A HYBRID COMPUTATIONAL APPROACH TO THE BENZOIN SYNTHESIS IN DIFFERENT MEDIA

by

Kevser Göçmen Topal

B.S., Chemistry, Middle East Technical University, 1995

M.S., Chemistry, Abant İzzet Baysal University, 1999

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Chemistry

Boğaziçi University

2008

To my lovely daughter Ayşegül İzem

ACKNOWLEDGEMENTS

I am gratefully indebted to my advisor Prof. Viktorya Aviyente for her unfailing help and guidance during my studies. As a scientist, she was always available to listen to my ideas. Beyond all, I thank her for walking with me while I explore my limits.

I would like to express my sincere gratitude to Prof. Ayhan S. Demir for initiating and continuously supporting my study, and to the members of his group, especially Ömer Reis, for his valuable scientific discussions and friendship. Without them this work could not have been completed.

Further thanks are due to the examining committee members for their valuable and constructive comments, and especially to Prof. Canan Atılgan for her commentary and supply of motivation on molecular dynamics. Her thorough scientific approach will always be an inspiration.

A number of people also share responsibility for the completion of this thesis work: Onur Temizsoylu, who is not only a good friend but also a Beowulf in the face of software and computing problems and whose endless help and friendship has proved precious; Dr. Birgit Schiott and Leyla Çelik from Aarhus University who gave crucial support to initiate this study; all members of the Bogaziçi University Computational Chemistry (BUCC) group, especially Nihan Ölçüm and Şaron Çatak, who extended to me their valuable friendship; Hülya Metiner who provided invaluable assistance in sorting out the administrative matters and gave her friendship; my husband, Dr. Uğur Topal, without whose unceasing support I may have never completed this work, and to whom this successful outcome belongs as much as it does to me; my daughter İzem who inspires my motivation; my parents Ahmet and Fazile Göçmen who have always backed me with their approval and words of encouragement in endeavors that I have undertaken; and last, but not least my sisters, Ülkü Göçmen and Nursel Erdoğan who have offered me their love and all the other minute joys of being a family.

ABSTRACT

A HYBRID COMPUTATIONAL APPROACH TO THE BENZOIN SYNTHESIS IN DIFFERENT MEDIA

Benzoin derivatives are widely used in organic chemistry, mainly as starting materials in drug synthesis and photoinitiators in polymer chemistry, due to their bifunctionality, the presence of the asymmetric center, and the photolabile benzoyl group. In this study, the synthesis of benzoin has been modeled in two different media in the presence of two different catalysts. In the first part of the study, the reactions of benzil derivatives (donor aldehydes) with benzaldehyde derivatives (acceptor aldehydes) have been modeled with PM3 and B3LYP/6-31G*. The mechanism of the reaction and the effect of ortho substitutions on the aromatic phenyl rings on benzil and on benzaldehyde have been discussed. The rate determining steps have been rationalized based on the effect of the substituents. Among the substituted benzaldehyde molecules o-fluoro benzaldehyde has been found to decrease the activation barrier of the reaction, as a result of the electron donating nature of fluorine.

In the second part of the reaction, the enantioselective benzoin synthesis in benzoylformate decarboxylase (BFD) environment has been investigated with molecular dynamics. The effect of the surrounding residues in the vicinity of the active center has been elucidated. Ten models with various protonation states of the amino acids and their mutated counterparts near the active center have been devised, modeled and analysed. Our studies indicate that H70, S26 and H281 are the catalytically important amino acids besides E47 cited in the literature. The role of these residues in the catalytic function of the enzyme has been rationalized. Furthermore, the experimentally observed enantioselectivity was explained on the basis of the face selectivity of the enamine/carbanion produced during the reaction.

ÖZET

BENZOİN SENTEZİNE FARKLI ORTAMLARDA HİBRİD HESAPLAMALI BİR YAKLAŞIM

Benzoin içerdiği fonksiyonel gruplar sayesinde organik sentez kimyasında yaygın olarak kullanılan ve kolay sentezlenen bir moleküldür. Bu çalışmada iki farklı ortam ve farklı reaksiyon koşullarında benzoin sentezi ele alınmıştır. Laboratuvar ortamında benzil ve benzaldehitten siyanür iyonu ile O-benzoillenmiş benzoin sentezi kuvantum mekanik yöntemlerle ele alınmıştır. Bu çalışmada farklı takılı gruplar taşıyan benzaldehit ve benzil molekülleri ele alınarak kinetik ve termodinamik olarak önemli tepkime adımları ile takılı grupların tepkimeye etkisi incelenmiştir. Deneysel olarak önerilen mekanizma PM3 yöntemi ile modellendikten sonra B3LYP/6-31G* yöntemi ile tepkime enerjileri hesaplanmıştır. Siyanohidrin molekülünün oluşması ve nükleofilik saldırısını içeren reaksiyon adımlarının takılı gruplara göre değişerek reaksiyon hızını etkiledikleri tespit edilmiştir. Takılı grup içeren benzaldehit molekülleri arasında o-flor benzaldehidin, florun elektron çekme özelliğinden dolayı, aktivasyon enerjisinin daha düşük olduğu saptanmıştır.

Çalışmanın ikinci kısmında benzoilformat dekarboksilaz enzimi ortamında enantioseçici benzoin sentezi moleküler dinamik yöntemi ile modellenmiştir. Temel olarak aktif merkezi çevreleyen amino asitlerin katalitik tepkimeye etkisi incelenmiştir. Bunun için farklı protonlanma durumlarındaki amino asitler ve bu aminoasitlerin değiştirilmesi ile elde edilen mutasyonlarını içeren on model hazırlanmıştır. Modelleme çalışmamız H70, S26 ve H281'in literatürde de belirtilen E47'ye ek olarak katalitik etkisi olan amino asitler olduğunu göstermiştir. Bu çalışmada, deneysel çalışmalarda enzim ortamında gözlenen enansiyo seçiciliğin, tepkime esnasında oluşan enamin/karbanyon'nun hangi taraftan saldırdığına bağlı olarak ta değiştiği gösterilmiştir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv
ABSTRACT v
ÖZET vi
LIST OF FIGURES ix
LIST OF TABLES xii
LIST OF SYMBOLS/ABBREVIATIONS xiv
1. BENZOIN CONDENSATION 1
2. OBJECTIVE AND SCOPE 10
3. THEORETICAL BACKGROUND 11
3.1. Quantum Mechanics 11
3.1.1.Basis Sets 14
3.1.2.Semiempirical Methods 16
3.1.3.Density Functional Theory 17
3.1.4.Continuum Solvation Model 21
3.1.5.Chemical Reactivity Indices 24
3.2. Molecular Dynamics 25
3.2.1.Fluctuations of Single Atoms 30
4. COMPUTATIONAL STUDY OF THE SYNTHESIS OF BENZOIN DERIVATIVES FROM BENZIL 31
4.1. Introduction 31
4.2. Computational Methodology 33
4.3. Results and Discussion 34
4.3.1.Elucidation of the Reaction Mechanism with Unsubstituted Molecules. 39
4.3.2 Effect of Substituent on Benzaldehyde 46
4.3.3.Effect of the Substituent on Benzil 48
4.3.4.Chemical Reactivity Indices 49
4.4. Conclusion 51

5. UNDERSTANDING THE MODE OF ACTION OF ThDP IN BENZOYLFORMATE DECARBOXYLASE (BFD) 53
5.1. Introduction 53
5.2. Computational Methodology 55
5.2.1.Simulation Details for MD 55
5.2.2.Possible Reaction Mechanisms 56
5.2.3.Model Systems 59
5.2.4.System construction for MD 65
5.2.5.Parameterization of MD Simulations 69
5.2.6.Quantum Mechanics 79
5.3. Results and Discussions 80
5.3.1.Proton Donors/Acceptors in the First Nucleophilic Attack 93
5.3.2.Proton Donors/Acceptors in the Second Nucleophilic Attack 98
5.3.3. Interaction of ThDP with the Surroundings in Model Systems 105
5.4. CONCLUSIONS 108
6. GENERAL CONCLUSIONS 110
7. FUTURE WORK 113
REFERENCES 115

LIST OF FIGURES

Figure 1.1. 2-hydroxy ketones; a) structure of an acyloin, *R* group can be aliphatic or aromatic, b) benzoin. 1

Figure 1.2. Conventional benzoin condensation. 2

Figure 1.3. General mechanism for the function of a catalyst in the acyloin condensation. 4

Figure 1.4. Synthesis of O-benzoylated benzoin from benzil..... 5

Figure 1.5. Solvent effect on the reactions of O-benzoylated cyanohydrin. 6

Figure 1.6. ThDP catalyzed acyloin condensation, I: ThDP, II: Ylide-ThDP, III: Benzaldehyde, IV: Enamine/Carbanion, V: Acceptor aldehyde (Acetaldehyde), VI: Acyloin (S-2-Hydroxypropiophenone (S-HPP)). 7

Figure 1.7. Main enzymatic function of BFD..... 8

Figure 1.8. Mandalate metabolism in *Pseudomonas mobilis* [47]..... 9

Figure 4.1. Computationally modeled mechanism for the benzoin condensation from benzil..... 35

Figure 4.2. Energy barriers for the formation of cyanohydrin (PM3) 37

Figure 4.3. Energy barriers for the condensation of cyanohydrin with benzaldehyde (PM3). 37

Figure 4.4. Various approaches of the cyanohydrin intermediate 7 to the si or re face of the acceptor benzaldehyde 8. 38

Figure 4.5. Energy barriers for the condensation of O-benzoylated cyanohydrin with benzaldehyde: compound 10 is in *SR* conformation (PM3). 39

Figure 4.6. 3-D view of O-benzoylated cyanohydrin 7 formation and its condensation with benzaldehyde 8. 40

Figure 4.7. C7-C19 bond distances in the transition states 9H, 9F, 9Me, 9pyr. 48

Figure 4.8. Atom numbers for the substituted aldehyde molecules and benzil molecule. 50

Figure 5.1. a) BFD representative unit, b) ThDPs at chain interface and c) 3D structure of ThDP, d) catalytically active amino acid residues around ThDP 54

Figure 5.2. S-HPP (VI) condensation in BFD environment. Pre and postreactive interactions for the first and second nucleophilic half reactions. Reactants within dashed rectangles represent the two basic nucleophilic reactions. 57

Figure 5.3. N4′ H-C2 distance in Ylide ThDP with *Model 1* solvated by 8 Å thick water layer and solvated by water box (PBC). 66

Figure 5.4. Dihedrals ϕ_P vs ϕ_T for *Model 1* vs. *Model 1* (PBC). 67

Figure 5.5. Positional fluctutation of the C_α atoms in the reaction zone for the simulation, in *Model 1* and *Model 1*(PBC). The two sets of data are plotted against each other in the inset; the $y = x$ line is shown in blue to guide the eye. 68

Figure 5.6. RMSD layer vs. PBC for *Model 1*. 69

Figure 5.7. Positional fluctuations of the C□ atoms for *Model 1* and BFD crystal structure (PDB code: 1MCZ). 81

Figure 5.8. RMSD vs. time graphs for *Model 1*. 82

Figure 5.9. RMSD vs. time graphs for *Model 2*. 82

Figure 5.10. RMSD vs. time graphs for *Model 3*. 83

Figure 5.11. RMSD vs. time graphs for *Model 4*. 83

Figure 5.12. RMSD vs. time graphs for *Model 5*. 84

Figure 5.13. RMSD vs. time graphs for *Model 6*. 84

Figure 5.14. RMSD vs. time graphs for *Model 7*. 85

Figure 5.15. RMSD vs. time graphs for *Model 8*. 85

Figure 5.16. RMSD vs. time graphs for *Model 9*. 86

Figure 5.17. RMSD vs. time graphs for *Model 10*. 86

Figure 5.18. Dihedral angles ϕ_P vs. ϕ_T for *Models 1 and 2*..... 87

Figure 5.19. Dihedral angles ϕ_P vs. ϕ_T for *Models 3 and 4*..... 88

Figure 5.20. Dihedral angles ϕ_P vs. ϕ_T for *Models 5 and 6*..... 88

Figure 5.21. Dihedral angles ϕ_P vs. ϕ_T for *Models 7 and 8*..... 89

Figure 5.22. Dihedral angles ϕ_P vs. ϕ_T for *Models 9 and 10*..... 89

Figure 5.23. Binding of ThDP and the octahedral geometry around Mg (II) cation. 91

Figure 5.24. Nucleophilic attack of Ylide-ThDP to benzaldehyde (*Model 3)*..... 95

Figure 5.25. H_{donor} transfer bridge in *Model 3*. Average bond distances within the NAC limits for the first nucleophilic attack on the trajectory of the model..... 96

Figure 5.26. Enantioselectivity in the nucleophilic attack of enamine/carbanion, IV to the acceptor aldehyde, V. R_1= CH_3 for S-HPP and R_1= Ph for R-benzoin. .. 99

Figure 5.27. Atom numbers for charge distribution in QM models, (a) BFD-BA, (b) BA, (c) AA, (d) BFD-BA+AA, (e) BFD-BA+BA 101

Figure 5.28. Prereactive interactions before the second nucleophilic attack (*Model 8*).. 104

LIST OF TABLES

Table 4.1. Yields of O-benzoylated benzoins synthesized by Demir [30]. 32

Table 4.2. Relative energies (B3LYP) and thermodynamic contributions (kcal/mol) for the formation and the condensation of O-benzoylated cyanohydrin intermediate 7 with aldehyde 8 (T=298K)..... 42

Table 4.3 Mulliken charge distribution and geometric parameters for unsubstituted molecules (B3LYP/6-31G*). 45

Table 4.4. Chemical Reactivity Indices (B3LYP/6-31+G*)..... 50

Table 5.1. Model systems generated for MD simulation†..... 61

Table 5.2. Tested protonation scenarios and corresponding model systems in the study. 64

Table 5.3. Acetaldehyde atom names utilized in *Model 10*..... 70

Table 5.4. Benzaldehyde atom names utilized in *Models 3, 4, 5, 7,8* and *9*..... 70

Table 5.5. Ylide ThDP atom names utilized in *Models 1, 2,3,4,5* and *7*..... 71

Table 5.6. Enamine/carbanion atom names utilized in *Model 5,8,9* and *10*..... 73

Table 5.8. Enamine/carbanion and benzaldehyde atom names in *Model 10*..... 77

Table 5.9. Assigned protonation states of charged amino acids (PropKa predictions) 79

Table 5.10. Geometric parameters representing the octahedral geometry of the Mg (II) cation and the interactions of ThDP with 92

Table 5.11. NAC percentages for nucleophilic attacks. 93

Table 5.12. H_{bond} NAC percentages of hydrogen bonds in model systems. Refer to Figure 5.28 for the definition of bonds. 94

Table 5.13. Energetics of the second nucleophilic attack. 100

Table 5.14. Charge Distribution in QM Models. 102

LIST OF SYMBOLS/ABBREVIATIONS

ThDP	Thiamine diphosphate
BFD	Benzoylformate decarboxylase
MD	Molecular Dynamics
QM	Quantum Mechanics
DFT	Density Functional Theory
LYP	Lee- Yang-Parr
PME	Particle Mesh Ewald
PBC	Periodic Boundary Conditions
(*S*)-2-HPP	(*S*)-2-hydroxy propiophenone
DMB	3′-5′ dimethoxybenzoinyl
PDC	Pyruvate decarboxylase
BAL	Benzaldehyde lyase
GTO	Gaussian type orbitals
STO	Slater type orbitals
LCAO	Linear Combination of Atomic Orbitals
SCF	Self-consistent Field
AM1	Austin Model 1
PM3	Parametrized Model 3
LDA	Local Density Approximation
PCM	Polarized Continuum Model
SCRF	Self consistent Recation Field
IPCM	Isodensity Polarized Continuum Model
LST	Linear Syncronous Transit
HSAB	Hard Soft Acid Base
PDB	Protein Data Bank
VMD	Visual Molecular Dynamics
NAMD	Nanoscale Molecular Dynamics
IRC	Internal Reaction Coordinate
RMSD	Root Mean Square Deviation

NAC	Near-Attack Conformations
QM/MM	Quantum Mechanics/Molecular Mechanics
QM/MD	Quantum Mechanics/Molecular Dynamics

1. BENZOIN CONDENSATION

Acyloin, with a secondary hydroxyl group, and a neighboring hydrogen atom on the 2-position of the carbonyl compound, bears an important and common structural unit among many biologically active natural products (Figure 1.1).

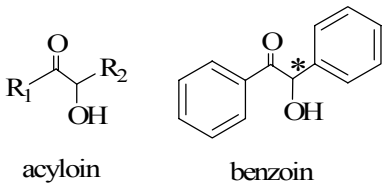


Figure 1.1. 2-hydroxy ketones; a) structure of an acyloin, R′ group can be aliphatic or aromatic, b) benzoin.

Desirable building blocks in synthetic organic chemistry, [1-6] acyloins are used in different fields due to their versatile functional groups, which may be easily transformed to other functionalities, e.g. diols, halo or amino derivatives and epoxides [7]. 2-hydroxy carbonyl compounds have been shown to be the inhibitors of proteases, and have been used for the treatment of Alzheimer's disease [8]. *Bupropion* synthesis is another example of the reactions utilizing 2-hydroxy ketones in their synthetic route. *Bupropion* is the active reagent of *Wellbutrin* ® (Glaxo Wellcome), and is synthesized by the stereo specific conversion of the corresponding 2-hydroxy ketone to 2-amino ketone. *Wellbutrin*® is marketed for the treatment of depression [9]. *Bupropion* was approved by the FDA for use as a smoking cessation aid under the name *Zyban* ®. 2-hydroxy ketones have also been reported to act as inhibitors for urease enzyme, which otherwise may lead to peptic ulcer in the presence of *Helicobacter pylori* [10].

Specifically, benzoin is mostly used in synthetic organic chemistry owing to its bifunctionality, asymmetric center, and the photolabile benzoyl group (in the case of O-benzoylated benzoin [3, 2]. It is an efficient photoinitiator [5], a starting material in

synthetic organic chemistry and is a conventional drug for healing cough. A useful feature of benzoinyl esters –namely 3',5'-dimethoxybenzoinyl (DMB) esters include extremely rapid photolysis, high quantum yield of conversion and generation of only a single photoproduct [3]. DMB esters are also utilized in lithographic synthesis [2] and as photolabile linkers [11].

Conventionally, benzoin is synthesized from benzaldehyde in the presence of cyanide ion and this method is one of the oldest carbon-carbon bond forming reactions (Figure 1.2). The first investigations on the synthesis of benzoin date back to 1832 when Wöhler and Liebig discovered the so-called benzoin condensation [6]. The mechanism of the conventional benzoin condensation reaction was first reported by Lapworth at the beginning of the last century [12]. The mechanism suggested by Lapworth was analyzed by Kuebrich [13] where, the kinetic steps and the factors affecting the course of the benzoin condensation reaction have been investigated. A detailed reaction mechanism has been suggested, however, neither of the steps has been clearly described as rate determining.

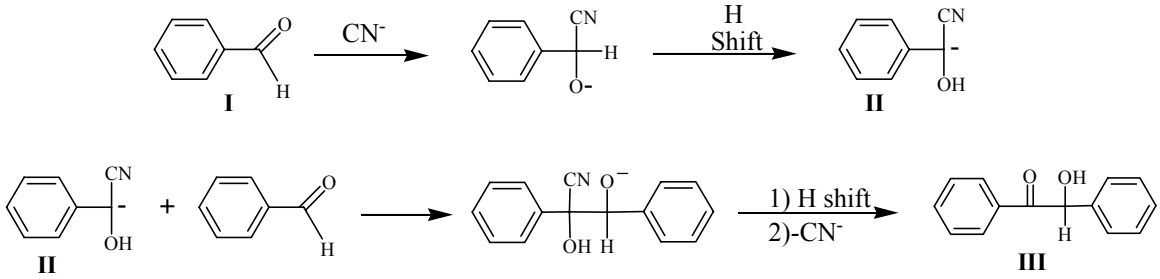


Figure 1.2. Conventional benzoin condensation.

The cyanide ion appears to be one of the strongest nucleophilic agents to catalyze the benzoin condensation [13, 14]. Baevsky in his report on the cleavage of α-diketones states that the cyanide ion sterically suits well for the attack to the carbonyl group at even hindered carbon and its cylindrical shape is particularly advantageous since it binds easily to the carbonyl group [14].

Although the conventional method provides easy access to the synthetically important benzoin molecules, the scope of this reaction is quite narrow. In the synthesis of

benzoin from two different aldehydes, Staudinger classifies aldehydes into two subgroups: aldehydes having a reactive carbonyl group and a non-mobile hydrogen atom (acceptor aldehyde) and aldehydes with a mobile hydrogen atom without an active carbonyl group (donor aldehydes) [15]. A benzaldehyde derivative to be utilized in benzoin condensation should be reactive enough to attack another benzaldehyde yet stable enough not to decompose under the reaction conditions [12, 14, 16, 17]. Therefore, depending on the electron donating or withdrawing ability of the substituents a limitation to the formation of some unsymmetrical benzoin derivatives arises [18]. The reaction produces a stable isomer where the carbonyl group is attached to the electron rich aromatic ring [19]. There are many examples in the literature, which present alternative synthesis methods for benzoin molecules [20, 21]. The synthesis of benzoin molecules takes place via two general mechanisms; either via the cyanohydrin or via hydrolyses of nitriles after treatment with Grignard reagents [22]. Synthetic routes to benzoin include usage of chiral or nonchiral thiazolium salts, oxidation of titanium enolates and silyl ethers, dihydroxylation of enolates, enzymatic reduction of α-diketones [23]. Furthermore, enzyme mediated benzoin condensation is reported to be an efficient synthetic route for enantioselective benzoin/acyloin products [18, 22, 24, 29].

Typically, an acyloin synthesis reaction in both enzymes mediated and cyanide ion catalyzed reactions, is initiated with the attack of the nucleophilic catalyst, **II** to the carbonyl carbon of the donor aldehyde (O_{donor}), **I** (Figure 1.3). Attack of the catalyst generates the alkoxide, **III,** which rearranges to *umpolung* -the German word for reversed polarity- reagent of the aldehyde, which is a resonance-stabilized carbanion called "*active aldehyde*" **IV**. This intermediate bears a negative charge on the carbon atom and behaves as a nucleophile at the second stage of the reaction [18].

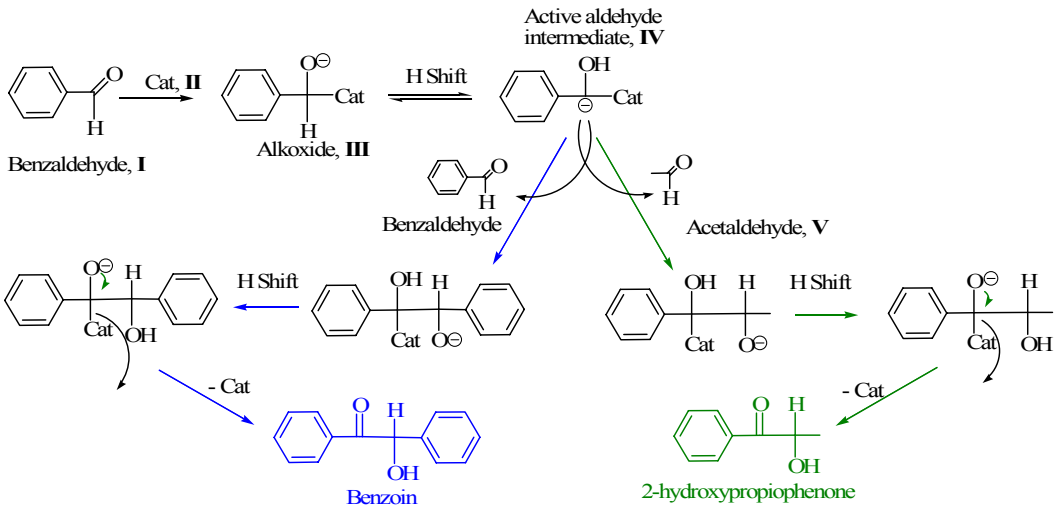


Figure 1.3. General mechanism for the catalytic acyloin condensation.

