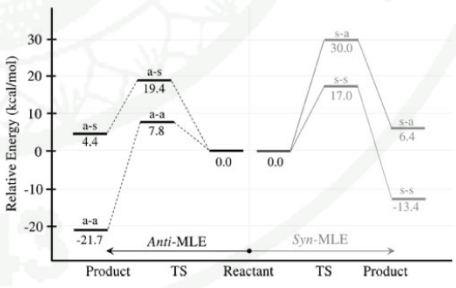


Fig. 2 Relative M05/6-31G(d):UFF optimized energies for the *anti*- and *syn*-products in both *anti*- and *syn*-MLEs. a-a refers to *anti*-product and a-s refers to *syn*-product of *anti*-MLE. For *syn*-MLE, s-s refers to *syn*-product and s-a refers to *anti*-product

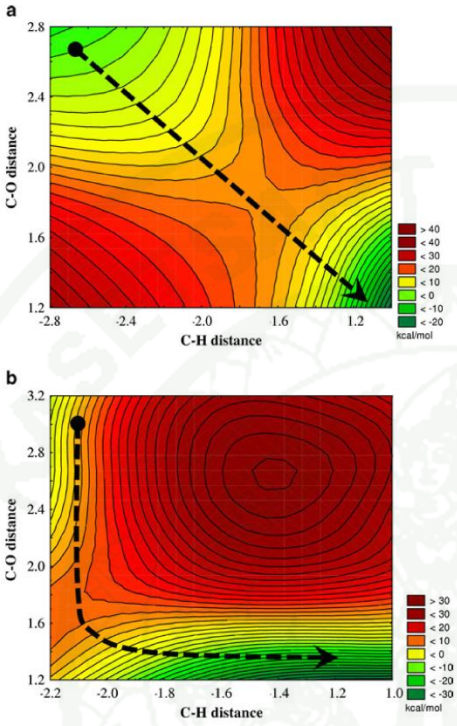


Fig. 3 Potential energy surfaces for the *anti-anti* (left) and *syn-syn* catalyzed reactions. In the former, P162 acts as the catalytic acid and in the latter, P168. C-O and C-H distances correspond to d_1 and d_2 in scheme 1, respectively. Energies are at the M05/6-31G(d):UFF level of theory. The energy performed at M05/6-31G(d) is correlated well with M05/6-311+G(d,p)

an analysis of Fig. 3 it is also apparent that the reaction proceeds along a concerted pathway with both C-O bond breaking and proton transfer to Lys-162. In contrast the process in *syn*-MLE appears to proceed along a more stepwise route, with C-O bond breaking occurring before proton transfer to Lys-168.

The theoretical QM/MM DFT models employed have helped to confirm the identity of the base in the reaction using. We were also therefore interested in understanding whether an explicit enolate anion exists in the active site of either protein. To determine whether the enolate was a stable stationary point within the protein we took the relevant structure obtained from the QM/MM potential energy scans (i.e, the structure having a C-O bond formed but a long C-H bond). This constrained structure was subsequently fully optimized QM/MM using the conditions as used for the other stationary points obtained in this study. In the case of both *syn*- and *anti*-MLE this enolate-like structure decomposed to the muconolactone by accepting a proton from the active site lysine residue which leads to the forming of the C-O bond. While this result would appear to suggest that an explicit enolate anion does not exist in these proteins further work is needed to prove whether this high energy structure is truly a stable stationary point or not. These calculations would necessitate additional polarization and diffuse functions as well as full frequency analyses given that the structure is likely to occupy a shallow energy minimum.

Conclusions

In this paper we discuss the application of DFT QM/MM calculations on *cis,cis*-muconate lactonizing enzymes, to our knowledge the first reported in the literature. QM/MM methods have been used to determine the reaction energetics associated with the conversion of *cis,cis*-muconate to muconolactone in both *anti*- and *syn*-MLEs. We have investigated the proposal that the base involved in the catalytic reaction is the Lys residue found at the end of the 2nd strand, rather than a Lys residue 6th which is almost equally well positioned.