Active aldehyde intermediate, **IV**, is an example of one of the rare negatively charged carbanions in synthetic organic chemistry. Nucleophiles with the negative charge on the carbon atom are rarely encountered in organic reactions; this emphasizes the peculiarity of this intermediate and explains the special features of the benzoin condensation, where the difficult task of C-C bond formation and polarity reversal occur spontaneously.

The mobility of the hydroxyl proton in some cases introduces a practical difficulty in the utilization of the cyanohydrin intermediate. The limitations of the conventional method can be overcome by the usage of the O-benzoylated cyanohydrin to synthesize O-benzoylated benzoin. O-benzoylated cyanohydrin is generated from benzil and Baevsky has suggested a plausible mechanism for the synthesis of benzoin via O-benzoylated cyanohydrin [13, 16, and 26].

In the synthesis of the O-benzoylated benzoin **VI**, the benzil molecule **IV** and an acceptor aldehyde -benzaldehyde- are mixed in the presence of cyanide ion (Figure 1.4). The reaction is initiated by the attack of cyanide ion to the benzil molecule [26, 27, 28] After the addition of the cyanide ion to the benzil molecule; a rearrangement takes place ending up with the O-benzoylated cyanohydrin **V**, which is very similar to the cyanohydrin **II** formed in the conventional benzoin condensation. Besides its ability as a protecting

group [4], the presence of the benzoyl group assists the formation and the stability of cyanohydrin.

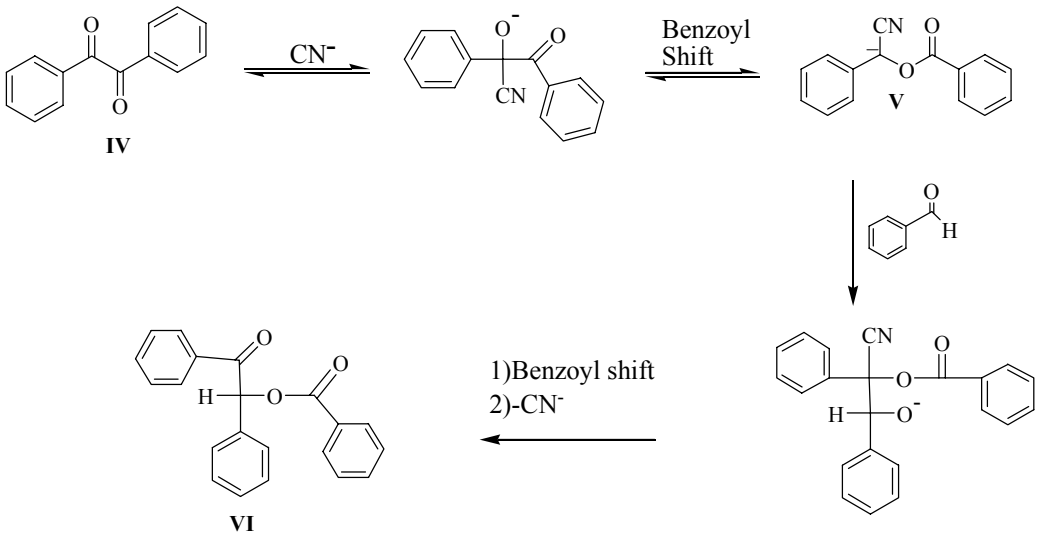


Figure 1.4. Synthesis of O-benzoylated benzoin from benzil.

Once it is formed, O-benzoylated cyanohydrin **V** can undergo a cleavage or a condensation reaction in the presence of an acceptor aldehyde depending on the solvent used. When the reaction is carried out in a polar protic solvent, the intermediate **V** cleaves into benzaldehyde and the corresponding ester of benzoic acid, whereas in polar aprotic solvents, it rearranges to yield O-benzoylated benzoin **VI** (Figure 1.5) [14].

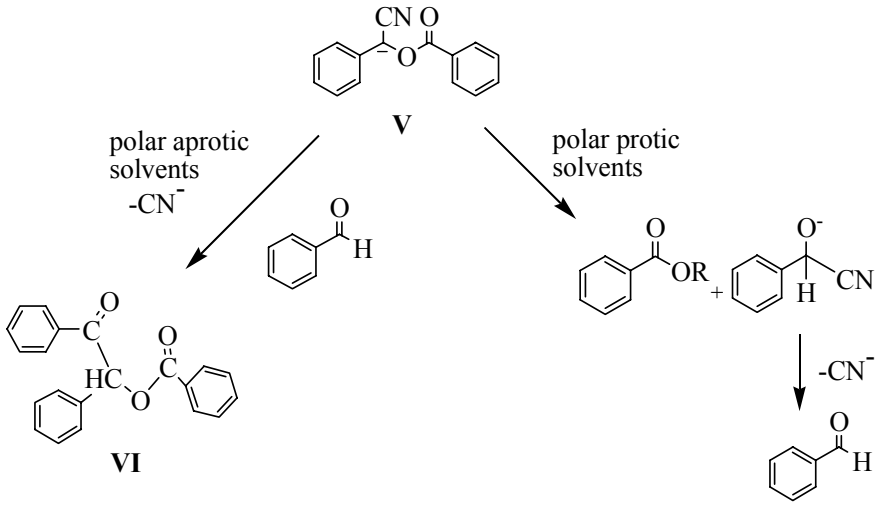


Figure 1.5. Solvent effect on the reactions of O-benzoylated cyanohydrin.

As an alternative to chemical methods, enantiomerically pure acyloins are also prepared enzymatically with thiamin diphosphate (ThDP) as the catalytic cofactor at the active center [29-33] of the enzyme. Breslow proposed a mechanism for the thiazolium catalyzed synthesis of acyloins (Figure 1.6) [31]. ThDP dependent enzymes are involved in many different pathways and catalyze a broad range of reactions. Carbon-carbon bond breaking (lyase activity) and unusual carbon-carbon bond formation (ligase activity) can be stated among the numerous capabilities of ThDP dependent enzymes [29, 30, 25, 32, 33]. Acyloin degradation and decarboxylation are examples for the former, whereas acyloin condensation is an example of the latter. The reaction catalyzed by ThDP dependent enzymes point out their potential as biocatalysts for the production of acyloins of high optical purity [33, 34, 35].

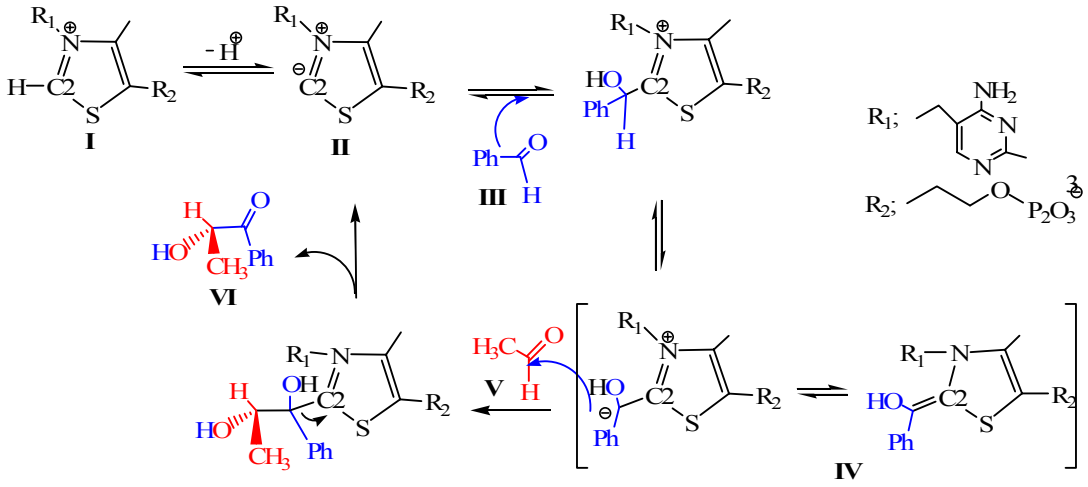


Figure 1.6. ThDP catalyzed acyloin condensation, **I**: ThDP, **II**: Ylide-ThDP, **III**: Benzaldehyde, **IV**: Enamine/Carbanion, **V**: Acceptor aldehyde (Acetaldehyde), **VI**: Acyloin ((*S*)-2-Hydroxypropiophenone, (*S*)-HPP).

Benzoylformate decarboxylase (BFD, E.C. 4.1.1.7) is a ThDP dependent enzyme that catalyzes the nonoxidative conversion of benzoylformate ($PhCOCO_2H$) to benzaldehyde (PhCOH) and carbon dioxide (CO_2) (Figure 1.7). Besides, BFD catalyzes the acyloin/benzoin reaction between benzoylformate or benzaldehyde acting as donor aldehyde, in the presence of an acceptor aldehyde with decarboxylation in the case of benzoylformate. A wide range of acyloin-reactions (where acceptor aldehyde is acetaldehyde) and benzoin reactions (where the acceptor aldehyde is benzaldehyde) are catalyzed by BFD, leading to the corresponding (*S*)-acyloins and (*R*)-benzoins, respectively. BFD also catalyzes ligation of a broad range of different aromatic, heteroaromatic, and even cyclic aliphatic and conjugated olefinic aldehydes as donor aldehydes to yield acyloin type products [34, 36].

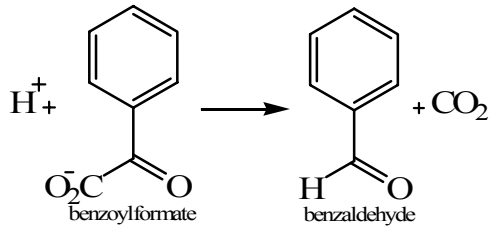


Figure 1.7. Main enzymatic function of BFD.

BFD has been found in bacteria such as *Pseudomonas putida*, *Acinetobacter calcoaceticus*, and *Pseudomonas aeruginosa* [33, 37-39]. The crystal structure of BFD from *Pseudomonas putida* is available at 1.6 Å resolution [33]. The spatial arrangement of the four subunits in the crystal is similar to pyruvate oxidase from *Lactobacillus plantarum* [40] and to pyruvate decarboxylase (PDC) from *Zymomonas mobilis* [41]. BFD is a component of the mandelate pathway, in which enzymes enable bacteria to grow using R-mandelic acid as carbon and energy source by converting it to benzoic acid [42-44], which is then metabolized by the α-ketoadipate pathway and the citric acid cycle (Figure 1.8).

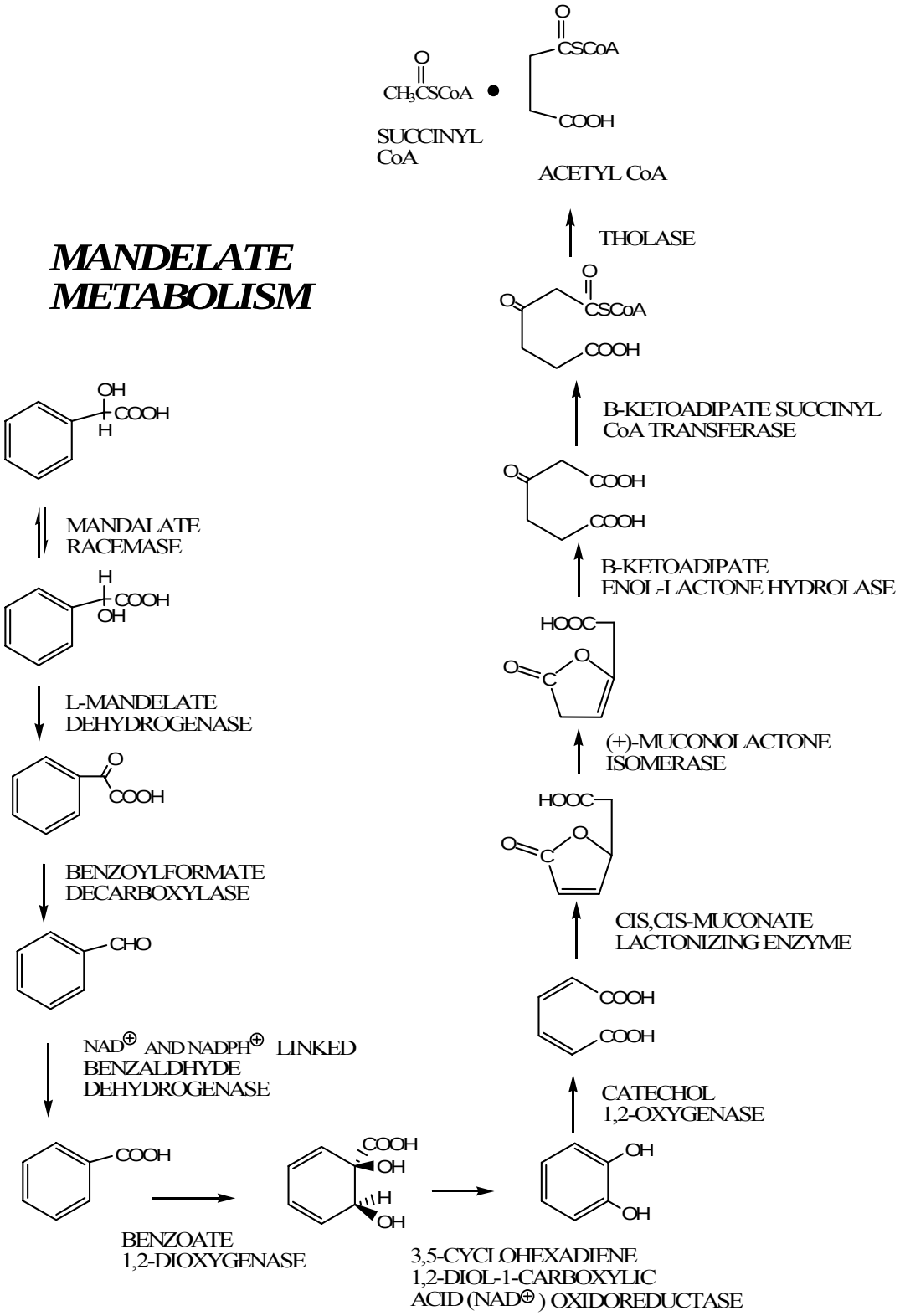


Figure 1.8. Mandalate metabolism in *Pseudomonas mobilis* [47].

2. OBJECTIVE AND SCOPE

In the first part of this study we use quantum mechanical tools to model the reactions of benzil derivatives, representing the donor aldehyde in the conventional benzoin condensation, with benzaldehyde derivatives as acceptor aldehydes. The mechanism of the reaction and the effect of ortho substitutions on the aromatic phenyl rings on benzil and on benzaldehyde have been discussed.

In the second part of this study, the ThDP catalyzed acyloin condensation in BFD environment is modeled with molecular dynamics and quantum mechanics. The protonation states of the catalytically active residues, the enantioselectivity of the intermediates are investigated, and the proton transfer mechanisms in the catalytic path are examined.

3. THEORETICAL BACKGROUND

3.1. Quantum Mechanics

Electrons and nuclei are considered as wave-like particles whose "waviness" and future state are mathematically represented by a set of wave functions obtained by an approximate solution of Schrödinger's equation. The Schrödinger equation is a fundamental postulate of quantum mechanics and can not be derived. Its predictions give excellent agreement with experimental results.

The concept of state function and time dependent state function for an n-particle system were introduced by the Austrian physicist Erwin Schrödinger (1887-1961) in 1926 [42]. The time dependent Schrödinger equation is;

$$-\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m_1}\left(\frac{\partial^2\Psi}{\partial x_1^{\,2}} + \frac{\partial^2\Psi}{\partial y_2^{\,2}} + \frac{\partial^2\Psi}{\partial z_3^{\,2}}\right) - \ldots - \frac{\hbar^2}{2m_n}\left(\frac{\partial^2\Psi}{\partial x_n^{\,2}} + \frac{\partial^2\Psi}{\partial y_n^{\,2}} + \frac{\partial^2\Psi}{\partial z_n^{\,2}}\right) + V\Psi \qquad (3.1)$$

in this equation $\hbar$ is Planck's constant divided by 2π, i is $\sqrt{-1}$, m is the mass of particle, x, y, z are coordinates and V is the potential energy of the system.

When the external potential of the system is independent on time then the wavefunction can be written as the product of the time part and the spatial part;

$$\Psi(r,t) = \psi(r)T(t) \qquad (3.2)$$

The time independent-Schrödinger equation can be written as;

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V \right] \Psi(r) = E\Psi(r) \tag{3.3}$$

for the cases where the potential is independent of time. It is usual to abbreviate the equation as:

$$H\Psi = E\Psi \qquad \text{where} \qquad H = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V \tag{3.4}$$

where H is named as Hamiltonian operator and E is the electronic energy. The solutions of the time-independent Schrödinger equation (3.3) are the time-independent wave functions.

The traditional *ab initio* techniques are based on the solution of the time-independent Schrödinger equation. However, Schrödinger equations can not be solved exactly for any molecular systems, mainly because of the failures in representing the exact wavefunction and spins of the electrons at the same orbital. Exact solutions are not computationally practical for large molecules. Approximations are introduced to electronic structure methods.

The masses of the nuclei are much greater than the masses of the electrons. The electrons can adjust almost instantaneously to any changes in the positions of the nuclei. The electronic wave function thus depends only on the positions of the nuclei and not on their momenta. This approximation is named as Born-Oppenheimer approximation [43]. Under Born-Oppenheimer approximation the total wave function for the molecule can be written as:

$$\Psi_{tot}(nuclei, electrons) = \Psi(electrons)\Psi(nuclei) \tag{3.5}$$

and the total energy equals to the sum of the nuclear energy (the electrostatic repulsion between the positively charged nuclei) and the electronic energy.

If the kinetic energy of nuclei is neglected then the Hamiltonian in atomic units is:

$$H = \sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\right)-\sum_{i=1}^{N}\frac{1}{r_i}+\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}+\sum_{\alpha=1}^{N}\sum_{\beta>\alpha}^{N}\frac{1}{r_{\alpha\beta}} \tag{3.6}$$

where the first term is the kinetic energy operator for the electrons, the second term is the operator for the attraction between electrons and nuclei, the third term is the repulsion between the electrons and the final term is the internuclear repulsion operator for an N particle system.

For a many-electron system, the interelectronic repulsion term in the potential energy makes it impossible to solve the Schrödinger equation exactly. The most widely used approximation to overcome this difficulty is the variation method. In the variational calculations a trivial wavefunction is selected and variational integral is calculated for the average value of the energy with the energy operator Hamiltonian. The variational theorem shows that well-behaved function giving the lowest value of variational integral provides the closest approximation to the ground state energy. The variation methods states that,

$$\int \phi^* \hat{H} \phi d\tau \geq E_{gs} \tag{3.7}$$

where E_{gs} is the ground state energy and ϕ is any normalized wave function.

3.1.1. Basis Sets

The molecular modeling programs for quantum chemical calculations build a set of functions to be occupied by the electrons assigned to the molecule. An electron in a molecule is described using a molecular orbital. The molecular orbital is expanded as a linear combination of atomic orbitals or, "*basis functions*" with the weights or coefficients to be determined. Basis functions collectively are called *basis set*.

There are two types of basis functions, commonly used in electronic structure calculation: gaussian type orbitals (GTO) and slater type orbital (STO). Slater orbitals correspond to a set of functions which decayed exponentially with distance from nuclei. Slater type orbitals can be approximated as linear combinations of Gaussian orbitals. Calculation of the integrals with the Gaussian Basis functions, leads to computational time savings.

The minimal basis set uses a minimum number of Gaussian functions to accommodate the electrons of the neutral atom. For example, in STO-3G minimal basis set, the Slater-type orbitals are represented by three Gaussian functions. The next improvement in the basis set is referred as the split-valance or double zeta basis sets. In this basis set, the valance orbitals are represented by more than one basis function. These orbitals with different spatial extentions allow the electron density to adjust to the particular molecular environment. In this study we have used the 6-31G* and 6-31+G*basis set in all of the calculations with 6-31+G* for single point corrections and solvent effect modeling studies. The number before the hypen indicates the number of gausian functions used to represent the core electrons, where valance electrons are represented by two orbitals with three and one gaussian orbitals. Another improvement accounts for distortions, e.g. polarization and diffusion caused by the interactions of neighboring atoms. As atoms are brought close together, the influence of the other nuclei will distort the electron density (the shape of the atomic orbitals) near the nucleus. This charge redistribution causes a polarization effect. For improved flexibility the spilt-valance basis set can be augmented with polarization functions. In polarization basis sets, the d orbitals are added to all heavy atoms. The polarization function is designated with a *, as in

the case of basis set that we have used.The additional basis functions are called *polarization functions*.

In some cases the basis functions used do not give adequate results. This is particularly the case in excited states and in anions where the electron density is more spread out over the molecule. To model this correctly, some basis functions, which themselves are more spread out have to be used. These additional basis functions are called *diffuse functions*. The diffuse functions are s and p – functions, denoted by + for heavy atoms and ++ for hydrogens and consequently go before the G.

In the linear combination of atomic orbitals (LCAO) assumption molecular orbitals are described as

$$\psi_i = \sum_{j=1,n} C_{ij}\phi_j \tag{3.8}$$

where ϕ_j represents basis functions, centered on each of the constituent atoms in a molecule, and C_{ij} is the coefficient of the j_{th} basis function in the i_{th} molecular orbital. *Ab initio* and semi-empirical calculations treat the linear combination of orbitals by iterative computations in which orbitals are improved from cycle to cycle until the electronic energy reaches a constant minimum value and the orbitals no longer change. At the end, self-consistent electric field (SCF) is established and energy of the system is minimized. The SCF does not allow the electrons to avoid each other but, assumes that their instantaneous positions are independent of one another. At the SCF level the electron-electron repulsion is actually overestimated. However, the overestimation is reasonably consistent so that its effects can be made to cancel by the use of proper comparison. In *ab initio* calculations, electron-electron repulsion is specifically taken into account.

3.1.2. Semiempirical Methods

The greatest portion of the time required to perform a computational calculation is spent calculating and manipulating the integrals encountered in solving the Schrödinger equation (3.1). Semiemprical methods neglect or approximate these integrals by use of parameters derived from experimental data and explicitly considering only the valance electrons of the system; the core electrons are subsumed into the nuclear core. Semiempirical methods are inherently dependent on the choice of parameters. The parameterization is tested against a limited set of molecules to ensure its accuracy.

There are a variety of semiemprical methods, characterized by their use of parameters derived from experimental data in order to simplify the approximation to the Schrödinger equation. The best known are AM1 (Austin Model 1), PM3 (Parametrized Method 3), MNDO (Modified Intermediate Neglect of differential overlap). Semiemprical methods are advantageous for a variety of modeling tasks especially for very large systems. Besides, they are used for the first treatment of large systems before handling them with computationally expensive methods.

However, semiemprical methods may only be used for systems where parameters have been developed for all of their component atoms. Semiemprical models are known to fail for treatment of hydrogen bonds, transition states and molecules containing atoms for which they are poorly parameterized.

Dewar and Thiel introduced the modified neglect of diatomic overlap (MNDO), which was based on NNDO [44]. MNDO utilizes monoatomic parameters derived from experimental molecular geometries, heats of formation, dipole moments and ionization potentials. MNDO however has a tendency to overestimate the repulsions between atoms separated by a distance equal to the sum of their van der Waals radii.

The inability of MNDO has led to Austin Model 1 (AM1) [45, 46]. In this model a term was added to MNDO to correct the excessive repulsions at van der Waals distances. For this purpose, each atom was assigned a number of spherical Gaussians, which were intended to mimic long-range correlation effects.

In this study, the semi-empirical PM3 (parameterized model 3, AM1 being considered the second) is used [47]. The parameters for PM3 model were derived using an automated parameterization procedure, using a large set of reference molecular data.

3.1.3. Density Functional Theory

DFT is based on theorems proposed by Kohn and Hohenberg [48, 49]. The first of the Kohn-Hohenberg theorems states that the electron density $\rho(r)$ determines the potential due to the nuclei. The second theorem introduces the variational principle. The electron density in this formalism is defined as:

$$\rho(x) = N\int ...\int \left|\Psi(x_1, x_2, ...x_n)\right|^2 dx_1 dx_2 ...dx_n \tag{3.9}$$

where ψ is the wave function and x is spin and spatial coordinates of electrons.

The electronic energy can be expressed as a functional of the electron density:

$$E[\rho] = \int v(r)\rho(r)dr + T[\rho] + V_{ee}[\rho] \tag{3.10}$$

where T[ρ] is the kinetic energy of the interacting electrons and $V_{ee}[\rho]$ is the interelectronic interactions. This equation can be rephrased as:

$$E[\rho] = \int v(r)\rho(r)dr + T_s[\rho] + J[\rho] + E_{xc}[\rho] \tag{3.11}$$

with $J[\rho]$ being the coulomb energy due to nuclear electron attraction, $T_s[\rho]$ being the kinetic energy of the non-interacting electrons and $E_{xc}[\rho]$ being the exchange-correlation

energy functional. The exchange-correlation functional is expressed as the sum of an exchange functional $E_x[\rho]$ and a correlation functional $E_c[\rho]$, although it contains also a kinetic energy term arising from the kinetic energy difference between the interacting and non-interacting electron systems. The last three terms in this equation are universal; they do not depend on the identity of the system.

In Kohn-Sham density functional theory, a reference system of independent non-interacting electrons in a common, one-body potential V_{KS} yielding the same density as the real fully-interacting system is considered. More specifically, a set of independent reference orbitals ψ_i satisfying the following independent particle Schrödinger equation are imagined:

$$\left[-\frac{1}{2}\nabla^2 + V_{KS}\right]\psi_i = \varepsilon_i\psi_i \tag{3.12}$$

with the one-body potential V_{KS} defined as:

$$V_{KS} = v(r) + \frac{\partial J[\rho]}{\partial \rho(r)} + \frac{\partial E_{xc}[\rho]}{\partial \rho(r)} \tag{3.13}$$

$$V_{KS} = v(r) + \frac{\rho(r')}{|r - r'|}dr' + v_{xc}(r) \tag{3.14}$$

where $v_{xc}(r)$ is the exchange-correlation potential. The independent orbitals ψ_i are known as Kohn-Sham orbitals and give the exact density by:

$$\rho(r) = \sum_{i}^{N} |\psi_i|^2 \tag{3.15}$$

if the exact form of the exchange-correlation functional is known. Various approximations of E_{xc}, which is a function of $\rho(r)$ separate the different DFT methods from each other, starting with the local density approximation (LDA). This approximation assumes the system is a homogeneous electron gas and E_{xc} $[\rho(r)]$ depends only on the local value of the electron density. The energy expression is:

$$E[\rho(r)] = T_s[\rho(r)] + \int \rho(r) v(r) dr + J[\rho(r)] + E_{xc}[\rho(r)] + E_b \tag{3.16}$$

where E_b is the electrostatic energy of the positive background. Since the positive charge density is the negative of the electron density due to uniform distribution of particles, the energy expression is reduced to:

$$E[\rho] = T_s[\rho] + E_{xc}[\rho] \tag{3.17}$$

$$E[\rho] = T_s[\rho] + E_x[\rho] + E_c[\rho] \tag{3.18}$$

The kinetic energy functional can be written as:

$$T_s[\rho] = C_F \int \rho(r)^{5/3} dr \tag{3.19}$$

where C_F is a constant equal to 2.8712. The exchange functional is given by:

$$E_x[\rho] = -C_x \int \rho(r)^{4/3} dr \tag{3.20}$$

with C_x being a constant equal to 0.7386. The correlation energy, $E_c[\rho]$, for a homogeneous electron gas comes from the parameterization of the results of a set of quantum Monte Carlo calculations.

LDA can successfully determine the optimized geometries [50, 51] and harmonic frequencies [52, 53]. However, the energies calculated with the LDA functional are rather poor due to the assumption of a homogeneous electron gas in the system [54].

The closed shell Lee-Yang-Parr (LYP) correlation functional [55] is given by:

$$E_c = -a \int \frac{1}{1 + d\rho^{-1/3}} \left\{ \rho + b\rho^{-2/3} \left[C_F \rho^{5/3} - 2t_w + \left(\frac{1}{9} t_w + \frac{1}{18} \nabla^2 \rho\right) \right] e^{-c\rho^{-1/3}} \right\} dr \tag{3.21}$$

where

$$t_w = \frac{1}{8} \frac{|\nabla \rho(r)|^2}{\rho(r)} - \frac{1}{8} \nabla^2 \rho \tag{3.22}$$

and a=0.04918, b=0.132, c=0.2533 and d=0.349.

The mixing of LDA, B88, E_x^{exact} and the gradient-corrected correlation functionals to give the hybrid functionals [56] involves three parameters:

$$E_{xc} = E_{xc}^{LDA} + a_0 \left(E_x^{exact} - E_x^{LDA}\right) + a_x \Delta E_x^{B88} + a_c \Delta E_c^{non-local} \tag{3.23}$$

where ΔE_x^{B88} is the Becke's gradient correction to the exchange functional. In the B3LYP functional, the gradient-correction ($\Delta E_c^{non-local}$) to the correlation functional is included in LYP. However, LYP contains also a local correlation term which must be subtracted to yield the correction term only:

$$\Delta E_c^{non-local} = E_c^{LYP} - E_c^{VWN} \tag{3.24}$$

where E_c^{VWN} is the Vosko-Wilk-Nusair correlation functional, a parameterized form of the LDA correlation energy based on Monte Carlo calculations. The empirical coefficients are a_0=0.20, a_x=0.72 and a_c=0.81 [56]

3.1.4. Continuum Solvation Model

Electronic structure methods are aimed at solving the Schrödinger equation for a single or a few molecules. Physically, this corresponds to the situation occurring in the gas phase. Experimentally, however, the majority of chemical reactions are carried out in solution. In continuum methods, the solvent is represented as a continuous polarizable medium of fixed dielectric constant ε, and solute (M) is embedded in a cavity of certain shape and size. The solute charge distribution interacts with the medium (dispersion interactions), polarizing it and creating a reflection charge distribution on the cavity surface (the *reaction field*) which, in turn, will interact electrostatically with the solute leading to the net stabilization [57]. The solvation (free) energy may be written as:

$$\Delta G_{solvation} = \Delta G_{cavity} + \Delta G_{dispersion} + \Delta G_{electrostatic} \tag{3.25}$$

In this representation, ΔG_{cavity} represents the energetic cost of creating a cavity in the medium producing a destabilization effect. Dispersion interactions between solvent and

solute add stabilization to solvation free energy term expressed as $\Delta G_{dispersion}$. The latter electrostatic term, $\Delta G_{electrostatic}$ has a stabilization effect and furthermore it appears to be responsible for the main structural changes of the solute.

Continuum salvation models are usually based on the Self-Consistent Reaction Field (SCRF) methods. SCRF methods can be classified based on the following characteristics [58]:

1. How the size and shape of the solute cavity is defined.
2. How the cavity creation and the dispersion contributions is calculated.
3. How the charge distribution of solute **M** is represented.
4. How the solute **M** is described classically or quantum mechanically.
5. How the dielectric medium is described.

The simplest SCRF model is the *Onsager* reaction field model. The basic assumption made in this model is that the solute is placed in a spherical cavity of radius a_0 inside the solvent, cavity/dispersion effects are neglected and only the net charge and dipole moment of the molecule are taken into account. The latter is described as a homogeneous, polarizable medium of constant dielectric constant. The solute dipole moment induces a dipole moment of opposite direction in the surrounding medium. Polarization of the medium in turn polarizes the charge distribution in the solute. The reaction field generated in this manner is proportional to the molecular dipole moment and inversely proportional to the third power of the radius of the solute cavity:

$$\Delta G_{electrostatic} = -\frac{(\varepsilon - 1)\mu^2}{(2\varepsilon + 1)a_0^3} \tag{3.26}$$

where a_0 is the cavity radius, ε is the dielectric constant and μ is the molecular dipole moment.

In polarized continuum solvation models (PCM), the solute is embedded in a cavity surrounded by an infinite polarizable dielectric. The Schrödinger equation for such systems is written as:

$$\left[H^0 + V_R\right]\Psi = E\Psi \tag{3.27}$$

where H^0 is the Hamiltonian of the solute in vacuo, V_R is a perturbation due to solvent and Ψ is the solute wave function in solution. The term V_R depends on the nuclear and electronic charge distribution.

IPCM with B3LYP geometries are known to yield results closer to the experimental data [59]. In this methodology, the cavity is defined by an isosurface of the total electron density calculated at the level of theory being applied. The solvent is considered as a uniform dielectric and the solute is assumed to occupy a cavity determined self consistenly from an isodensity surface. The IPCM method involves the calculation of point charges on the cavity surface which mimics the reaction field. The magnitude of these charges is proportional to the derivative of the solute electrostatic potential at each surface point. The point charges are then included in the one electron Hamiltonian which includes polarization of the solute. This process is repeated until the surface charges are self consistently equilibriated in the solute charge distribution.

This approach has both advantages and disadvantages. The advantages are considerable, since it allows the geometry and dipole moment to change under the influence of the medium, provides the possibility of calculating vibrational frequencies of solvated species quantum mechanically and does not require any empirical data as opposed to all force field methods. However, this approach can not reproduce specific solute-solvent interactions.

3.1.5. Chemical Reactivity Indices

The sensitivities of electron density to structural perturbations and responses to the changes in the external conditions are rather more important in reflecting the reactivity of a system. Many of the well known empirical but important chemical concepts such as the chemical potential (μ), electronegativity (χ), hardness (η), softness (*S*) appear naturally within the framework of DFT [60]. In general, these quantities correspond to the linear responses of the electron density with respect to changes in external potential or number of electrons. Responses with respect to the external potential are local properties and they reflect site selectivity, such as Fukui function (*f)* and local softness (s). Responses with respect to the number of electrons are global properties and are related to the overall molecular stability and reactivity, such as global softness (S). The parameters mentioned above emerge as a useful tool for rationalizing, interpreting and predicting diverse aspects of chemical bonding and reaction mechanisms. The mathematical formalism of these quantities and their applications to concrete problems has been well presented [61]. Nucleophilic Fukui functions, f^+ are calculated in this study. This function is introduced as the left hand side derivative of the electron density due to the fact that the electron density is discontinuous with respect to the number of electrons. The usual finite difference approximation approximation [62] has been followed in order to evaluate *f, S, s*. The following relationships have been used in this study.

$$f^+ \approx q(N+1) - q(N) \tag{3.49}$$

$$f^- \approx q(N) - q(N-1) \tag{3.50}$$

$$\eta \approx \frac{(I-A)}{2} \tag{3.51}$$

$$S = \frac{1}{\eta} \tag{3.52}$$

$$s = fS \tag{3.53}$$

where $q(N+1)$, $q(N-1)$ and $q(N)$ are the charges on the concerned atom of the molecule with *N+1, N-1* and *N* electrons respectively, *I* is the ionization potential and *A* is the electron affinity. The B3LYP functional was also shown to yield a good performance in the computation of atomic charges, dipole moments, Fukui functions, ionization energies, electron affinities and hardnesses [63]. Mulliken population analysis was used in the calculation of atomic charges. B3LYP/6-31+G* methodology has been used to derive chemical reactivity indices.

3.2. Molecular Dynamics

Molecular dynamics (MD) method has the main justification that statistical ensemble average of a system parameter is equal to the time average of that parameter, which is known as *Ergodic theory*. One can calculate the time averages by molecular dynamics simulation, but the experimental observables are assumed to be the ensemble averages. In statistical mechanics the time averages are defined as the ensemble averages. The average value of any property during time evolution is:

$$\langle A \rangle = \frac{1}{M} \sum_{1}^{M} A(t_i) \tag{3.54}$$

where M is the number of times the property is sampled.

If the system is allowed to evolve in time indefinitely, it will eventually pass through all possible states and the above equation becomes [64,65]:

$$\langle A \rangle = \lim \frac{1}{t} \int_{t_0}^{t_0+t} A(\tau) d\tau \tag{3.55}$$

assuming Ergodic hypothesis to be valid and independent of choice of t_0.

The state of any classical mechanical system can be completely described by means of specifying the positions and momenta of the all particles. The relationship between two positions in any time interval is given:

$$q(t_2) = q(t_1) + \int_{t_1}^{t_2} \frac{p(t)}{m} dt \tag{3.56}$$

$$v = \frac{p}{m} \tag{3.57}$$

Similarly, using Newton's Second Law of Motion, the relationship between any two momentum vectors is:

$$p(t_2) = p(t_1) + m \int_{t_1}^{t_2} a(t) dt \tag{3.58}$$

$$a = \frac{F}{m} \tag{3.59}$$

To measure an observable quantity in an MD study, one must be able to express this observable as a function of the position and momenta of the particles in the system. For example, the instantaneous value of the temperature is related to the kinetic energy via the particle's momenta as follows:

$$Kinetic\ Energy = \sum_{i=1}^{N} \frac{|p_i|^2}{2m_i} = \frac{k_B T}{2} 3N \tag{3.60}$$

In MD, successive configurations of the system are generated by integrating Newton's law of motion. For this purpose, the forces acting on the atoms should be calculated, and these are derived from a potential energy *V(R)*. The result is a trajectory that specifies how the positions and velocities of the particles in the system vary with time.

$$F_{x_i} = \left[-\frac{\partial V(r)}{\partial r} \right] \tag{3.61}$$

The interaction between particles is described by a force field. The force fields provide approximations for the forces acting on each atom. A force field is a simple representation of the intra and intermolecular forces within the system. A force field refers to the functional form and parameter sets used to describe the potential energy of a system of particles. Typical examples of force fields are AMBER [66] CHARMM [67] GROMOS [68]. Parameters of these force fields are determined by quantum chemical calculations combined with thermodynamical data. The force filed is composed of various contributions like bonded or valance terms (bond stretching, angle bending and torsion angle) and non bonded terms (van der Waals and Coulomb forces) all of which contain empirical parameters fitted to the results of experimental studies or high level calculations. The potential energy V(R) is defined as:

$$V(r) = E_{bond} + E_{angle} + E_{torsion} + E_{non-bonded} \tag{3.62}$$

where specific contributions are explained in the following paragraphs.

The potential function representing the non-bonded energy is formulated as a sum over interactions between the particles of the system. The simplest choice is the pair potential in which the total potential energy can be calculated from the sum of energy contributions between the pairs of atoms. An example is the non-bonded Lenard-Jones potential used for calculating van der Waals forces. The Lennard-Jones [69] potential is the most commonly used form of potentials representing continuous and differentiable pair potentials for van der Waals type interactions. In the presence of electrostatic charges, an appropriate Coulomb potential should be added.

$$E_{LJ}(i,j) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{3.63}$$

$$E_{Coulomb}(i,j) = \frac{1}{4\pi\varepsilon_o}\frac{q_i q_j}{r_{ij}} \tag{3.64}$$

where σ_{ij} is the separation distance at which the energy is zero and ε_{ij} is the potential energy well depth, q_i and q_j are the charges and ε_0 is the permittivity of space.

In the definition of bonding interactions in potential energy, the bonds and angles are typically described as harmonic oscillators. Thus, the potential energy terms are; *i)* bond stretching, E_{bond}, between two covalently bonded atoms:

$$E_{bond} = \sum_{bonds} K_r (r - r_{eq})^2 \tag{3.65}$$

where r is the bond length between bonded atoms, and the parameters K_r and r_{eq} are defined for each type of pair atoms, such that the former represents the spring constant, and the latter is the average bond length, *ii)* bond-angle bending, E_{angle}, term ensures that the bond angle fluctuates around its equilibrium value θ_{eq}:

$$E_{angle} = \sum_{angles} K_\theta (\theta - \theta_{eq})^2 \tag{3.66}$$

where θ is the angle between the atom triplet *i-j-k*, with atoms *i-j* and *j-k* being covalently bonded. K_θ controls the strength of the fluctuations around θ_{eq} and they are defined for each type of atom triplets. *iii)* The dihedral-angle term, $E_{dihedral}$ controls the rotations around the torsional angles and the barrier crossings between its n minima:

$$E_{dihedral} = \sum_{dihedrals} \frac{V_n}{2}\left[1 + \cos(n\varphi - \gamma)\right] \tag{3.67}$$

to set the interactions for the quadruple of atoms *i-j-k-l*. The angle ψ is defined as the angle between planes *i-j-k* and *j-k-l*. V_n, γ and n are set for each type of atom quadruplets.

The modes associated with stretching of bonds of polar hydrogens in the bio-molecules and the bond stretching and angle bending in the solvent (water) molecules can

be eliminated altogether with proper constraining using SHAKE [70] and RATTLE [71]. RATTLE is an iterative method involving constraining the velocities in addition to SHAKE- an iterative method for solving constrained differential equations in MD such that the positions of atoms satisfy the distance constraints.

During a classical MD simulation, the most expensive (or CPU intensive) task is the evaluation of the forces as a function of the internal coordinates of the particles. By employing Particle Mesh Ewald (PME) method this computational cost can be reduced [72]. The PME provides a method for computing a full representation of long-range electrostatic interactions [73]. Its applications have led to stable nucleic acid dynamics trajectories [74] without need for other restraints.

3.2.1. Fluctuations of Single Atoms

X-ray crystallographic B-factors, B_k, arise from the fluctuations of the individual atoms k around their space-group symmetric positions as given by

$$< \Delta r_k^{\,2} >= \frac{3}{8\pi^2} B_k \tag{3.68}$$

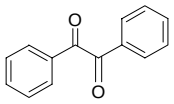
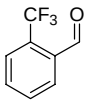
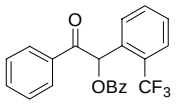
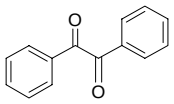
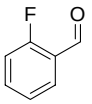
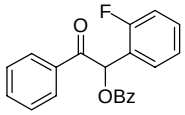
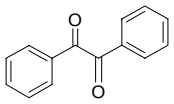
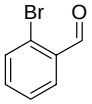
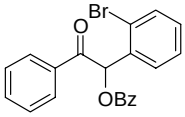
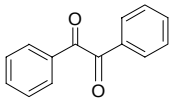
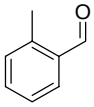
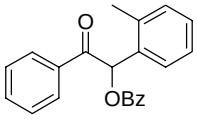
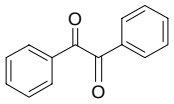
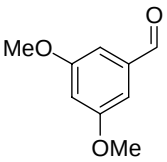
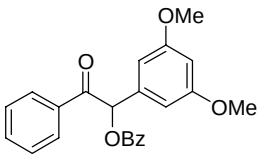
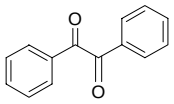
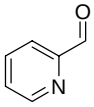
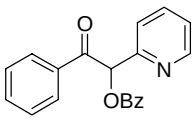
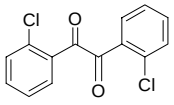
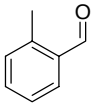
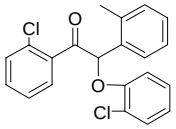
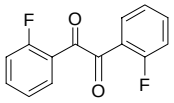
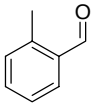
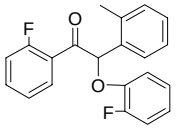
The fluctuations provide information on the translational and rotational mobility of individual atoms in the equilibrium state. The protein fluctuations were derived from the trajectories, and the results are compared with the experimental B-factors in the text to check the consistency of the proposed systems with the experimental data.

4. COMPUTATIONAL STUDY OF THE SYNTHESIS OF BENZOIN DERIVATIVES FROM BENZIL

4.1. Introduction

Benzil (1,2-diphenylethane 1,2-dione) undergoes cyanide catalyzed condensation with benzaldehyde to yield O-benzoylated benzoin (2-benzoyl-1,2-diphenylethanone). The studies of Kuebrich [13] on the kinetics and mechanism of the benzoin condensation (from benzaldehyde) and on the nature of cyanohydrin, [16] the results of Baevsky [14] on the cleavage of α-diketones and the investigation of Demir on the steric and electronic effects of the substituents have shed light on the formation of benzoin from benzil [28] The experimental results of Demir are summarized in Table 4.1.

Table 4.1. Yields of O-benzoylated benzoins synthesized by Demir [28].

Entry	Benzil	Aldehyde	Benzoin	% Yield
1				98
2				99
3				95
4				93
5				83
6				78
7				82
8				72

In this study, the kinetic and thermodynamic features of the experimentally suggested benzoin condensation mechanism has been modeled with PM3 and verified with B3LYP. It is the task of the first part of this study to model the reactions of benzil derivatives, representing the donor aldehyde in the conventional benzoin condensation, with benzaldehyde derivatives as acceptor aldehydes. The effect of the substituent on the reaction yield has been rationalized by considering two benzil derivatives; 1,2-bis(2-chlorophenyl)ethane-1,2-dione and 1,2-bis(2-fluorophenyl)ethane-1,2-dione and three benzaldehyde derivatives; o-fluoro benzaldehyde, o-methyl benzaldehyde and 2-pyridinecarboxaldehyde. The effect of the solvent has been modeled by using the isodensity-surface polarizable continuum (IPCM) model. Reactivity descriptors have been

utilized to justify the reactivity differences of the various substituents. The reaction of unsubstituted benzil molecule with unsubstituted benzaldehyde has been modeled for comparative purposes. Within each class of compounds, the energy barriers of the rate determining steps have been considered in order to explain the experimentally observed trends by Demir *et al.* [28]. The reactivity descriptors, such as softness and Fukui functions, have also been utilized in order to justify the findings based on the energy barriers. The content of this section has been published in the *International Journal of Quantum Chemistry* in 2006 [75].

4.2. Computational Methodology

The reaction of benzil with benzaldehyde has been modeled with PM3 [49,76] The stationary points have been located along the addition of cyanide ion to benzil yielding an O-protected cyanohydrin intermediate, which reacts further with benzaldehyde. For all the ground state molecules on the reaction path, a conformer search has been done with the semi-empirical PM3 [49, 76] method and 20 low energy conformers have been optimized with B3LYP/6-31G* without any symmetry consideration. In each case, the structure corresponding to the global minimum is considered in the text. The transition states were located by using the linear synchronous transit (LST) option and for each transition structure a conformational search has been performed with PM3. For this purpose the critical distances have been frozen and all the possible conformers have been located. Then, the constraints have been released and these structures have been used for reoptimization with B3LYP/6-31G*.Gaussian 98, was used for full optimization of the stationary geometries located with PM3 [77, 78]. In combination with the proper choice of the basis set, B3LYP is known to be adequate for nucleophilic reactions, usually being superior to ab initio methods [77]

Unless otherwise stated the geometrical parameters and energetics in the text are based on the results with the B3LYP methodology. Single point energy calculations (B3LYP/6-31+G*//B3LYP/6-31G*) have also been carried out to refine the energetics. Vibrational frequency calculations were used to characterize all stationary points as either

minima (no imaginary frequencies) or transition states (with only one imaginary frequency). Vibrational frequencies were also used to evaluate zero point energies and the thermodynamic data. Intrinsic reaction coordinate (IRC) calculations were traced to ensure that the transition states connect the proper minima [79].

The effect of solvation on reaction energetics was determined by means of single point self consistent reaction field (SCRF) calculations using the isodensity-surface polarized continuum model (IPCM) at the B3LYP/6-31G* level with a relative permitivity of 36.75 for dimethyl formamide as the solvent.

4.3. Results and Discussion

To the best of our knowledge we report the first computational modeling of this reaction in Figure 4.1. The reaction starts with the attack of the nucleophilic cyanide ion **2** to the carbonyl carbon of benzil, **1**, to yield the oxyanion **5**. Then, compound **5** rearranges to the O-benzoylated cyanohydrin, **7**, after the migration of the benzoyl group through a 3-membered transition structure **6**. This migration is initiated by the attack of alkoxy oxygen atom to the neighboring carbonyl carbon atom. The formation of **7** is a process in which the electrophilic carbonyl carbon of the benzil molecule **1** is converted into a nucleophilic carbon. The next stage of the reaction has been postulated as the attack of the nucleophilic cyanohydrin **7** to the electrophilic center in the benzaldehyde molecule **8.** The C-O double bond of the aldehyde opens while a new carbon-carbon single bond forms yielding another oxyanion, **10**. In compound **10**, an intramolecular cyclization through a five membered cyclic ring takes place, yielding compound **12**. This step constitutes the first stage of the migration of the benzoyl group towards the alkoxy oxygen. In the transition state **13** the neighboring C-O bond lengthens with the assistance of the electron withdrawing cyanide group and yields compound **14**. Finally, the catalyst cyanide ion leaves and the O-benzoylated benzoin **16** is produced.