Our QM/MM results show that the expected muconolactone *anti*-product derived from *anti*-MLE is (a) closer to the X-ray structure in terms of RMSD, (b) lower in energy and (c) has a lower barrier to reaction than the corresponding *syn*-product. Our results also show that the *syn*-product derived from *syn*-MLE also displays a lower RMSD to the original X-ray coordinates, is lower in energy, and has a lower barrier to reaction than the corresponding *anti*-product. Thus, for *anti*-MLE we find Lys-162 to results in the lowest energy reaction and Lys-168 for *syn*-MLE.

The theoretical calculation performed here have helped to confirm the identity of the basic residues involved in the MLE reaction originally proposed by Sakai et al. through indirect means from an analysis of the products produced in experiments performed in deuterated solvents. Our results show that although the basic Lys residues located on the 6th strand in both MLEs are almost equally well positioned to accept a proton, however, it is the base located on the 2nd strand that is thermodynamically and kinetically more favorable, as well as giving the optimized QM/MM product closest to the original X-ray structure.

Acknowledgments This work is supported by the Thailand Research Fund (RTA5380010) and in part supported through the National Research University research fund of Kasetsart University. The National Center of Excellence for Petroleum, Petrochemical, and

Advanced Materials (NCE-PPAM) and Faculty of Science, Kasetsart University is grateful to T.S. for scholarship. The computational resource at Laboratory for Computational & Applied Chemistry (LCAC), Kasetsart University is also acknowledged.