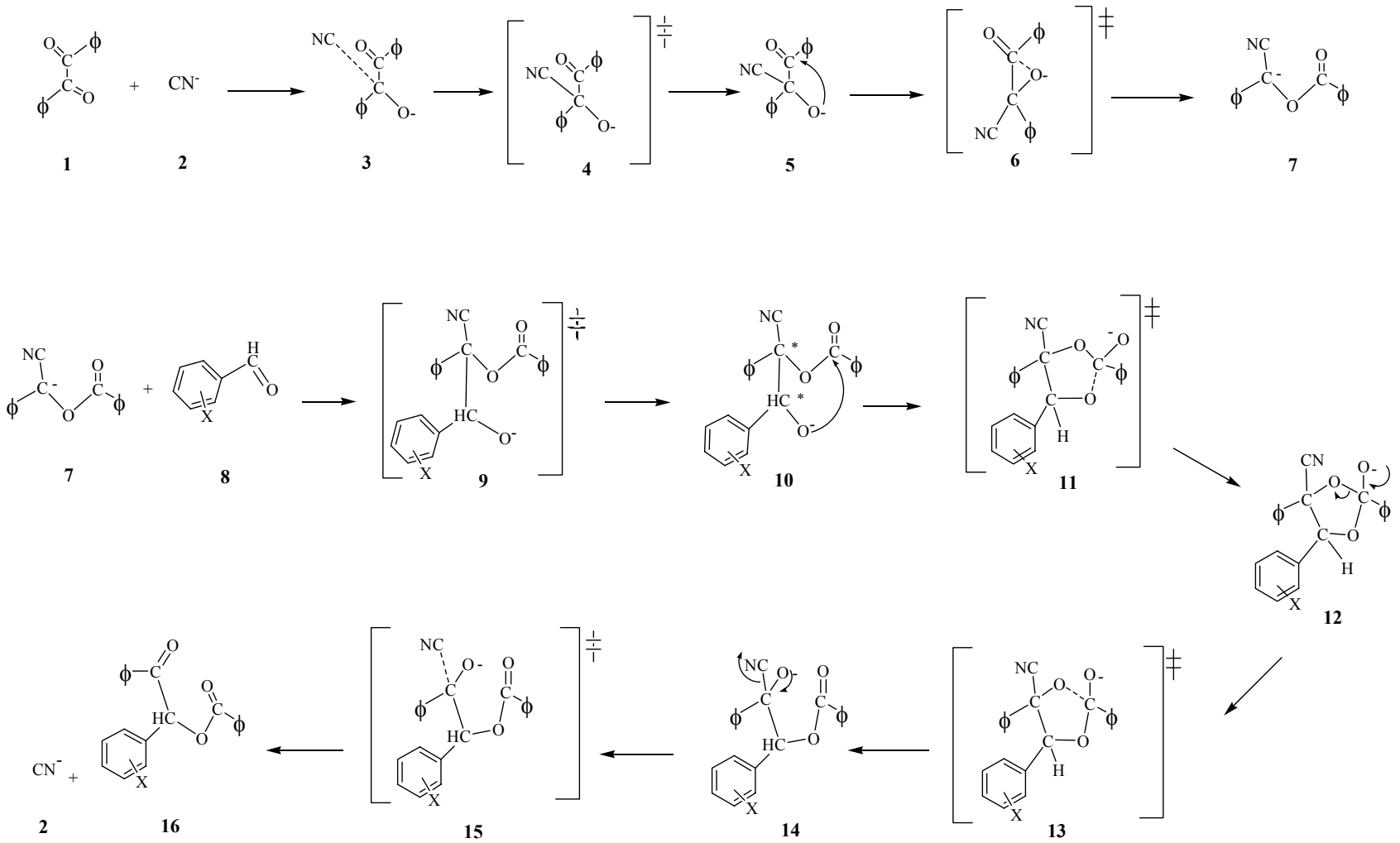


Figure 4.1. Computationally modeled mechanism for the benzoin condensation from benzil.

The reaction path shown in Figure 4.1 has been modeled with PM3 and the results are presented in Figure 4.2 and 4.3. The high-energy barrier for the formation of O-benzoylated cyanohydrin is due to the energy used to break the C-C bond and to overcome an angle strain. For this step the activation barrier is 24.0 kcal/mol. The formation of compound **7** is exothermic by 13.2 kcal/mol, indicating the stable nature this compound. A stable epoxy-like structure corresponding to a local-minimum between molecules **5** and **7** could not be located.

In the next step, the two achiral molecules **7** and **8** condense to form compound **10** with two asymmetric centers (Figure 4.3). For the unsubstituted benzaldehyde, the calculated activation barrier for the formation of compound **10** is 19.2 kcal/mol. The second high activation barrier belongs to this step due to the forming carbon-carbon single bond. The heat of reaction for the formation **10** amounts to 18.7 kcal/mol.

The activation barriers of the subsequent steps are lower up to the compound **14**. These steps are rearrangements on the way to produce O-benzoylated benzoin **16**. The process ends up with the removal of the catalyst; the cyanide ion. The activation barrier for the intramolecular nucleophilic attack of negatively charged oxygen atom to the neighboring carbonyl group at compound **10** to yield **12** is 4.3 kcal/mol and heat of reaction is 8.2 kcal/mol. The formation of a **5** membered ring does not suffer from ring or angle strain; therefore the activation barrier for this step is low. The C-O bond in this cyclic intermediate cleaves with an energy of activation of 5.5 kcal/mol to yield **14**. The electron withdrawing cyanide ion and the negative charge on the oxygen atom assist this cleavage. The activation barrier for the step describing the removal of the cyanide ion from compound **14** is postulated to be artificially high. The negatively charged cyanide ion is expected to be stabilized by the counter ion or the solvent molecules.

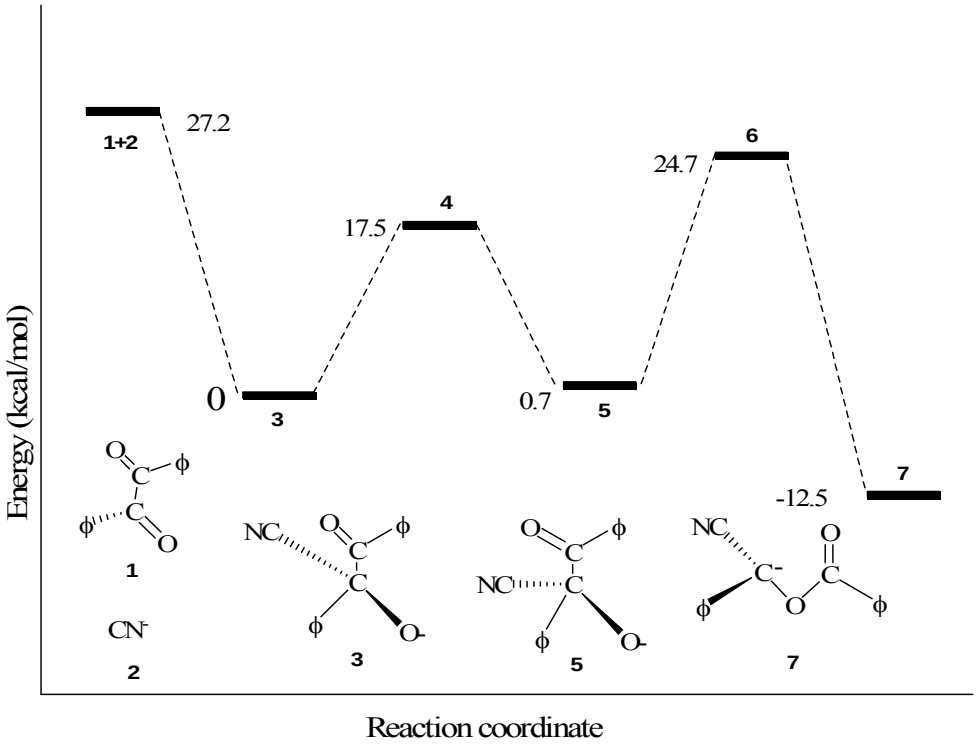


Figure 4.2. Energy barriers for the formation of cyanohydrin (PM3

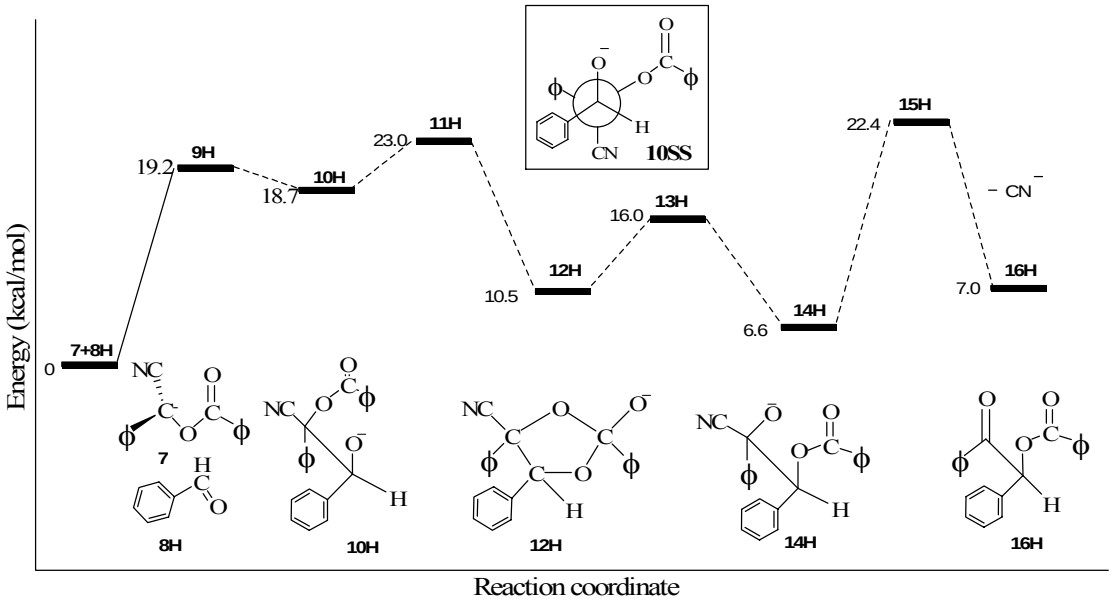


Figure 4.3. Energy barriers for the condensation of cyanohydrin with benzaldehyde (PM3).

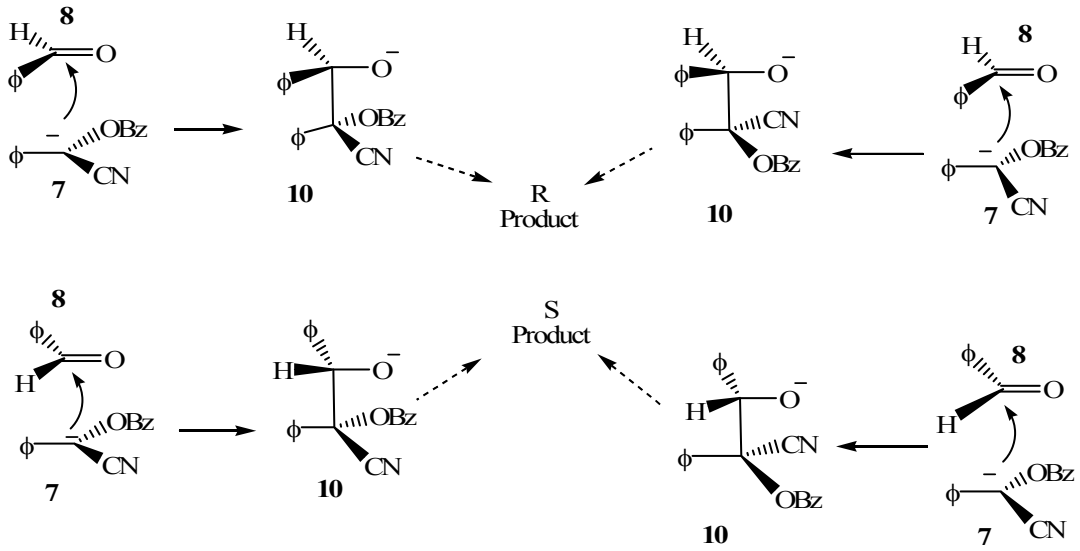


Figure 4.4. Various approaches of the cyanohydrin intermediate **7** to the si or re face of the acceptor benzaldehyde **8**.

Since the synthetic route does not include any chiral entity for the reaction to be streoselective the product is expected to be a mixture of streoisomers (*R* and *S*) (Figure 4.4). The side of attack of O-benzoylated cyanohydrin to the benzaldehyde molecule has to be rationalized in order to understand the role of the various streoisomers on the yield of the reaction. In this study, the reaction path yielding the *S* stereoisomer of O-benzoylated benzoin (*SS* in the case of compound **10**) has been modeled. For the sake of completeness, we have also modeled the reaction path yielding the *SR* diastereomer of compound **10** (Figure 4.5). The trend in the energy barriers for the stationary structures in Figure 4.5 resembles the one in Figure 4.3 the condensation of **7** with **8H** is still the slowest step along the reaction path.

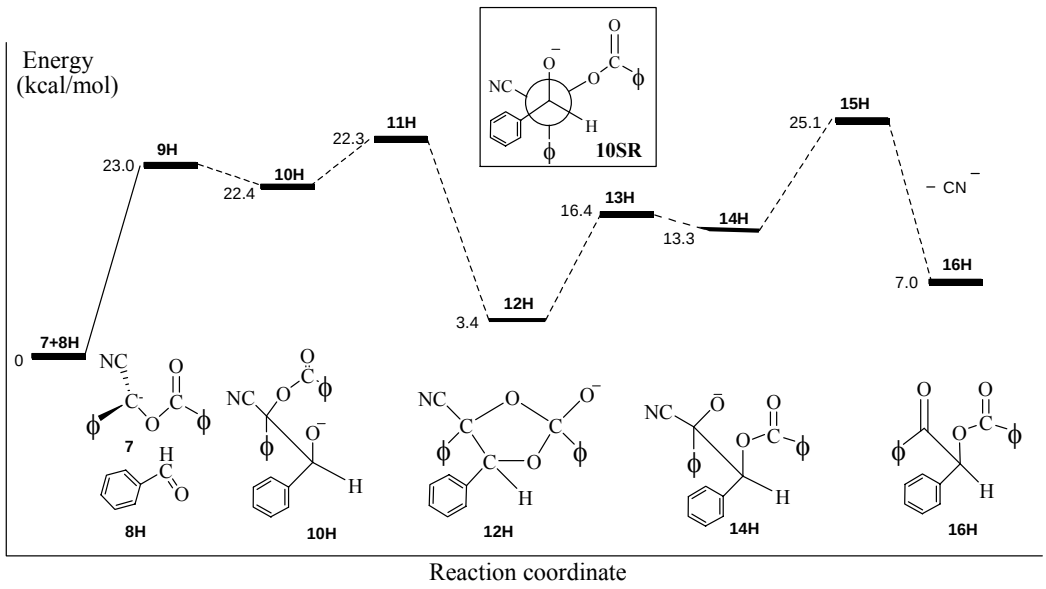


Figure 4.5. Energy barriers for the condensation of O-benzoylated cyanohydrin with benzaldehyde: compound **10** is in *SR* conformation (PM3).

4.3.1. Elucidation of the Reaction Mechanism with Unsubstituted Molecules

The reaction path displayed in Figure 4.1 exhibits six barriers among which four are high (**3** to **5**, **5** to **7**, **8** to **10** and **14** to **16**) and two are low (**10** to **12** and **12** to **14**). The latter belong to facile intra-molecular rearrangements. Addition and removal of the cyanide ion (**3** to **5** and **14** to **16**) are supposed to be triggered by the presence of the counter ion and the solvent. The addition of cyanide is experimentally stated to be fast; therefore the step where the addition of cyanide ion takes place is not expected to be the slowest step [13]. Furthermore, in the rearrangement of **3** to **5** and of **14** to **16** the substituent is located very far from the reactive center to induce any effect. Thus, in order to explain the effect the substituent on the benzaldehyde has on the reaction mechanism, we have focused on the step that yields compound **10**. The rearrangement of the intermediate **5** to the O-benzoylated cyanohydrin intermediate **7** and the condensation of **7** with **8** to yield **10** are expected to determine the rate of the reaction. In the experimental study of Kuebrich, neither the formation nor the nucleophilic attack of **7** are cited as completely rate determining steps [16]. In order to

understand the mechanism of the condensation reaction and to assess the effect of the substituent on benzil and benzaldehyde on the reaction mechanism, the step involving the formation of **7** (structures **5**, **6**, **7**) and its nucleophilic attack (structures **7**, **8**, **9**, **10**) have been further considered with B3LYP (Figure 4.6 and Table 4.2).

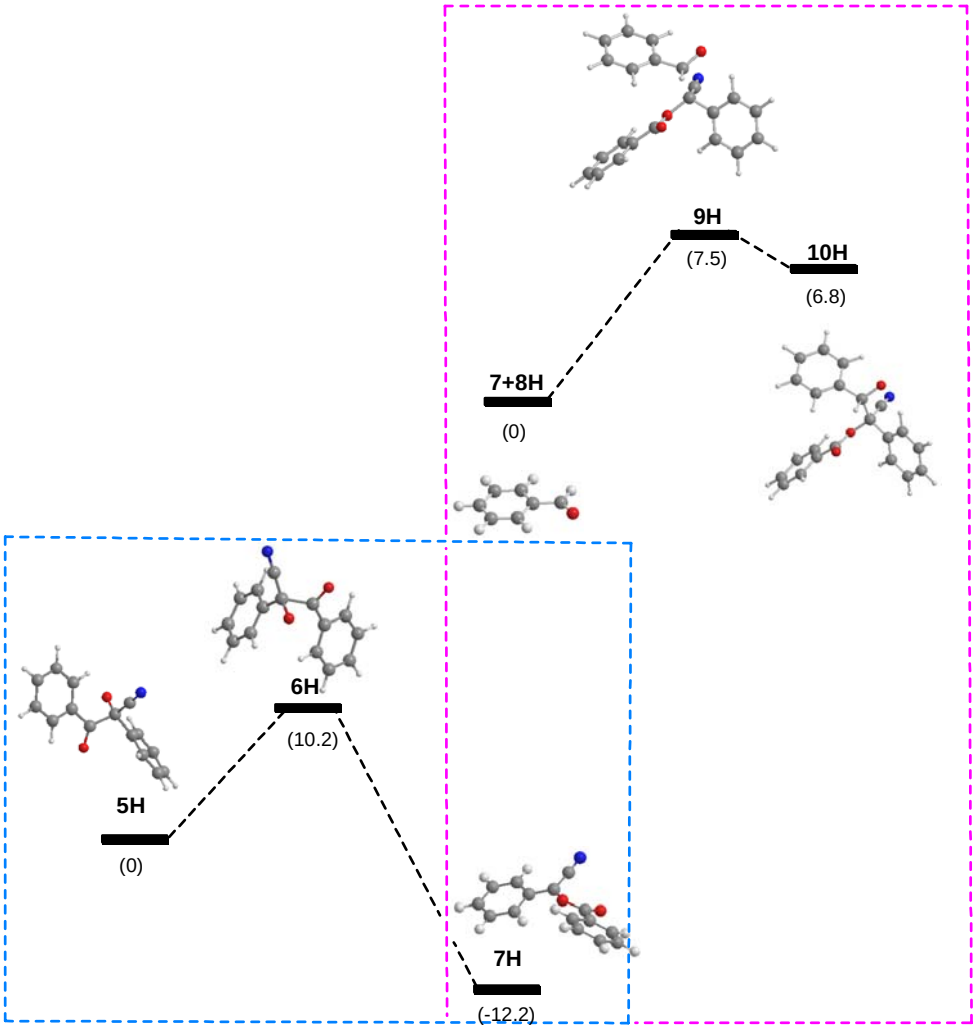
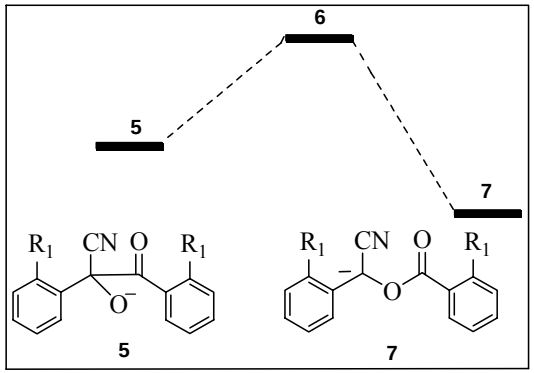


Figure 4.6. 3-D view of O-benzoylated cyanohydrin **7** formation and its condensation with benzaldehyde **8**.

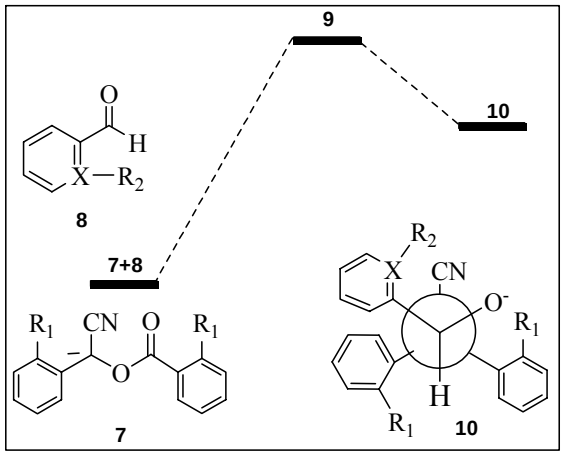
The energetics in the gas phase is considered first; the effects of solvation are discussed later. In the formation of O-benzoylated cyanohydrin intermediate **7**, the 1,2 C-O shift of benzoyl is an intramolecular rearrangement with 10.1 kcal/mol activation barrier (B3LYP/6-31+G*//B3LYP/6-31G*). The driving force for this rearrangement of compound **5** to **7** is the presence negative charge, which is resonance stabilized on compound **7**. The stability of O-benzoylated cyanohydrin intermediate **7** is well correlated with the exothermicity of this step (-12.6 kcal/mol). The negative charge on the compound **7** is expected to play a key role in the reaction. Any factor that stabilizes the negative charge on compound **7** or destabilizes the same charge on compound **5** is expected to increase the rate of the reaction at this step. The condensation of **7** with benzaldehyde **8** is an intermolecular reaction in which an ion-dipole type of interaction takes place. The activation barrier of this step is 7.5 kcal/mol and the step is endothermic by 6.8 kcal/mol. Any effect that increases the electrophilicity of compound **8** and the nucleophilicity of compound **7** are expected to increase the rate of this step.

Table 4.2. Relative energies (B3LYP) and thermodynamic contributions (kcal/mol) for the formation and the condensation of O-benzoylated cyanohydrin intermediate **7** with aldehyde **8** (T=298K). [e]



		Total Energy in Hartrees	**ΔE[a]**		**ΔE+ZPE[b]**		**ΔG[c]**		**ΔH[d]**	
X	**R_1**	**5**	**6**	**7**	**6**	**7**	**6**	**7**	**6**	**7**
C	H	-782.8192	9.7,10.1(3.5)	-12.6	9.3	-12.0	10.5	-12.1	8.8	-12.1
C	F	-981.2857	9.9, 9.0(10.3)	-8.2	9.4	-7.7	10.0	-9.0	9.0	-7.7
C	Cl	-1702.0062	8.3,8.5(8.3)	-8.6	7.8	-8.0	8.0	-9.4	7.4	-8.0

Table 4.3. Relative energies (B3LYP) and thermodynamic contributions (kcal/mol) for the formation and the condensation of O-benzoylated cyanohydrin intermediate **7** with aldehyde **8** (T=298K) (cont'd). [e]



			Total Energy in Hartrees	ΔE[a]		ΔE+ZPE[b]		ΔG[c]		ΔH[d]	
X	**R_1**	**R_2**	**7+8**	**9**	**10**	**9**	**10**	**9**	**10**	**9**	**10**
C	H	H	-1128.4126	7.5 ,9.7(11.9)	6.8	8.6	8.4	22.4	22.1	8.2	8.2
C	H	F	-1227.6444	5.8,8.6(10.7)	4.0	6.8	5.5	21.0	19.8	6.4	5.2
C	H	Me	-1167.7295	9.6 ,12.6(22.8)	9.0	10.7	10.6	24.2	24.5	10.3	10.4
N	H	H	-1144.5544	5.0,7.1(7.7)	3.7	6.0	5.3	19.9	19.3	5.6	5.0
C	F	Me	-2086.9101	10.7,13.0(22.2)	-	11.5	-	26.2	-	11.2	-
C	Cl	Me	-1366.1890	12.9,14.4(22.1)	9.7	13.5	11.3	28.2	25.9	13.2	11.1

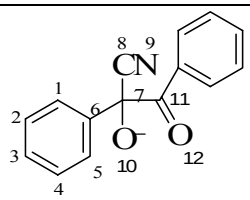
[a] Relative energy with respect to reactants (**5** or **7**+**8**).
[b] Relative energy with respect to reactants with zero point correction,
[c] Relative free energy with respect to reactants.
[d] Relative enthalpy of reaction with respect to reactants.
[e] B3LYP/6-31+G*//B3LYP/6-31G* energies are in bold and energetics in solution (IPCM) are in parenthesis.

The central carbon-carbon bond elongates from 1.570 Å in compound **5** to 2.415 Å in compound **7**, though it is 1.564 Å in the transition state **6** (Table 4.3). This shortening of the carbon-carbon bond in the transition state **6** can be explained by the charge delocalization over the atoms; the charge on O10 is –0.695 in **5**, it is –0.511 in **6** and –0.608 in **7**; the charge on O12 is –0.511 in **5**, -0.609 in **6** and –0.503 in **7** (Table 4.3) . At the end of this step, in compound **7**, the negative charge accumulates on the nitrogen atom of the cyanide group. The charge on C7 is 0.219 in **5**, 0.128 in **6** and 0.117 in **7**; that on C8 is 0.271 in **5**, 0.342 in **6** and 0.247 in **7**; the charge on N9 is -0.546 in 5, -0.545 in **6** and -0.598 in **7**. The length of the C11-O12 bond, which is 1.227 Å in **5**, 1.248 Å in **6** and 1.221 Å in **7**, can also confirm this delocalization- migration of electrons through nitrogen.

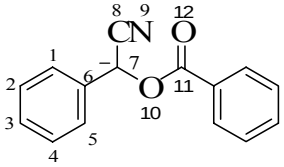
The negative charge on compound **7** that is responsible for its nucleophilicity is localized on the nitrogen atom (N9). As compounds **7** and **8** combine to yield compound **10** electrons on nitrogen atom (N9) migrate towards the atoms of the forming bond (C8: 0.247 in **7**, 0.364 in **9** and N9: -0.598 in **7** and –0.541 in **9**). An increase in the electron density of the nucleophilic carbon at the transition state (C7) is observed (0.117 in **7** and 0.114 in **9**). In compound **10** the charge is cumulated on the oxygen atom (O20). The most dramatic changes along the reaction path are on the acceptor aldehyde. Since the negative charge on compound **7** is ready to be donated for bond formation, the changes on compound 7 are not as pronounced as the ones on compound **8**.

The changes in the bond lengths also display a characteristic behavior. The C8-N9 bond length in compound **7** (1.178 Å) shortens in the transition state **9** (1.164 Å) and is the shortest in compound **10** (1.162 Å) indicating an increase in the bond order. As opposed to the decrease in the C8-N9 bond length, the C7-C8 bond length gradually increases as the reaction progresses (1.394Å in **7**, 1.455Å in **9** and 1.477Å in **10**). The forming C7-C19 single bond is 1.869 Å in the transition state and gains a single bond character in compound **10** with a length of 1.644 Å.

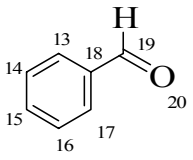
Table 4.4 Mulliken charge distribution and geometric parameters for unsubstituted molecules (B3LYP/6-31G*).



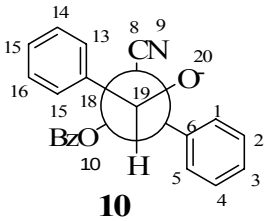
5



7



8



10

Atom	Charges					
	5H	**6H**	**7H**	**8 H**	**9 H**	**10 H**
C1	-	-	-0.221	-	-0.172	-0.159
C2	-	-	-0.132	-	0.11	0.114
C6	-	-	0.173	-	-0.144	-0.147
C7	0.219	0.128	0.117	-	0.114	0.167
C8	0.271	0.342	0.247	-	0.364	0.373
N9	-0.546	-0.545	-0.598	-	-0.541	-0.532
O10	-0.695	-0.608	-0.511	-	-	-
C11	0.388	0.441	0.578	-	-	-
O12	-0.511	-0.609	-0.503	-	-	-
O18	-	-	-	0.102	-0.63	-0.685
C16	-	-	-	-0.131	0.139	0.131
C13	-	-	-	0.194	-0.21	-0.147
C16	-	-	-	-0.119	-0.136	-0.136
C17	-	-	-	-0.132	-0.319	-0.165
	Distances					
C19-**C**20	-	-	-	1.216	1.284	1.322
C8-**N**9	-	-	1.178	-	1.164	1.164
C7-**C**8	-	-	1.394	-	1.456	1.455
C7-**C**19	-	-	-	-	1.869	1.644

Increasing the basis set by addition of polarization functions to 6-31G* is expected to alter the energies of the charged species. Compound **5** has been stabilized by 29.9 kcal/mol by including polarization functions; this stabilization is of 29.5 kcal/mol for the transition structure **6**. These two effects cancel each other as these two species resemble each other in terms of the negative charge. In the condensation of the cyanohydrin intermediate **7** with benzaldehyde **8**, the addition of polarization functions to B3LYP/6-31G*, stabilizes compound **8** by 9.6 kcal/mol, compound **7** by 27.4 kcal/mol and the transition state **9** by 34.8 kcal/mol. The overall effect of polarization functions is to increase the activation barrier by ~2 kcal/mol.

The gas phase dipole moments for the species **5**, **6**, **7**, **8**, **9**, **10** involved in the rate determining steps are 3.73, 6.19, 9.56, 1.31, 5.89 and 3.52 D respectively, at the B3LYP/6-31G* level. Since the experiment has been carried out in solution and the molecules of interest have high dipole moments, it is essential to consider the effect of solvation in the calculations. The net effect of solvation on the activation barrier of the cyanohydrin formation step is a decrease by 6.2 kcal/mol. The presence of negative charge causes, the same amount of stabilization on both compound **7** and the transition state **9** (~61 kcal/mol). Compound **8H** is stabilized by 4 kcal/mol, leading to an increase in the activation barrier.

4.3.2 Effect of Substituent on Benzaldehyde

Any substituent group on the aromatic rings is expected to change the electronic nature of the molecule and to affect the yield of the reaction. Experimentally, the electron withdrawing substituents at the ortho position of the acceptor aldehyde, such as CF_3, F, and Br have been found to increase the yield of the reaction and the electron-donating substituents such as OCH_3 and CH_3 have been found to decrease the yield (Table 4.1). On the other hand even though 2-pyridinecarboxaldehyde, which bears an electron poor aromatic ring, is expected to increase the yield, a decrease in the yield has been observed by Demir et.al. who have attributed this anomaly to a side reaction [80].

The condensation of O-benzoylated cyanohydrin intermediate **7** with **8** to yield compound **10** is expected to reflect the relationship between the substituent on benzaldehyde and the reaction yield. o-fluoro benzaldehyde has approximately the same activation barrier as 2-pyridinecarboxaldehyde and a lower activation barrier than o-methyl benzaldehyde. The endothermicity (ΔH_{rxn}) for the formation of compound **10** is similar for o-fluoro benzaldehyde and 2-pyridinecarboxaldehyde (~5.0 kcal/mol) and lower than the one for o-methyl benzaldehyde (10.4 kcal/mol). Among the energetically unfavored paths (ΔG is positive for all the aromatic aldehydes) o-fluoro benzaldehyde and 2-pyridinecarboxaldehyde are less endoergic than the others and their condensation is thermodynamically more favored. The fact that the experimental yield for o-fluoro benzaldehyde is higher than the one for o-methyl benzaldehyde has been attributed to the difference in barrier heights, with the known exception of 2-pyridinecarboxaldehyde. This molecule with its π-deficient aromatic pyridine ring has an activation barrier (5 kcal/mol) similar to the one in o-fluoro benzaldehyde, whereas its yield is lower than in the case of o-fluorobenzaldehyde.

The condensation of O-benzoylated cyanohydrin intermediate **7** with benzaldehyde **8** to yield compound **10** is endothermic for all the compounds studied and the transition states resemble the products obeying Hammond's postulate [81]. Hammond's postulate points out higher endothermicity correlated with high activation barrier. The path with low endothermicity will experience an earlier transition state. Among the transition states considered **9F** is the earliest transition state. It has the longest C7-C19 bond length, which justifies the lowest activation barrier. The transition state **9M**e has the shortest C7-C19 bond length suggesting a late transition state and a higher activation barrier (Figure 4.7). The C7-C19 bond length at the transition state **9** may be considered as a suitable indicator for the efficiency of reaction.

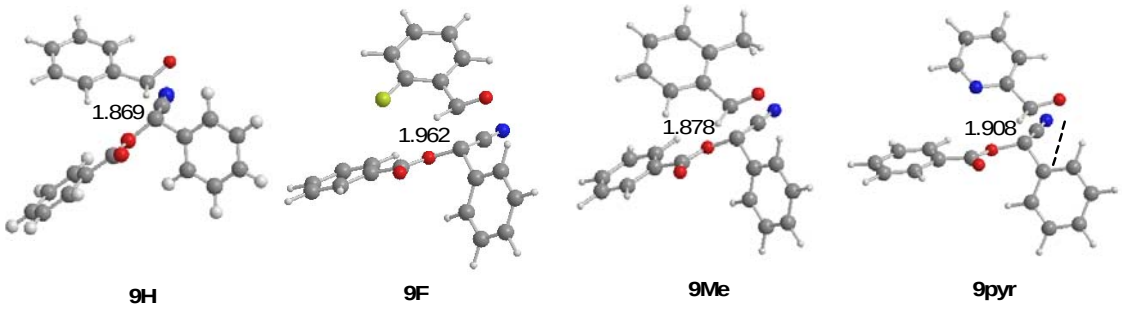


Figure 4.7. C7-C19 bond distances in the transition states **9H**, **9F**, **9Me**, **9pyr**.

Polarization functions are observed to increase the activation barrier of the condensation reaction of cyanohydrin **7** with aldehyde **8** for all the substituents by ~2.5 kcal/mol.

The solvent has increased the activation barriers for the condensation reaction of benzaldehyde **8** and o-fluoro benzaldehyde **8**F with O-benzoylated cyanohydrin intermediate **7** by approximately the same amount (~4.5 kcal/mol). The increase in the activation barrier arises from the shielding of the negative charge by the solvent dielectric, which inhibits its attack towards the electrophilic center. In the case of o-methyl benzaldehyde **8Me**, the destabilizing effect of the solvent on the activation barrier is much more pronounced (13.2 kcal\mol).

4.3.3. Effect of the Substituent on Benzil

In this part of the study, in order to understand the effect of the substituents on benzil, two benzil derivatives: 1,2-bis(2-chlorophenyl)ethane-1,2-dione and 1,2-bis(2-fluorophenyl)ethane-1,2-dione have been considered and both the formation of the corresponding O-benzoylated cyanohydrin intermediates (**7**) and their condensation with benzaldehyde **8Me** have been considered. Experimentally, o-methyl benzaldehyde has been chosen as the acceptor aldehyde in the condensation of substituted benzil molecules. Among the compounds chosen, 1,2-di-2-fluorophenylethane-1,2-dione has low reaction yield than 1,2-

di-2-chlorophenyl ethane-1,2-dione. The activation barrier for the rearrangement of **5** to **7** in the former molecule has been calculated to be 2 kcal/mol higher than the activation barrier for the later (B3LYP/6-31+G*//B3LYP/6-31G*) which accounts for the difference in the reactivities. However, the barrier for condensation reaction with **8Me** is higher by 1.4 kcal/mol in the case of 1,2-di-2-chlorophenylethane-1,2-dione as compared to 1,2-di-2-fluorophenylethane-1,2-dione (B3LYP/6-31+G*//B3LYP/6-31G*). Based on gas phase B3LYP/6-31G* calculations, the experimental behavior is only reproduced in the formation of cyanohydrin intermediate **7**. Since the negatively charged cyanohydrin intermediate is modified by substitution on benzil, solvation is expected to have a considerable effect on the energetics of the molecules concerned. The formation of the cyanohydrin intermediate from 1,2-di-2-fluorophenylethane-1,2-dione has been observed to have 2 kcal/mol higher activation barrier than 1,2-di-2-chlorophenylethane-1,2-dione in solution. This difference in the activation barriers justifies the lower reaction yield with 1,2-di-2-fluorophenylethane-1,2-dione. On the other hand, the activation barriers for the condensation of the cyanohydrin intermediates with **8Me** are more or less similar for both 1,2-di-2-fluorophenylethane-1,2-dione and 1,2-di-2-chlorophenylethane-1,2-dione in solution (~22 kcal/mol).

4.3.4. Chemical Reactivity Indices

In this section, chemical reactivity indices have been employed to investigate the most susceptible sites for nucleophilic/electrophilic attack and the reactivity variance among the acceptor aldehydes (Figure 4.8) using equations 3.49, 3.50, 3.51 and 3.52 Fukui functions for the nucleophilic attack, local and global softness values describing the local reactivity are presented in Table 4.4. When the HSAB principle is employed, global softness values can be used to verify the results derived from the potential energy surface analysis in the sense that interactions are favored, when the global softness values of the two interacting species are close to each other.

The global softness of the o-fluoro benzaldehyde (**8F**) is 5.369: this value is closest to the global softness value of compound **7** (S=53.476) among the aldehydes concerned. For the

acceptor aldehyde, the atom with the highest positive Fukui function and local softness will be the most susceptible site for the nucleophilic attack. For all the acceptor aldehydes this site is the C_{19} carbonyl carbon. It is interesting to note that; o-fluorobenzaldehyde (**8F**) has higher Fukui function (0.246) and local softness (1.321) than the ones (0.224, 1.190) for o-methylbenzaldehyde (**8Me**). This describes the high susceptibility of o-fluoro benzaldehyde towards O-benzoylated cyanohydrin in agreement with experimental results and the energy barriers.

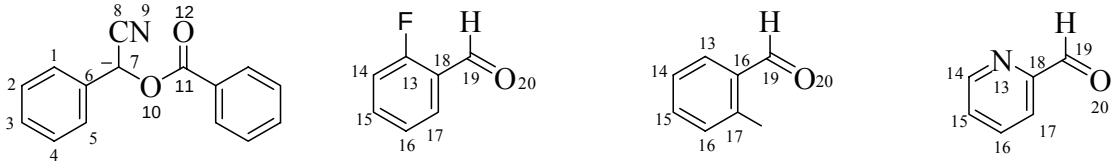


Figure 4.8. Atom numbers for the substituted aldehyde molecules and benzil molecule.

Table 4.5. Chemical Reactivity Indices (B3LYP/6-31+G*)

Atom	f^+_{MPA}	s^+_{MPA}	f^+_{MPA}	s^+_{MPA}		
	7di-Cl (S=28.317)		**7di-F(S=29.451)**			
7	0.592	16.764	0.422	11.94		
8	-0.803	-22.739	0.004	0.11		
9	0.226	6.400	0.179	5.06		
10	0.065	1.841	0.012	0.34		
11	-0.002	-0.057	0.239	6.67		
12	0.035	0.991	0.034	0.96		
	f^+_{MPA}	s^+_{MPA}	f^+_{MPA}	s^+_{MPA}	f^+_{MPA}	s^+_{MPA}
	8F(S=5.369)		**8Me(S=5.313)**		**8pyr(S=5.502)**	
13	0.038	0.204	0.187	0.994	0.084	0.462
17	0.120	0.644	0.234	1.243	0.103	0.567
18	0.023	0.124	0.008	0.043	0.086	0.473
19	0.246	1.321	0.224	1.190	0.255	1.403
20	0.163	0.875	0.144	0.765	0.160	0.880

The global softness of 1,2-bis(2-chlorophenyl)ethane-1,2-dione (**7di-Cl**) is 28.316 and that of 1,2-bis(2-fluorophenyl)ethane-1,2-dione (**7di-F**) is 29.451, whereas the global softness of **8Me** is 5.313 (Table 4.4). Therefore, 1,2-bis(2-chlorophenyl)ethane-1,2-dione (**7di-Cl**) with the lowest global softness value is expected to be more susceptible for a nucleophilic attack to **8Me**. Both of the benzil derivatives have the highest Fukui function and local softness at the negatively charged carbon atom as expected.

4.4. Conclusion

The mechanism and substituent effect for the synthesis of O-benzoylated benzoin from benzil have been investigated computationally with PM3 and B3LYP. The O-benzoylated cyanohydrin functions as the cyanohydrin in the conventional benzoin condensation. The O-benzoylated cyanohydrin behaves as a nucleophile and attacks benzaldehyde, where the product is the ester of the corresponding unsymmetrical benzoin. The mechanism of the condensation reaction is consistent with the mechanism suggested by Lapworth and is presented in details in this study.

Two steps are proposed as being effective on the rate of the reaction: the cyanohydrin formation and its condensation with aldehyde. The experimental reactivity difference between the substituted benzil molecules has been justified based on the difference in the activation barriers of the O-benzoylated cyanohydrin formation step (**5**→**7**) in solution, since the energetics of the condensation reaction are similar for both. Thus, the reactivity difference between 1,2-di-2-chlorophenylethane-1,2-dione and 1,2-bis(2-fluorophenyl)ethane-1,2-dione has been rationalized by comparison of the energy barriers along the formation of O-benzoylated cyanohydrin in solution. On the other hand, the activation barriers for the step yielding compound **10**, reproduce qualitatively the experimental yields of the various benzaldehyde derivatives subjected to condensation with O-benzoylated cyanohydrin. The nucleophilic attack of the O-benzoylated cyanohydrin to the electrophilic benzaldehyde is favored by the o-fluoro substituent, which withdraws electrons from the electrophilic center.

However for the o-methyl substituent, the path is comparatively less favored due to the weak electron donation. The effects of the substituents at the ortho position of the aldehyde are explained in terms of the electronic nature of the substituent. The discrepancy between the experimental and theoretical results for 2-pyridinecarboxaldehyde has been attributed to side reactions. In the synthesis of O-benzoylated unsymmetrical benzoins from benzil the reactivity of novel ortho-substituted aromatic aldehydes can conveniently be predicted by considering the heights of their activation barriers along the condensation step. In accordance with Kuebrich, we may state that neither the formation nor the condensation of the O-benzoylated cyanohydrin are rate-determining steps by themselves, both play a role on the rate of the reaction [16].

5. UNDERSTANDING THE MODE OF ACTION OF ThDP IN BENZOYLFORMATE DECARBOXYLASE (BFD)

5.1. Introduction

Molecular dynamics and quantum mechanics have been employed to explore the mechanism of all elementary steps involved in the catalytic cycle of BFD to produce an acyloin. Models involving different charge states of amino acids and/or mutants of critical residues were constructed in order to understand the involvement of the catalytically active residues and the enantioselectivity in this reaction. A model was the representative unit of BFD to yield an acyloin is composed of two protein chains, two ThDPs, three Mg (II) cations and all ions and water molecules within the interaction distance of these molecular entities (Figure 5.1). In such a model, one protein chain is placed near the other in hand-shake interaction. ThDPs are located at the two ends of the space between the two proteins. Being at the two ends, ThDPs will benefit from being exposed to any incoming substrate and other molecules. ThDPs are embedded into the enzyme from the pyrophosphate tail via two coordination covalent bonds between phosphate oxygens and Mg (II) cation. Mg (II) is also bonded to the surrounding residues occupying the other four locations of the octahedral geometry.

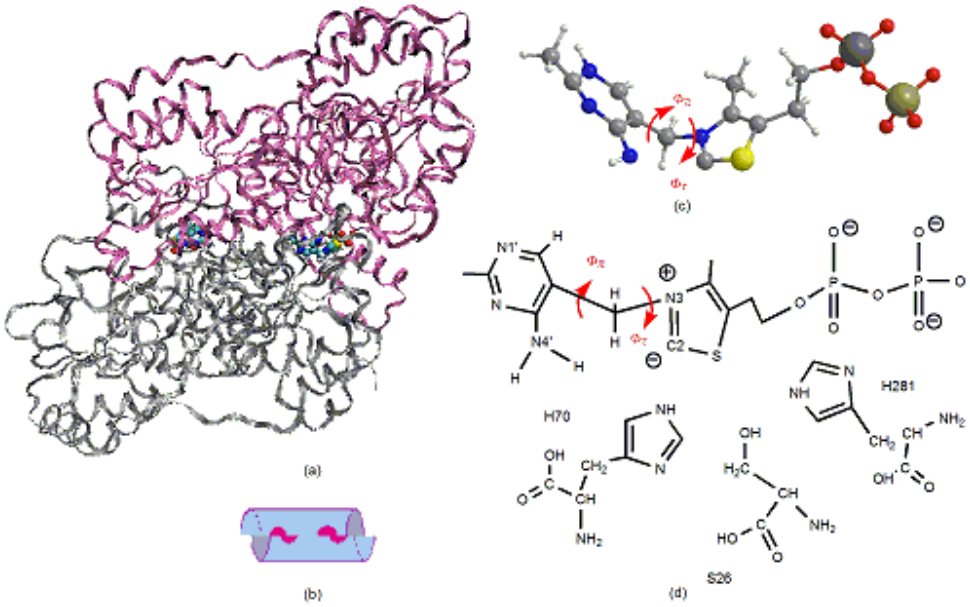


Figure 5.1. a) BFD representative unit, b) ThDPs at chain interface and c) 3D structure of ThDP, d) catalytically active amino acid residues around ThDP.

Several studies point out the importance of the active site residues and the structural features of ThDP in its catalytic function. Mutation studies assist the mechanistic studies by identifying the catalytically active amino acid residues. The residues that play a role in the catalysis by BFD are H70, S26, and H281. Replacing H70 and H281 with alanine essentially deactivates the enzyme [33, 35]. All ThDP-dependent enzymes appear to contain either a polar residue (arginine or glutamine), or a histidine in this position [82, 83]. Jordan et al. point out the importance of a histidine residue at this position in substrate binding or in the course of the reaction [84]. Another study conducted by Kenyon et al. identifies the active site residues in mandelate racemase, a ThDP dependent enzyme, and points out the evidence of the same catalytic action of a histidine residue [34]. Mutation of S26 to alanine results in a decreased enzymatic activity [35].

In this study, the ThDP catalyzed acyloin condensation in BFD environment is modeled. The protonation states of the catalytically active residues and the enantioselectivity of the intermediates are investigated, and the proton transfer mechanisms in the catalytic path are examined. Finally, energy barriers derived from quantum mechanical calculations were utilized to explain the reactivity differences for benzaldehyde and acetaldehyde along the second nucleophilic attack in the reaction. This part has recently been submitted for publication in *Biochemistry*.

5.2. Computational Methodology

5.2.1. Simulation Details for MD

All simulations were performed using NAMD version 2.6 [85] with CHARMM27 all-atom force field parameters [86, 87]. Coulombic terms were cutoff at 12 Å by the use of a switching function at 10 Å. The nonbonded atom list was updated every 10 steps for interactions within 13.5 Å of each atom. The SHAKE [70] and RATTLE [71] algorithms were used to constrain all bonds in the system, allowing for a time step of 1 fs. Long range electrostatics were handled using the Particle Mesh Ewald (PME) sum algorithm [72]. Possible high energy contacts between atoms were removed using 2500 steps of steepest descents minimization of the potential energy of each system. The structures produced by minimization were then subjected to an equilibration run of 0.5 ns. Finally, data production runs were carried out for 5 ns. The equations of motion were integrated by using the Verlet algorithm [88] and were treated as interacting Langevin atoms to control the temperature which was set to 310 K. Throughout the simulations the fluctuations in temperature were monitored and were found to be less than 5 K. The analyses of the results were performed using VMD [89].