References

1. Gerlt JA, Babbitt PC, Rayment I (2005) Divergent evolution in the enolase superfamily: the interplay of mechanism and specificity. Arch Biochem Biophys 433:59–70. doi:10.1016/j.abb.2004.07.034
2. Thompson TB, Garrett JB, Taylor EA, Meganathan R, Gerlt JA, Rayment I (2000) Evolution of enzymatic activity in the enolase superfamily: structure of o-succinylbenzoate synthase from escherichia coli in complex with Mg2+ and o-succinylbenzoate†,‡. Biochemistry 39:10662–10676. doi:10.1021/bi000855o
3. Vick JE, Schmidt DMZ, Gerlt JA (2005) Evolutionary potential of (β/α)8-Barrels: In vitro enhancement of a "new" reaction in the enolase superfamily†. Biochemistry 44:11722–11729. doi:10.1021/bi050963g
4. Sakai A, Fedorov AA, Fedorov EV, Schnoes AM, Glasner ME, Brown S, Rutter ME, Bain K, Chang S, Gheyi T, Sauder JM, Burley SK, Babbitt PC, Almo SC, Gerlt JA (2009) Evolution of enzymatic activities in the enolase superfamily: stereochemically distinct mechanisms in two families of cis, cis-muconate lactonizing enzymes†‡. Biochemistry 48:1445–1453. doi:10.1021/bi802277h
5. Sakai A, Fedorov AA, Fedorov EV, Schnoes AM, Glasner ME, Brown S, Rutter ME, Bain K, Chang S, Gheyi T, Sauder JM, Burley SK, Babbitt PC, Almo SC, Gerlt JA (2009) Evolution of enzymatic activities in the enolase superfamily: stereochemically distinct mechanisms in two families of cis, cis-muconate lactonizing enzymes. Biochemistry 48:2569–2570. doi:10.1021/bi900265w
6. Sakai A, Xiang DF, Xu C, Song L, Yew WS, Raushel FM, Gerlt JA (2006) Evolution of enzymatic activities in the enolase superfamily: N-succinylamino acid racemase and a new pathway for the irreversible conversion of d- to l-amino acids†. Biochemistry 45:4455–4462. doi:10.1021/bi060230b
7. Dapprich S, Komáromi I, Byun KS, Morokuma K, Frisch MJ (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. doi:10.1016/s0166-1280(98)00475-8
8. Vreven T, Byun KS, Komáromi I, Dapprich S, Montgomery JA, Morokuma K, Frisch MJ (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM. J Chem Theor Comput 2:815–826. doi:10.1021/ct050289g
9. Bruice TC (2006) Computational approaches: reaction trajectories, structures, and atomic motions. Enzyme reactions and proficiency. Chem Rev (Washington, DC US) 106:3119–3139. doi:10.1021/cr050283j
10. Lin H, Truhlar DG (2007) QM/MM: What have we learned, where are we, and where do we go from here? Chem Inform. doi:10.1002/chin.200722224
11. Senn HM, Thiel W (2007) Atomistic approaches in modern biology. In: Topics in current chemistry, vol 268. Springer, Berlin
12. Senn HM, Thiel W (2009) QM/MM methods for biomolecular systems. Angew Chem Int Edn 48:1198–1229. doi:10.1002/anie.200802019
13. Liu H, Zhang Y, Yang W (2000) How is the active site of enolase organized to catalyze two different reaction steps? J Am Chem Soc 122:6560–6570. doi:10.1021/ja9936619
14. Kamp MWvd, Perruccio F, Mulholland AJ (2008) High-level QM/MM modelling predicts an arginine as the acid in the condensation reaction catalysed by citrate synthase. Chem Commun (Cambridge, UK) 16:1874–1876
15. Drug Discovery Studio 2.5 (2010) Accelrys Inc, San Diego, CA, USA
16. Maestro, version 9.1 (2010) Schrödinger, LLC, New Yok, NY
17. Darden TYD, Pedersen L (1993) Particle mesh Ewald: An N•log (N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
18. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, 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 JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople J (2004) Gaussian 03, Revision B.05. Gaussian Inc, Wallingford, CT
19. Zhao Y, Truhlar DG (2006) Density functionals for noncovalent interaction energies of biological importance. J Chem Theor Comput 3:289–300. doi:10.1021/ct6002719
20. Zhao Y, Truhlar DG (2008) Density functionals with broad applicability in chemistry. Acc Chem Res 41:157–167. doi:10.1021/ar700111a
21. Plyasova LM, Vasilieva NA, Cherepanova SV, Shmakov AN, Chuvilin AL (1998) Structure investigation of defect MgO - high temperature process catalyst. Nucl Instrum Meth Phys Res A 405:473–475. doi:10.1016/s0168-9002(97)00161-7

CURRICULUM VITAE

NAME : Ms. Tuanjai Somboon

BIRTH DATE : March 24, 1986

BIRTH PLACE : Surat Thani, Thailand

EDUCATION :

YEAR	INSTITUTE	DEGREE
2009	Walailak University	B.Sc. (Computational Science)
2011	Kasetsart University	M.Sc. (Chemistry)

SCHOLARSHIP : In part grants: 2009-2011, Thailand Research Fund (RTA5380010), the National Research University research fund of Kasetsart University, and the National Center of Excellence for Petroleum, Petrochemical, and Advanced Materials (NCE-PPAM)

EXPERIENCE : Short-term research at College of Science, Ibaraki University, Japan, 2nd November – 21st December, 2010.

Attending The Great Tsukuba Seminar for Physical Chemistry, Chiba, Ibaraki Prefecture, Japan, 12th -13th November, 2010.

Attending the Regional Meeting of Physical Chemistry, Hitachi, Ibaraki Prefecture, Japan, 5th November, 2010.

Attending the 14th International Annual Symposium on Computational Science and Engineering, Mae Fah Luang University, Chiang Rai, Thailand, 23rd – 26th March, 2010.

SKILLS **:** Operating systems: Windows and Unix
Office software: Word, Excel, and PowerPoint
Special software: SPSS, Gaussian 03W, GaussView, HyperChem, Discovery Studio, Amber and CS ChemOffice