5.2.2. Possible Reaction Mechanisms

The proposed reaction mechanism, which serves as a model for the formation of *S*-HPP from benzaldehyde and acetaldehyde is depicted in Figure 5.2. Different from the mechanism presented in Figure 1.6, the mechanism that we propose includes the reactive intermediates that function in proton transfer. An oxyanion, **VII**, for example, is introduced to the path representing the proton acceptor that demands the proton from **HB1**. In the first nucleophilic attack, two successive proton migrations catalyzed by proton donor, **HB1**, and proton acceptor, **B2** take place.

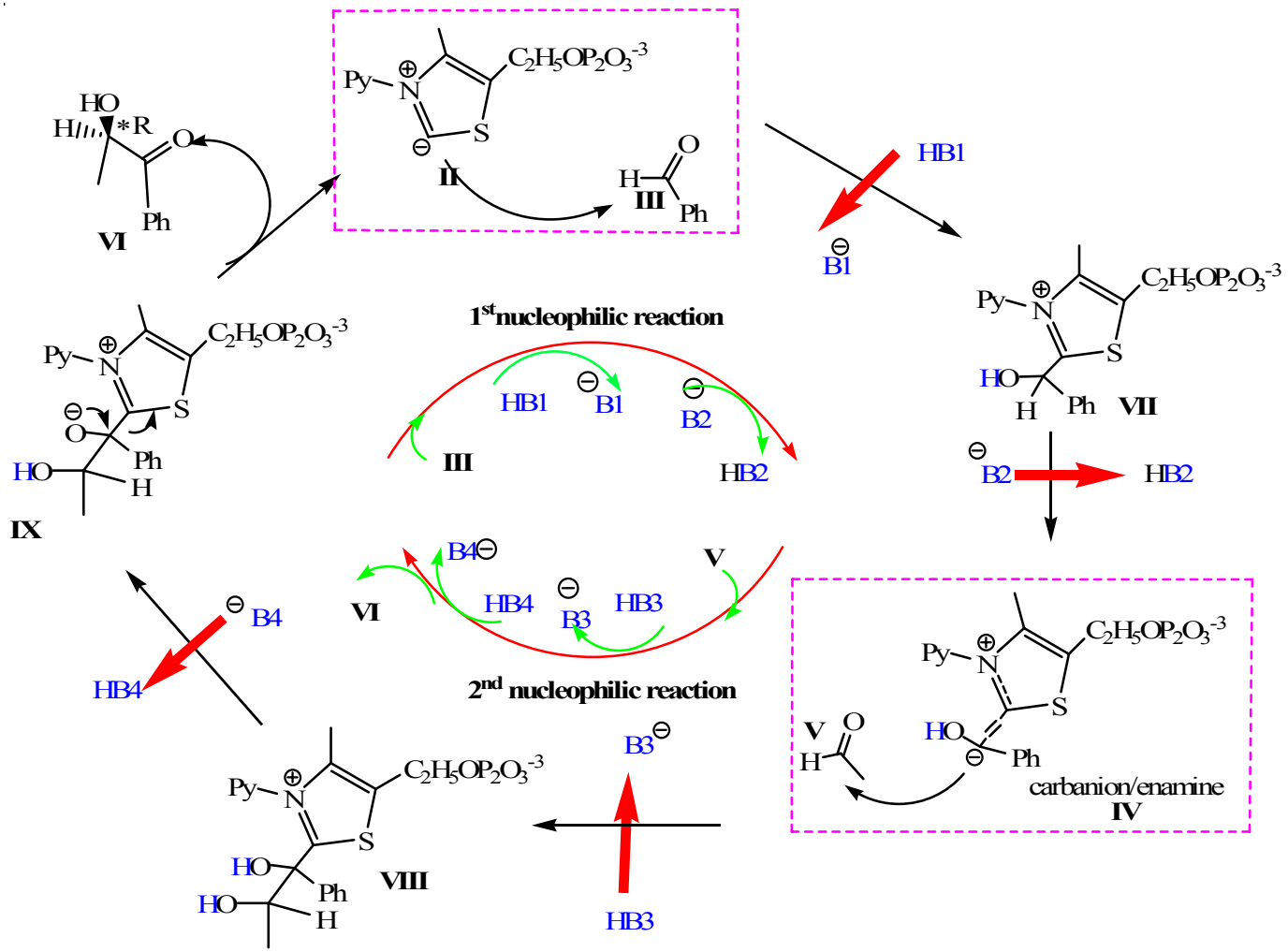


Figure 5.2. (*S*)-HPP (VI) condensation in BFD environment. Pre and postreactive interactions for the first and second nucleophilic half reactions. Reactants within the dashed rectangles represent the two basic nucleophilic reactions.

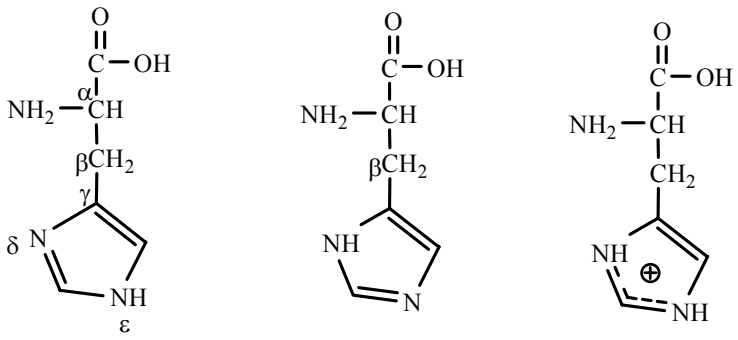
There are two types of proton transfers in the first nucleophilic attack. The first one is the protonation of O_{donor} and the second is the removal of the donor aldehyde proton (H_{donor}). At the end of the reaction, the protons should be donated back to their original owner or to the second substrate (acceptor aldehyde), for the reaction to be carried out with another pair of aldehydes. One more proton transfer takes place at the second stage to protonate $O_{acceptor}$ and in benzoin condensation, O_{donor} conserves its carbonyl character at the end of the reaction. Therefore, it is highly probable that **HB1** and **B4**, and **B2** and **HB3** are conjugate acid base pairs. At the active center, only H70, H281 and N4′ may provide necessary proton transfer. Thus, three possible mechanisms can be derived:

i) H70 as **HB1/B4** and H281 as **B2/HB3**,

ii) H281 as **HB1/B4** and H70 as **B2/HB3**,

iii) ThDP as HB1/B4 and H281 as B2/HB3.

HB1 and **HB3** represent the protonated forms and **B2** and **B4** the deprotonated forms of the bases. When it acts as **HB** type of a proton donor, a histidine residue is protonated and when it is a **B** type acceptor it is in its neutral form (Scheme 5.1). If the histidine is protonated, it is repesented by HSP where both nitrogen atoms are protonated. If only one proton is available, it can be bonded to nitrogen atom at the δ position of the histidine, it is represented by HSD, or if it is at the ε position, it is represented by HSE. The three different histidine residue types are presented in Scheme 5.1.



Scheme 5.1. Histidine protonation states.

According to the first proton donation mechanism, ***i***, **VII** is protonated by H70 and aldehyde the proton is removed by H281. In the second proton donation mechanism ***ii***, **VII** is protonated by H281 and the aldehyde proton is removed by H70. The third scenario depicts the protonation of oxygen atom in **VII** via N4′ and aldehyde proton removal via H281. In this case N4′ donates one of its protons to the aldehyde oxygen. Self proton donation and removal mechanism as in the case of conventional benzoin condensation is also possible for this reaction [75]. However, due to the proton rich and labile nature of the enzyme, this is less plausible.

5.2.3. Model Systems

We have designed ten model systems defining different local interactions in the active site, to discern which of the three scenarios outlined above is more likely. Each of the model systems has different protonation states of H70, H281 and N4′ in the presence of different substrates as outlined in Tables 5.1 and 5.2. ***Models 1*** and ***2*** represent the holoenzyme (catalytically active enzyme with two cofactors). In ***Model 1***, protonation mechanism ***i*** is mimicked with H70 in protonated (**HB1**) and H281 in deprotonated states (**B2**). Conversely, protonation mechanism ***ii*** is pursued in ***Model 2*** with H70 and H281 as modeled in deprotonated and protonated states, respectively. ***Model 1*** and ***2*** form the basis for the other

models. They are built to check the consistency of the literature data with the data produced through the model structures with simulations [35]. ***Model 1*** and ***2*** contradict each other. They both have ThDP at the active center without the donor aldehyde. ***Models 3*** and ***4*** are prepared to investigate the attack of ThDP, **II** to benzaldehyde, **III** so as to analyze the active center interactions in the first nucleophilic attack in scenarios ***i*** and ***ii***, respectively. These models were prepared as a continuation of the first two models with benzaldehyde placed near the active center. Interactions with N4′ and protonation mechanism, ***iii*** can also be investigated with these two models. In ***Model 4***, H70 is deprotonated; therefore, if N4′ is the proton donor **HB1**, it would interact more closely with the aldehyde oxygen. Other intermediate structures in Figure 3 are inevitably present on the reaction path, although they are not modeled. The enzyme environment is proton rich and labile; thus, any proton demand is served immediately, giving these intermediates short life times in the course of the reaction.

Table 5.1. Model systems generated for MD simulation.

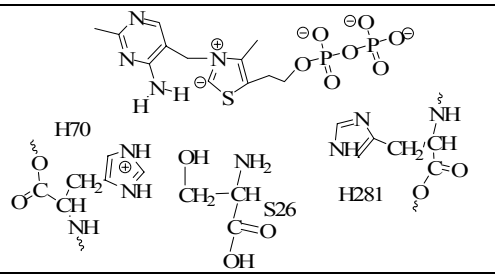
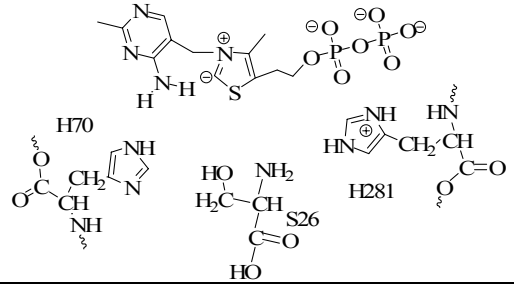
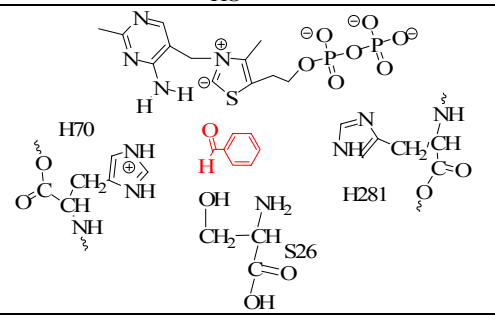
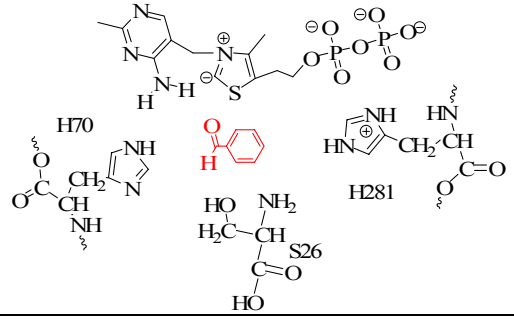
	Tested Scenario and Protonation States
Model 1	
Model 2	
Model 3	
Model 4	

Table 5.2. Model systems generated for MD simulation (cont'd).

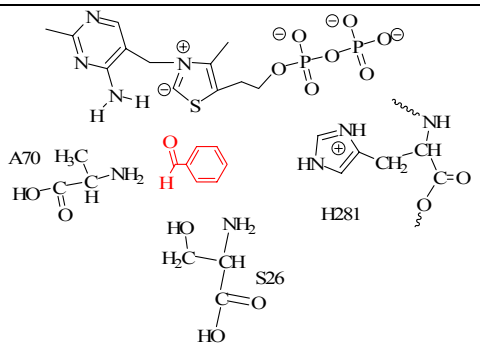
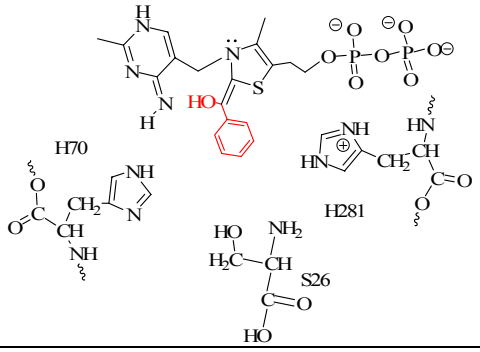
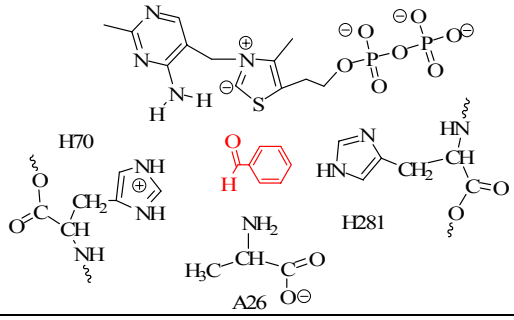
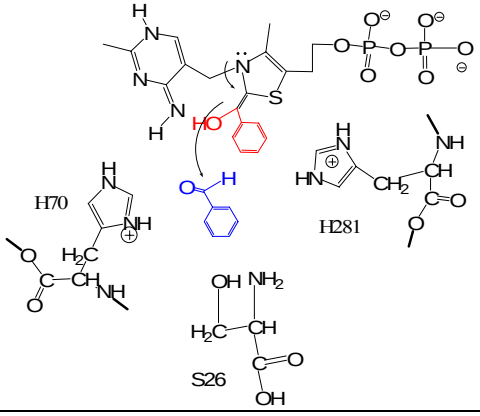
Model 5	
Model 6	
Model 7	
Model 8	

Table 5.3. Model systems generated for MD simulation (cont'd).

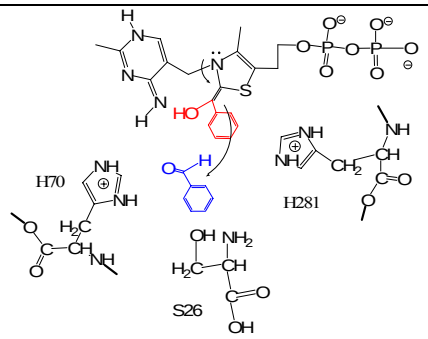
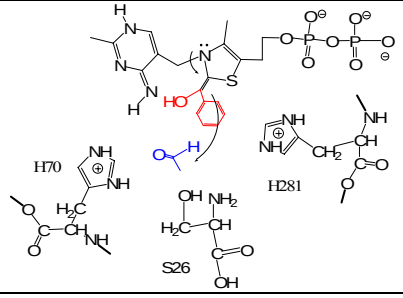
Model 9	
Model 10	

Table 5.4. Tested protonation scenarios and corresponding model systems in the study.

	Scenario and Protonation State	System
Model 1	H70 or N4′ as **HB1/B4** and H281 as **B2/HB3**	Enzyme and ThDP
Model 2	H281 as **HB1/B4** and H70 as **B2/HB3**	″
Model 3	H70 or N4′ as **HB1/B4** and H281 as **B2/HB3**	Enzyme, ThDP and BA
Model 4	H281 as **HB1/B4** and H70 as **B2/HB3**	″
Model 5	H70A mutation of ***Model 3*** and H281 as **B2/HB3**	″
Model 6	H70 or N4′ as **HB1/B4** and H281 as **B2/HB3**	″
Model 7	S26A mutation of ***Model 3*** and H281 as **B2/HB3**	″
Model 8	H70 or N4′ as **HB1/B4** and H281 as **B2/HB3**	Enzyme, enamine/carbanion (*re* face) and BA (*re* face)
Model 9	″	Enzyme, enamine/carbanion (*si* face) and BA (*si* face)
Model 10	″	Enzyme, enamine/carbanion (*si* face) and AA (*si* face)

With the aim of identifying the proton donors and acceptors we have prepared another model by mutating H70 to alanine. ***Model 5*** is a mutated version of ***Model 3***. In this model, benzaldehyde, **III** was docked into the active center and H281 is modeled as deprotonated.

Model 6 depicts the interaction of resonance stabilized enamine/carbanion intermediate, **IV** with the enzyme environment to represent the continuation of the first nucleophilic attack. ***Model 6*** represents the post reaction interactions after the first half reaction whereby O_{donor} atom is protonated by N4′ and the H_{donor} is removed by H281. Enamine replaces ylide-ThDP at the active site, N4′ is deprotonated and H70 is protonated. In the light of experimental predictions, we have prepared one more model, investigating the effect of S26A mutation on the first nucleophilic attack. ***Model 7*** is the S26A mutated version of ***Model 3***.

The second nucleophilic attack was elucidated both with QM and MD tools. In QM models, the enamine/carbanion, **IV**, benzaldehyde or acetaldehyde **V**, were modeled as ground state reactants. A transition state is located along the reaction path. The first MD model for this part of reaction is ***Model 8*** and represents the nucleophilic attack of enamine/carbanion, **IV**, to benzaldehyde. In this model, H70 and H281 are protonated, whereas N4′ is modeled as the second proton donor **B3** and H281 is modeled as the second proton acceptor, **HB4**. The product of this step has a chiral center. Therefore, we have prepared another model to represent the other enantiomer. We have prepared ***Model 9*** with similar protonation states but attacking from the outer face of the V shaped ThDP to produce the other enantiomer. ***Model 10*** is similar to ***Model 9***, with benzaldehyde replaced by acetaldehyde.

5.2.4. System construction for MD

All systems in this study were modeled based on the reported crystal structure (PDB code 1MCZ) [35]. Three of the co-crystallized dimers were removed from the original PDB file with all surrounding water molecules, ions and cofactors to mimic the minimal functional unit of BFD at different stages of acyloin condensation reaction (see Figure 5.1). A model system in this study contains 1050 amino acid residues (for the dimeric protein structure),

5899 water molecules, and various substrates whose effects were studied. Thus, each system consists of at least 33486 atoms.

After the addition of the substrates, the system is soaked in an 8 Å thick layer of water. We have compared the results from runs under periodic boundary conditions (PBC) at constant pressure, with those where the protein is immersed in a 8 Å thick water layer. The model system with PBC was prepared and immersed into a water box where there is a water molecule within at least 10 Å of the farthest atom from the center of mass of the model system. The characteristic properties of the two simulations are compared in Figures 5.3-5.6. The distance between C2 and one of the N4′ hydrogens, the dihedral angle values, RMSD values and positional fluctuations are in agreement in ***Model 1*** and ***Model 1*** with PBC, as presented in Figures 5.3 to 5.6, respectively (the latter model is depicted as ***Model 1*** (PBC)).

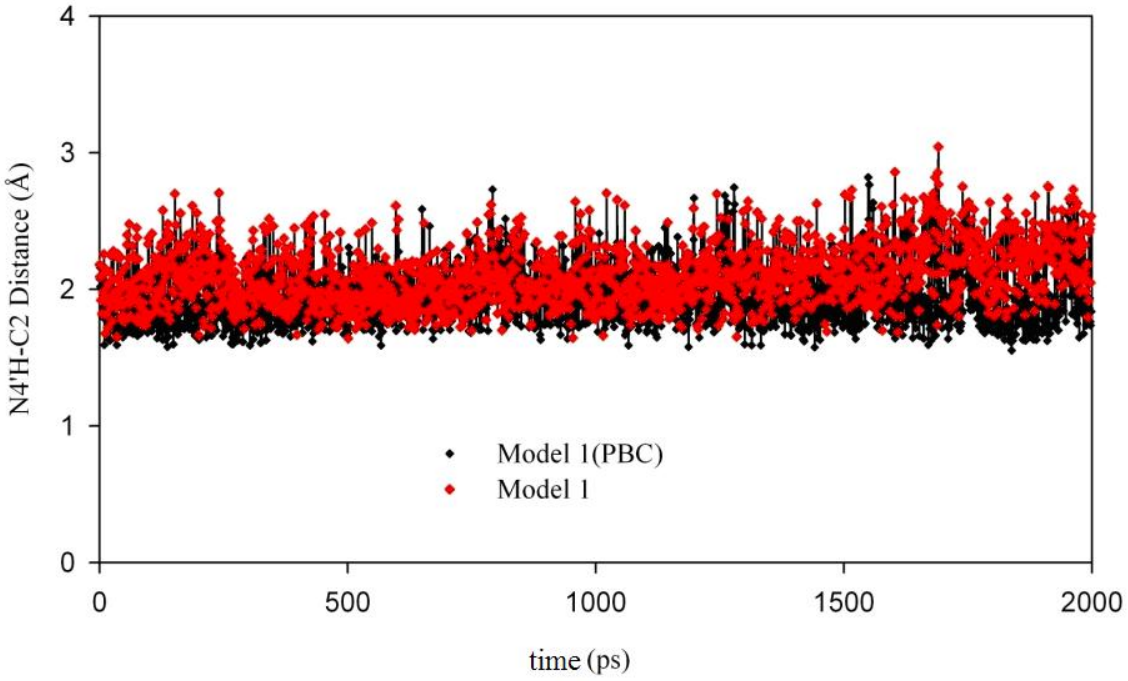


Figure 5.3. N4′ H-C2 distance in Ylide ThDP with *Model 1* solvated by 8 Å thick water layer and solvated by water box (PBC).

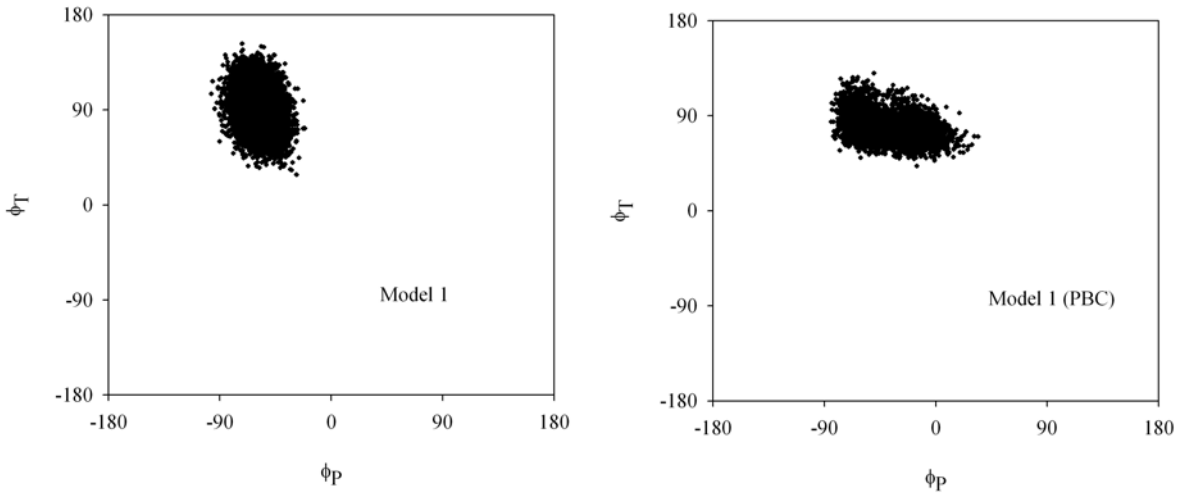


Figure 5.4. Dihedrals ϕ_P vs ϕ_T for ***Model 1*** vs. ***Model 1*** (PBC).

The positional fluctuations of C_α atoms in the systems are similar to each other (Figure 5.5).

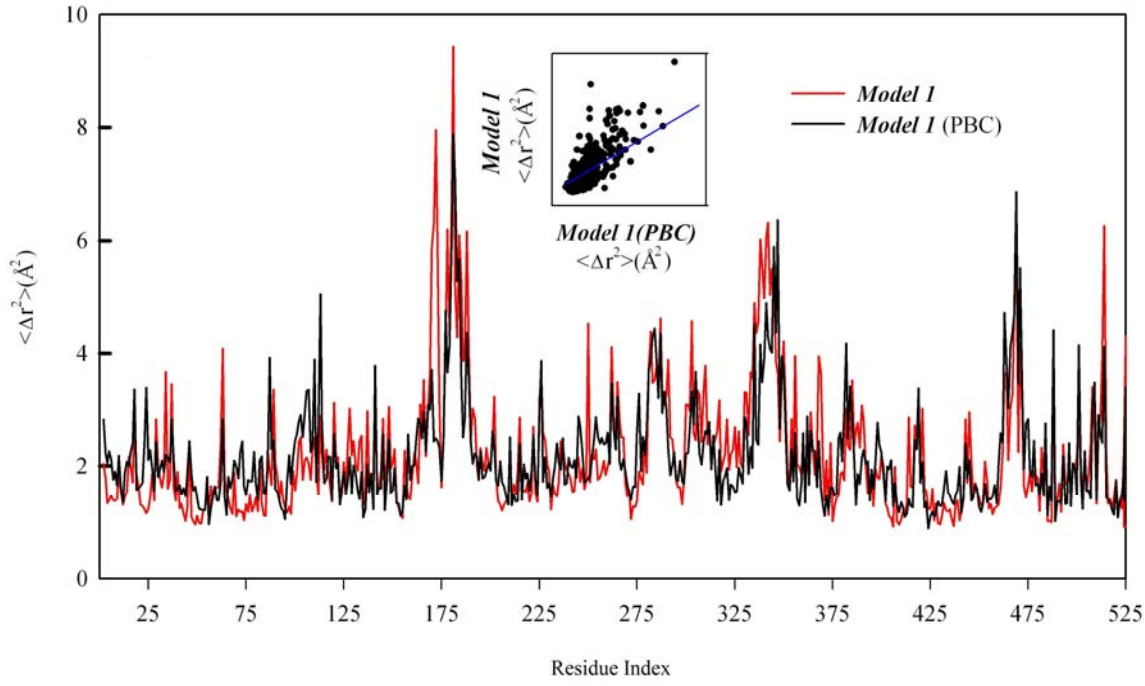


Figure 5.5. Positional fluctutation of the C_α atoms in the reaction zone for the simulation, in ***Model 1*** and ***Model 1***(PBC). The two sets of data are plotted against each other in the inset; the $y = x$ line is shown in blue to guide the eye.

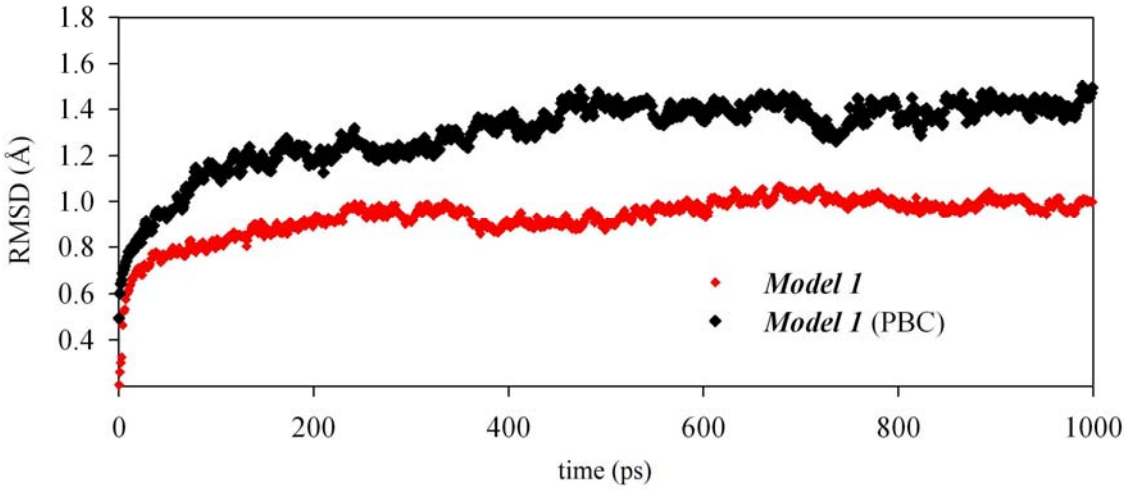


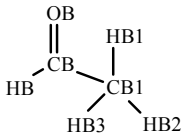
Figure 5.6. RMSD layer vs. PBC for *Model 1*.

We do not use periodic boundary conditions (PBC) considering the number of simulations performed, and the fact that we extensively study the local interactions in the active site, which is at least 24 Å away from the surface in order to save computer time.

5.2.5. Parameterization of MD Simulations

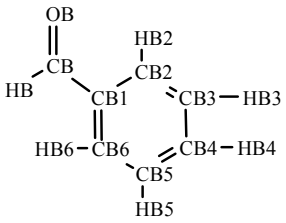
Ylide ThDP, **II**, and related aldehydes were docked into the active centers using Spartan [76] and VMD version 1.8.4 [89]. Initial structures for the corresponding (MD) nucleophilic attack models were optimized with B3LYP method and 6-31+G* basis set. Atomic properties and force field parameters of the ylide ThDP **II**, benzaldehyde, **III** and acceptor aldehyde (acetaldehyde or benzaldehyde), **V** were taken from the CHARMM27 all-atom force field parameters for the regular amino acids [86, 87]. MacKerrell *et al*. presented atom properties and charges (topology files) of acetaldehyde and benzaldehyde (Table 5.3 and 5.4 respectively). The atom names, types, and charges are listed below, according to the CHARMM format.

Table 5.5. Acetaldehyde atom names utilized in *Model 10*.



Atom Name	Atom Type	Charge	Atom Name	Atom Type	Charge
ATOM HB	HR1	0.06	ATOM HB1	HA	0.09
ATOM CB	CD	0.12	ATOM HB2	HA	0.09
ATOM OB	O	-0.32	ATOM HB3	HA	0.09
ATOM CB	CT3	-0.13			

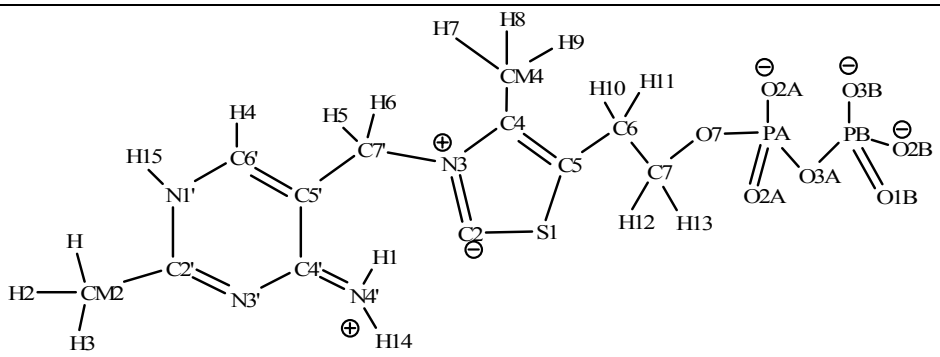
Table 5.6. Benzaldehyde atom names utilized in *Models 3, 4, 5, 7,8* and *9*.



Atom Name	Atom Type	Charge	Atom Name	Atom Type	Charge
ATOM HB	HR1	0.05	ATOM HB3	HP	0.12
ATOM CB	CD	0.16	ATOM CB4	CA	-0.12
ATOM OB	O	-0.33	ATOM HB4	HP	0.12
ATOM CB1	CA	0.12	ATOM CB5	CA	-0.12
ATOM CB2	CA	-0.12	ATOM HB5	HP	0.12
ATOM HB2	HP	0.12	ATOM CB6	CA	-0.12
ATOM CB3	CA	-0.12	ATOM HB6	HP	0.12

Schiott *et al.* have studied different conformers of ThDP and presented the atom types and related information in their study [90]. Those of ylide-ThDP are as shown in Table 5.5.

Table 5.7. Ylide ThDP atom names utilized in ***Models 1, 2,3,4,5*** and ***7***.



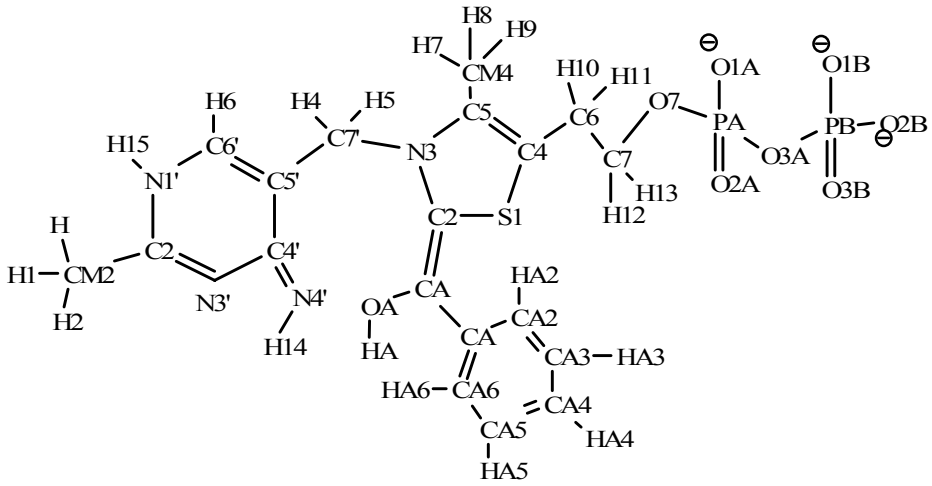
Atom Name	Atom Type	Charge	Atom Name	Atom Type	Charge
GROUP			ATOM CM4	CT3	-0.29
ATOM N1'	NN2	-0.45	ATOM C6	CT2	-0.21
ATOM C2'	CA	0.65	ATOM H5	HA	0.15
ATOM CM2	CN9	-0.54	ATOM H6	HA	0.15
ATOM N3'	N6R	-0.65	ATOM H	HA	0.09
ATOM C4'	CA	0.45	ATOM H8	HA	0.09
ATOM N4'	NC2	-0.58	ATOM H9	HA	0.09
ATOM C5'	CA	0.07	ATOM H10	HA	0.15
ATOM C6'	CA	-0.12	ATOM H11	HA	0.15
ATOM H	HN9	0.17	GROUP		
ATOM H2	HN9	0.17	ATOM C7	CN8	0.18
ATOM H3	HN9	0.17	ATOM H12	HN8	0.04
ATOM H4	HN3	0.17	ATOM H13	HN8	0.04
ATOM H15	HN2	0.36	GROUP		
ATOM H14	HC	0.37	ATOM O7	ON2	-0.54
ATOM H1	HC	0.25	ATOM PA	P	1.48
GROUP			ATOM O1A	ON3	-0.92
ATOM C7'	CT2	-0.42	ATOM O2A	ON3	-0.92
ATOM N3	NR3	0.58	ATOM O3A	ON2	-0.71
ATOM C2	CPH2	-0.69	ATOM PB	P2	1.60

Table 5.5. Ylide ThDP atom names utilized in ***Models 1, 2,3,4,5*** and ***7*** (cont'd).

ATOM S1	S5R	0.19	ATOM O1B	ON3	-0.91
ATOM C5	CPH1	-0.04	ATOM O2B	ON3	-0.91
ATOM C4	CPH1	-0.08	ATOM O3B	ON3	-0.91

VCharge [91] predictions for the enamine/carbanion intermediate generated for the model systems in this study are shown in the Table 5.6. In CHARMM format the molecular entities should be introduced as a group, the charge of the group as a whole should be an integer. This molecule is treated as a whole in the calculations, not separated into subgroups. This is a limitation arising from the calculations of VCharge.

Table 5.8. Enamine/carbanion atom names utilized in ***Model 5,8,9*** and ***10.***



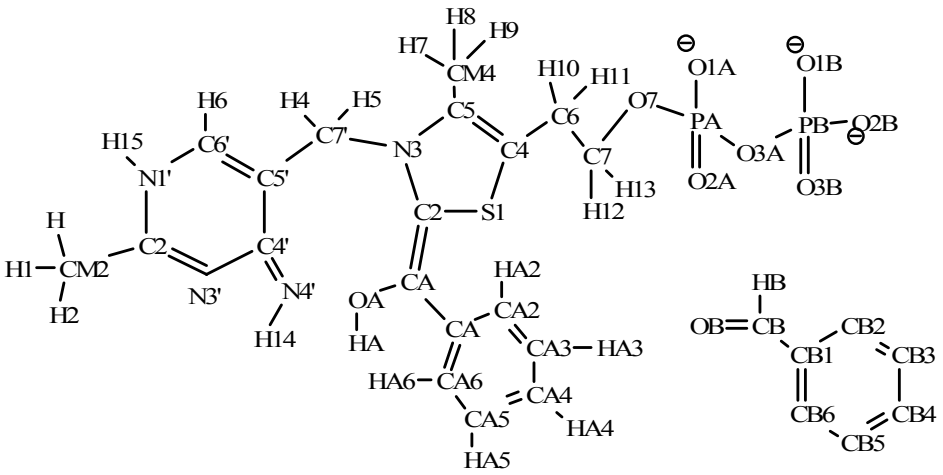
Atom Name	Atom Type	Charge	Atom Name	Atom Type	Charge
ATOM N1'	NN2	-0.29	ATOM H9	HA	0.09
ATOM C2'	CA	0.27	ATOM C7	CN8	0.03
ATOM CM2	CN9	-0.26	ATOM H12	HN8	0.08
ATOM N3'	N6R	-0.44	ATOM H13	HN8	0.08
ATOM C4'	CA	0.35	ATOM O7	ON2	-0.39
ATOM N4'	NN1C	-0.70	ATOM PA	P	1.13
ATOM C5'	CA	0.00	ATOM O1A	ON3	-0.86
ATOM C6'	CA	-0.07	ATOM O2A	ON3	-0.86
ATOM H	HN9	0.09	ATOM O3A	ON2	-0.43
ATOM H1	HN9	0.09	ATOM PB	P2	1.20
ATOM H2	HN9	0.09	ATOM O1B	ON3	-0.93
ATOM H3	HN3	0.09	ATOM O2B	ON3	-0.93
ATOM H14	HN1	0.32	ATOM O3B	ON3	-0.93
ATOM H15	HN2	0.25	ATOM CA	CE1	0.21
ATOM C7'	CT2	-0.12	ATOM OA	OH1	-0.54
ATOM N3	NR3	-0.12	ATOM HA	HO	0.38
ATOM C2	CE1	0.07	ATOM CA1	CA	-0.01
ATOM S1	S5R	-0.10	ATOM CA2	CA	-0.11

Table 5.6. Enamine/carbanion atom names utilized in ***Model 5, 8, 9*** and ***10*** (cont'd).

ATOM C4	CPH1	0.05	ATOM CA3	CA	-0.11
ATOM C5	CPH1	0.04	ATOM CA4	CA	-0.11
ATOM CM4	CT3	-0.26	ATOM CA5	CA	-0.11
ATOM C6	CT2	-0.16	ATOM CA6	CA	-0.11
ATOM H4	HA	0.09	ATOM HA2	HP	0.11
ATOM H5	HA	0.09	ATOM HA3	HP	0.11
ATOM H7	HA	0.09	ATOM HA4	HP	0.11
ATOM H8	HA	0.09	ATOM HA5	HP	0.11
ATOM H10	HA	0.09	ATOM HA6	HP	0.11
ATOM H11	HA	0.09			

Vcharge [91] predictions for the enamine/carbanion intermediate generated for ***Models 8, 9*** and ***10*** are displayed in Table 5.7 and Table 5.8 respectively. This molecule was also treated as a whole, because of the limitations of VCharge. We have faced another limitation of VCharge in this case, because of the negative charge on the carbon atom. Negative charges on a carbon atom are very rare in organic chemistry, and it was not possible to parameterize this atom. Instead, we have introduced the transition state geometry depicting the attack of enamine/carbanion to the acceptor aldehyde of the corresponding model system (Table 5.1, ***Models 8, 9*** and ***10***) to calculate the charges on the atoms. Atom types were generated from similar tautomers of nucleic acids in available CHARMM force fields [86, 87]. The *guesscoord* facility of VMD was used to add hydrogen atoms to the Protein Data Bank (PDB) structure derived from X-Ray diffraction study. The TIP3P water model was used in the solvation of model systems [92].

Table 5.7. Enamine/carbanion and benzaldehyde atom names in ***Models 8*** and ***9***.

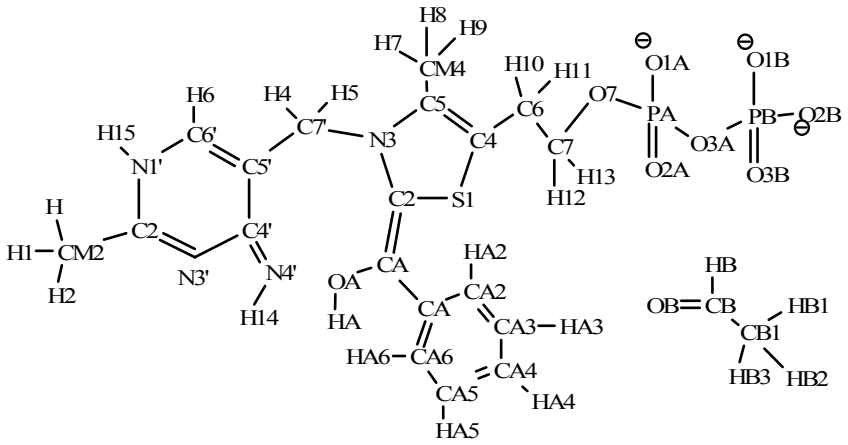


Atom Name	Atom Type	Charge	Atom Name	Atom Type	Charge
ATOM N1'	NN2	-0.29	ATOM PA	P	1.13
ATOM C2'	CA	0.27	ATOM O1A	ON3	-0.86
ATOM CM2	CN9	-0.26	ATOM O2A	ON3	-0.86
ATOM N3'	N6R	-0.44	ATOM O3A	ON2	-0.43
ATOM C4'	CA	0.35	ATOM PB	P	1.20
ATOM N4'	NN1C	-0.69	ATOM O1B	ON3	-0.93
ATOM C5'	CA	0.01	ATOM O2B	ON3	-0.93
ATOM C6'	CA	-0.07	ATOM O3B	ON3	-0.93
ATOM H	HN9	0.09	ATOM CA	CN8	0.25
ATOM H1	HN9	0.09	ATOM OA	OH1	-0.55
ATOM H2	HN9	0.09	ATOM HA	HO	0.39
ATOM H3	HN3	0.10	ATOM CA1	CA	-0.02
ATOM H4	HN1	0.32	ATOM CA2	CA	-0.11
ATOM H5	HN2	0.25	ATOM CA3	CA	-0.11
ATOM C7'	CT2	0.00	ATOM CA4	CA	-0.11
ATOM N3	NR3	0.09	ATOM CA5	CA	-0.11
ATOM C2	CE1	0.10	ATOM CA6	CA	-0.11
ATOM S1	S5R	-0.04	ATOM HA2	HP	0.11

Table 5.7. Enamine/carbanion and benzaldehyde atom names in ***Models 8*** and ***9*** (cont'd).

ATOM C5	CPH1	-0.01	ATOM HA3	HP	0.11
ATOM C4	CPH1	0.11	ATOM HA4	HP	0.11
ATOM CM4	CT3	-0.24	ATOM HA5	HP	0.11
ATOM C6	CT2	-0.15	ATOM HA6	HP	0.11
ATOM H6	HA	0.15	ATOM CB	CN8	0.19
ATOM H7	HA	0.15	ATOM OB	OH1	-0.88
ATOM H8	HA	0.09	ATOM HB	HO	0.08
ATOM H9	HA	0.09	ATOM CB1	CA	-0.02
ATOM H10	HA	0.09	ATOM CB2	CA	-0.11
ATOM HB5	HP	0.11	ATOM CB3	CA	-0.11
ATOM HB6	HP	0.11	ATOM CB4	CA	-0.11
ATOM H11	HA	0.09	ATOM CB5	CA	-0.11
ATOM H12	HA	0.09	ATOM CB6	CA	-0.11
ATOM C7	CN8	0.03	ATOM HB2	HP	0.11
ATOM H14	HN8	0.09	ATOM HB3	HP	0.11
ATOM H13	HN8	0.09	ATOM HB4	HP	0.11
ATOM O7	ON2	-0.39	ATOM HB5	HP	0.11

Table 5.9. Enamine/carbanion and benzaldehyde atom names in ***Model 10.***



Atom Name	Atom Type	Charge	Atom Name	Atom Type	Charge
ATOM N1'	NN2	-0.29	ATOMO7	ON2	-0.39
ATOM C2'	CA	0.27	ATOMPA	P	1.13
ATOM CM2	CN9	-0.26	ATOMO1A	ON3	-0.86
ATOM N3'	N6R	-0.44	ATOMO2A	ON3	-0.86
ATOM C4'	CA	0.35	ATOMO3A	ON2	-0.43
ATOM N4'	NN1C	-0.69	ATOMPB	P2	1.20
ATOM C5'	CA	0.00	ATOMO1B	ON3	-0.93
ATOM C6'	CA	-0.07	ATOMO2B	ON3	-0.93
ATOM H	HN9	0.09	ATOMO3B	ON3	-0.93
ATOM H1	HN9	0.09	ATOMCA	CN8	0.25
ATOM H2	HN9	0.09	ATOM OA	OH1	-0.55
ATOM H3	HN3	0.09	ATOM HA	HO	0.39
ATOM H4	HN1	0.32	ATOM CA1	CA	-0.02
ATOM H5	HN2	0.25	ATOM CA2	CA	-0.11
ATOM C7'	CT2	0.00	ATOM CA3	CA	-0.11
ATOM N3	NR3	0.09	ATOM CA4	CA	-0.11
ATOM C2	CE1	0.10	ATOM CA5	CA	-0.11
ATOM S1	S5R	-0.04	ATOM CA6	CA	-0.11
ATOM C5	CPH1	-0.01	ATOM HA2	HP	0.11

Table 5.10. Enamine/carbanion and benzaldehyde atom names in ***Model 10*** (cont'd)**.**

ATOM C4	CPH1	0.11	ATOM HA3	HP	0.11
ATOM CM4	CT3	-0.24	ATOM HA4	HP	0.11
ATOM C6	CT2	-0.15	ATOM HA5	HP	0.11
ATOM H7	HA	0.15	ATOM HA6	HP	0.11
ATOM H8	HA	0.09	ATOM CB	CD	0.18
ATOM H9	HA	0.09	ATOM OB	O	-0.88
ATOM H10	HA	0.09	ATOM HB	HO	0.08
ATOM H11	HA	0.09	ATOM CB1	CA	-0.27
ATOM H12	HA	0.09	ATOM HB1	HP	0.09
ATOM C7	CN8	0.03	ATOM HB2	HP	0.09
ATOM H13	HN8	0.08	ATOM HB3	HP	0.09
ATOM H14	HN8	0.08			

Protonation states of charged residues were predicted by propKa, a web based program (Table 5.9) [93, 94], those of catalytically active residues were predicted based on their interactions with surrounding residues according to the local interactions. The catalytically active E47 was assigned as deprotonated in all model systems, as it is necessary for ylide generation. Protonation states of H70 and H281 are listed in Tables 5.1 and 5.2.

Table 5.11. Assigned protonation states of charged amino acids (PropKa predictions)

Residue No.	Residue Name	Residue No.	Residue Name	Residue No.	Residue Name	Residue No.	Residue Name
5	HSP	141	ARG	245	ARG	356	GLU
10	GLU	144	HSP	246	HSE	360	ASP
13	ARG	161	ASP	250	ARG	364	ASP
14	ARG	162	ASP	265	GLU	368	GLU
18	ASP	164	ASP	267	HSP	375	GLU
28	GLU	165	LYS	268	ASP	386	ARG
33	LYS	166	ASP	279	ARG	390	ARG
34	ASP	168	ASP	281	-	417	GLU
37	GLU	172	HSP	284	ASP	419	GLU
38	ASP	173	HSP	290	LYS	420	ARG
40	ARG	176	ASP	294	ARG	428	ASP
47	GLU	177	ARG	301	ASP	462	ARG
55	ASP	178	HSP	304	GLU	469	GLU
62	ARG	184	ARG	307	ARG	471	GLU
63	LYS	187	ASP	312	ASP	477	ASP
70	-	189	ASP	317	ASP	482	ASP
89	HSD	191	ASP	330	GLU	484	ARG
101	ARG	195	LYS	331	GLU	488	LYS
107	GLU	210	ASP	334	ARG	496	LYS
114	ASP	212	ASP	342	GLU	498	ASP
120	ARG	219	ASP	345	LYS	501	GLU
124	LYS	225	GLU	347	ASP	504	LYS
128	GLU	226	ARG	349	ASP	509	GLU
134	GLU	228	LYS	352	ARG	514	LYS
137	HSD	239	ARG	354	HSP	520	GLU

5.2.6. Quantum Mechanics

The second nucleophilic attack of the reaction path was modeled with quantum mechanical tools. The geometrical parameters and energetics in the text were based on the results with the B3LYP methodology with 6-31+G* basis set. Vibrational frequency calculations were used to characterize all stationary points as either minima (no imaginary frequencies) or transition states (with only one imaginary frequency). Vibrational frequencies were also used to evaluate zero point energies and the thermodynamic data. Intrinsic reaction coordinate (IRC) calculations were traced to ensure that the transition states connect the

proper minima [79]. The QM optimized transition state geometry was used to generate atom types and parameters for the MD models representing the second nucleophilic attack (Tables 5.7 and 5.8). For the ***Models 8, 9*** and ***10*** we have used the corresponding transition state geometry to generate the necessary input for the MD calculations.

5.3. Results and Discussions

The complete machinery of ThDP action can only be elucidated with a detailed examination of the nature and behavior of key intermediates on the reaction path, determining their role in proton transfer mechanism to carry out the catalytic reaction. MD and QM tools were utilized in this study to provide the underlying mechanisms that lead to the experimental observations [34, 35, and 36] on the mode of action of ThDP in BFD.

We first check the extent to which the systems have equilibrated by monitoring the root mean square (RMS) deviations of the atomic positions as a function of time. The RMS fluctuations for the whole trajectory of all simulations were calculated for the backbone heavy atoms by fitting their positions to the positions in the X-Ray coordinates of the structure. The positional fluctuations of the C_α atoms should approximately match those extracted from the experimental temperature factor values. Those of ***Model 1*** are presented in Figure 5.7. The calculated values are in good agreement with the experimentally predicted ones. The exceptions are the regions of high fluctuations spanning the residue index ranges of 163-188 and 332-355. However, there are located far from the reaction center, in the solvent exposed outer region of the model system and are stabilized by the neighboring dimers in the crystal structures.

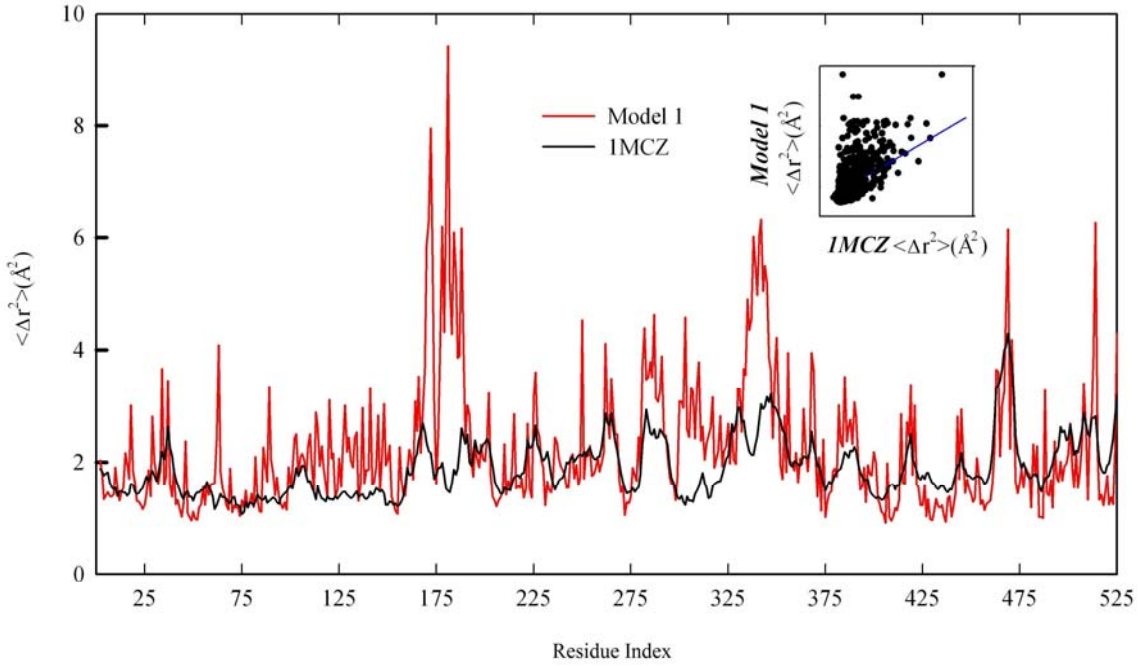


Figure 5.7. Positional fluctuations of the C_{α} atoms for *Model 1* and BFD crystal structure (PDB code: 1MCZ). The two sets of data are plotted against each other in the inset; the $y = x$ line is shown in blue to guide the eye.

We find that in all of the simulations (5 ns) the fluctuations were stabilized to a value of ca. 2 Å after 0.5 ns equilibration was completed. RMSD for the whole trajectory of all model systems as a function of simulation time for the backbone heavy atoms by fitting their positions to the positions in the X-Ray coordinates of the structure are shown in the Figure 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14, 5.15, 5.16 and 5.17. The picture is the same for all the 5 ns long simulations: The fluctuations are stabilized after 0.5 ns equilibration is completed.

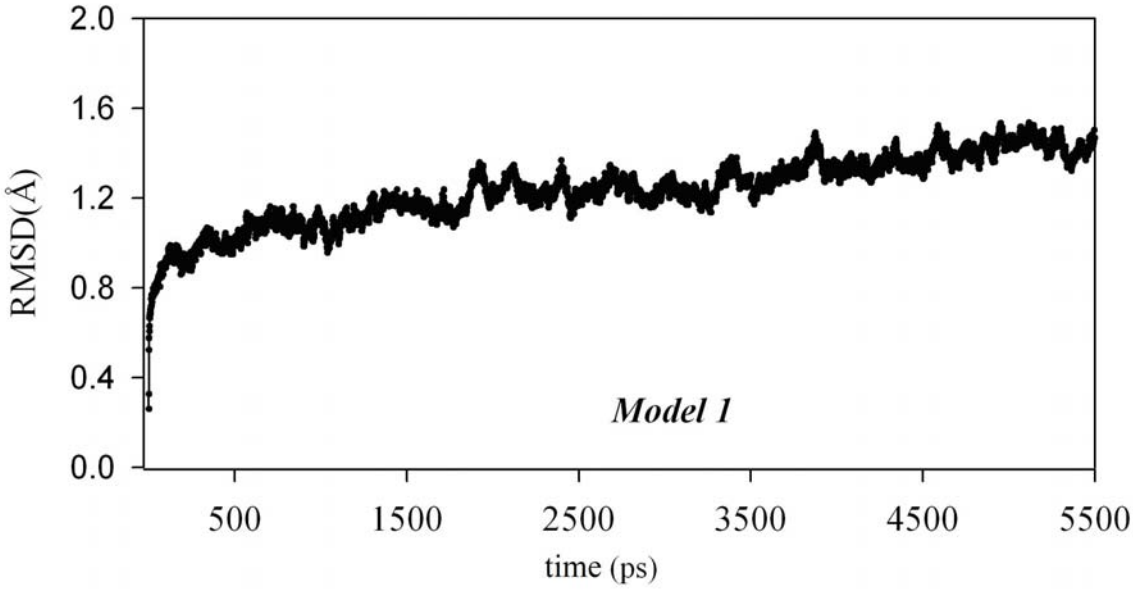


Figure 5.8. RMSD vs. time graphs for ***Model 1***.

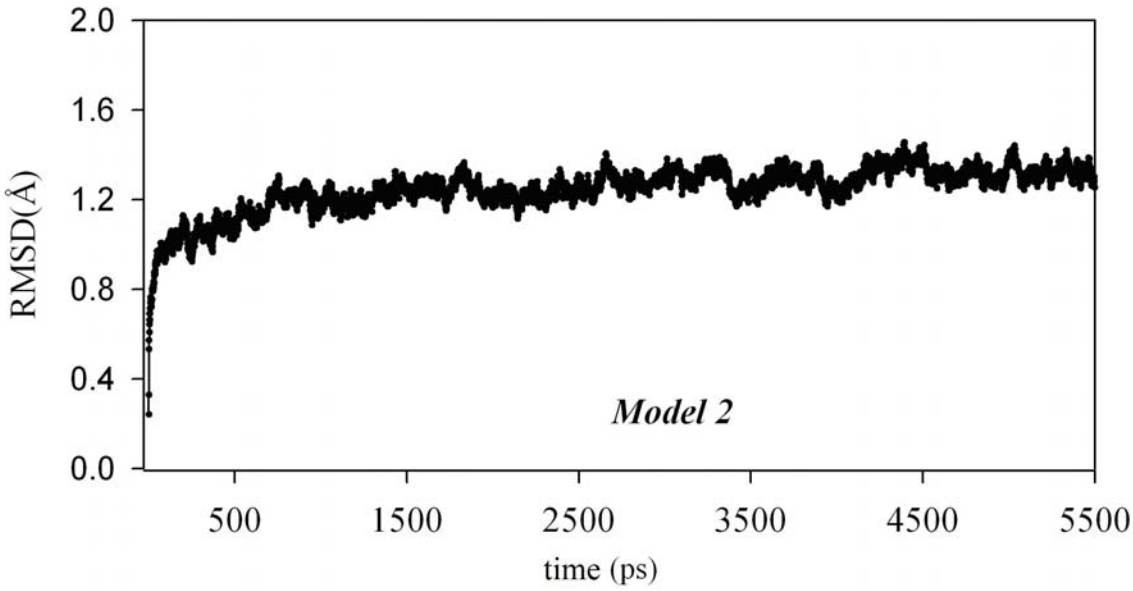


Figure 5.9. RMSD vs. time graphs for ***Model 2***.

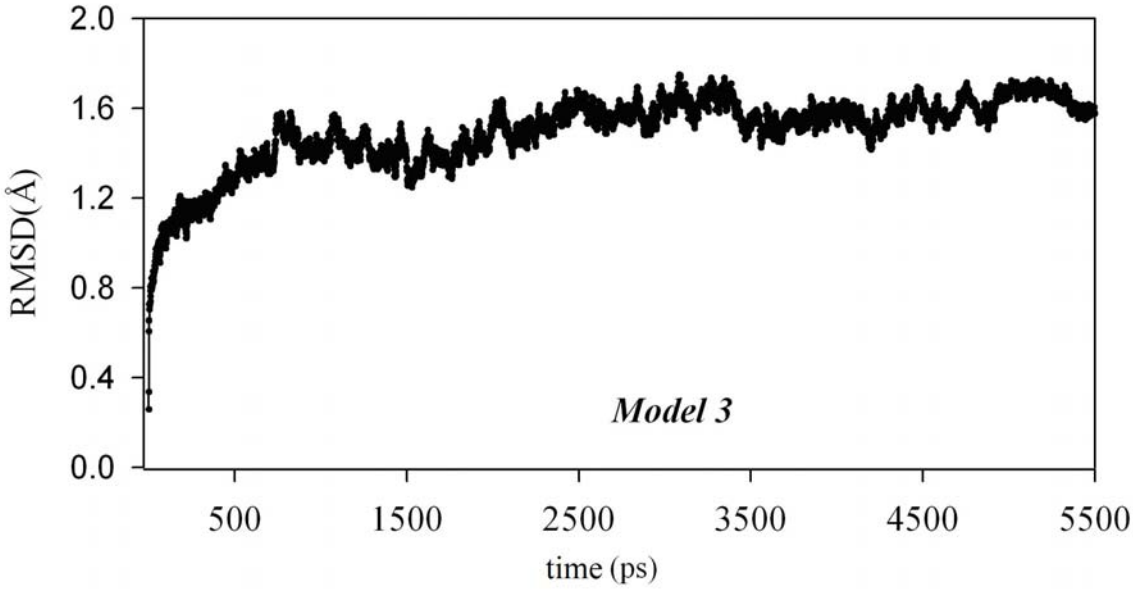


Figure 5.10. RMSD vs. time graphs for ***Model 3***.

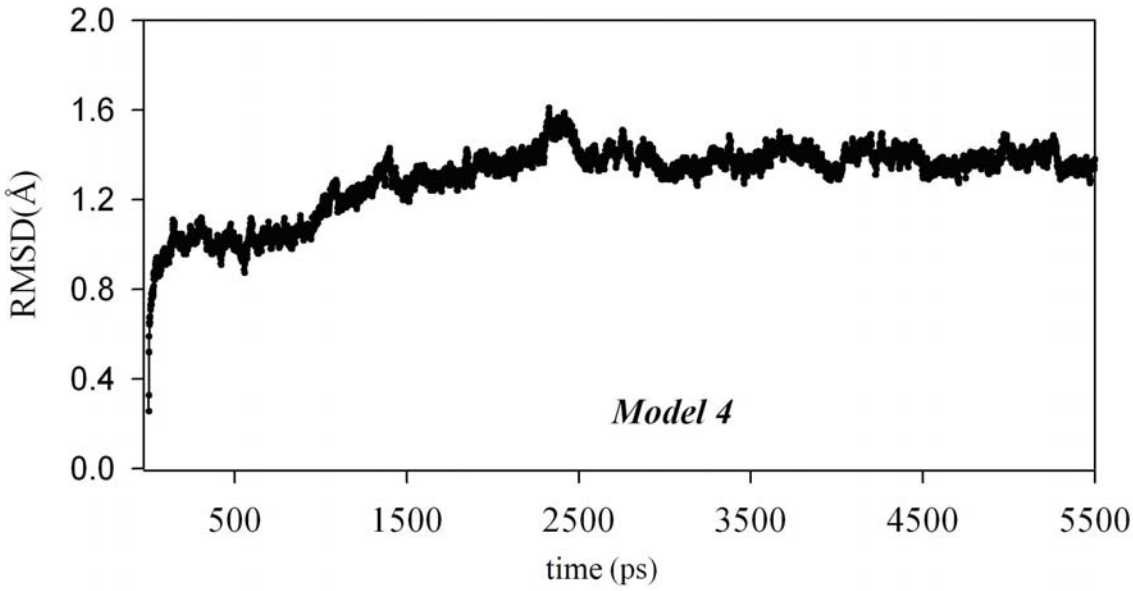


Figure 5.11. RMSD vs. time graphs for ***Model 4***.

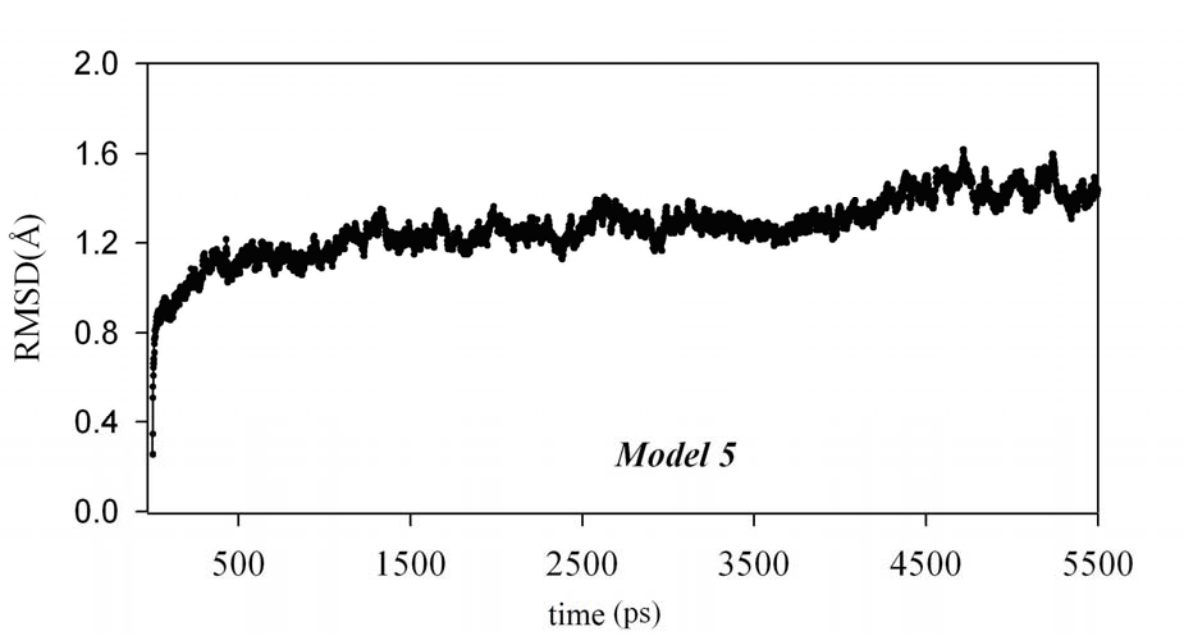


Figure 5.12. RMSD vs. time graphs for ***Model 5***.

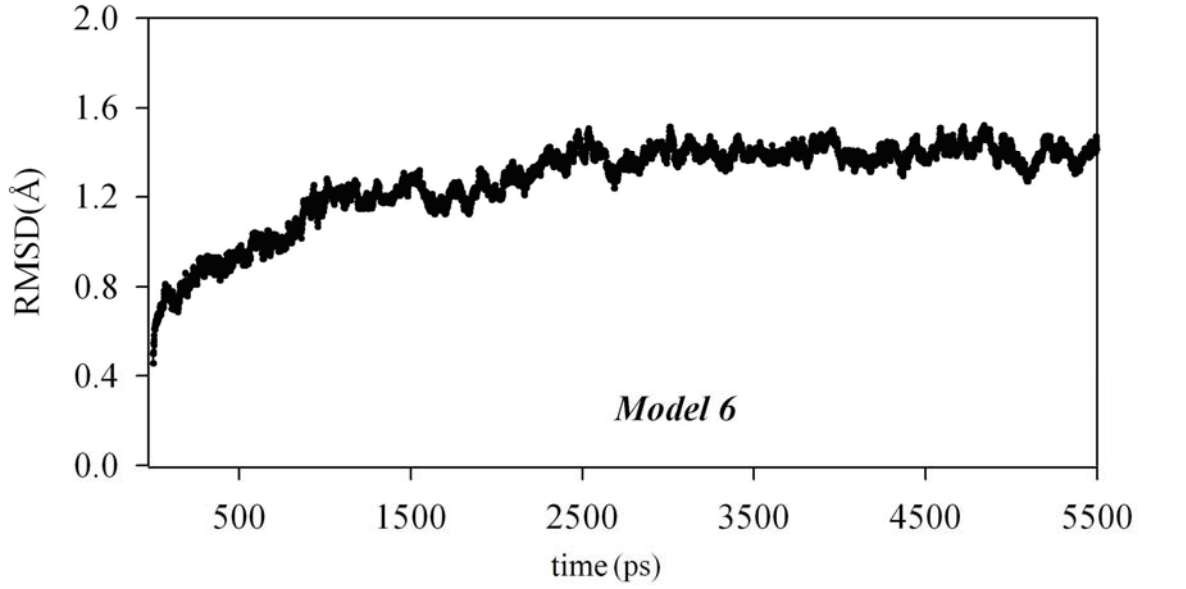


Figure 5.13. RMSD vs. time graphs for ***Model 6***.

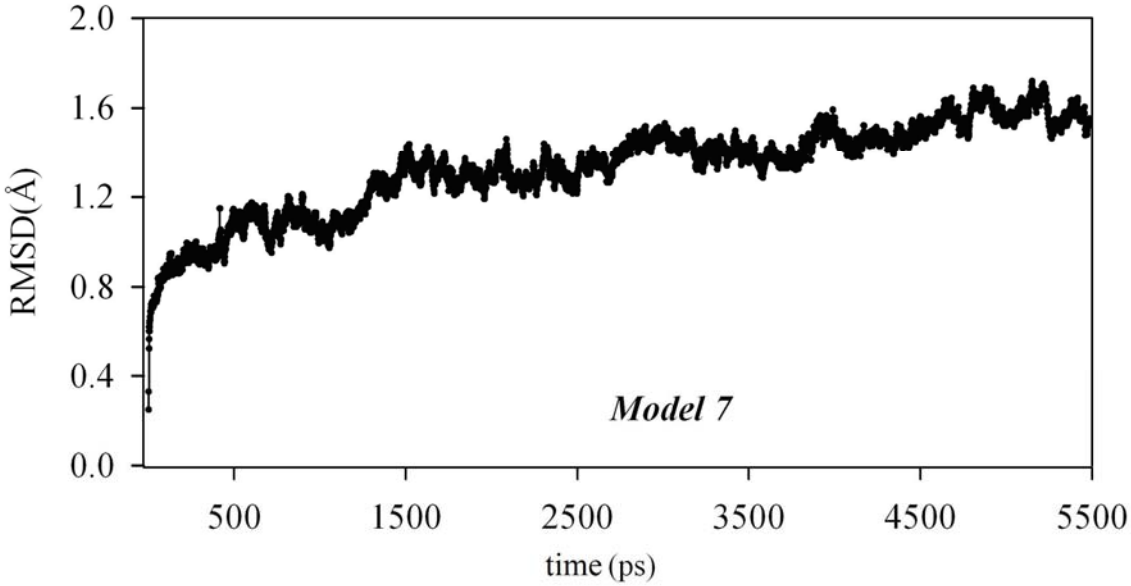


Figure 5.14. RMSD vs. time graphs for ***Model 7***.

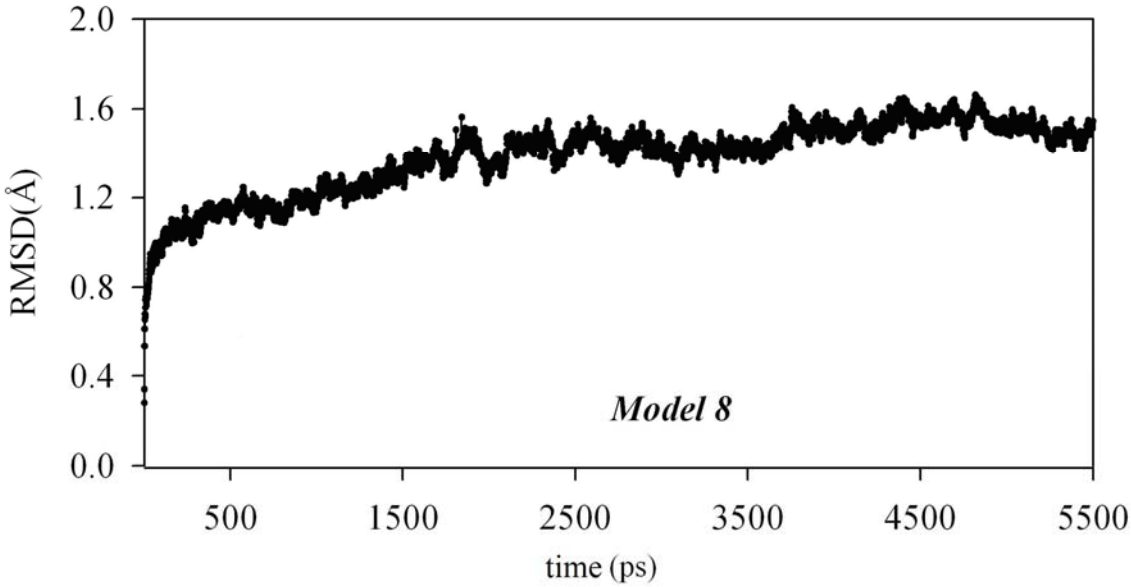


Figure 5.15. RMSD vs. time graphs for Model 8.

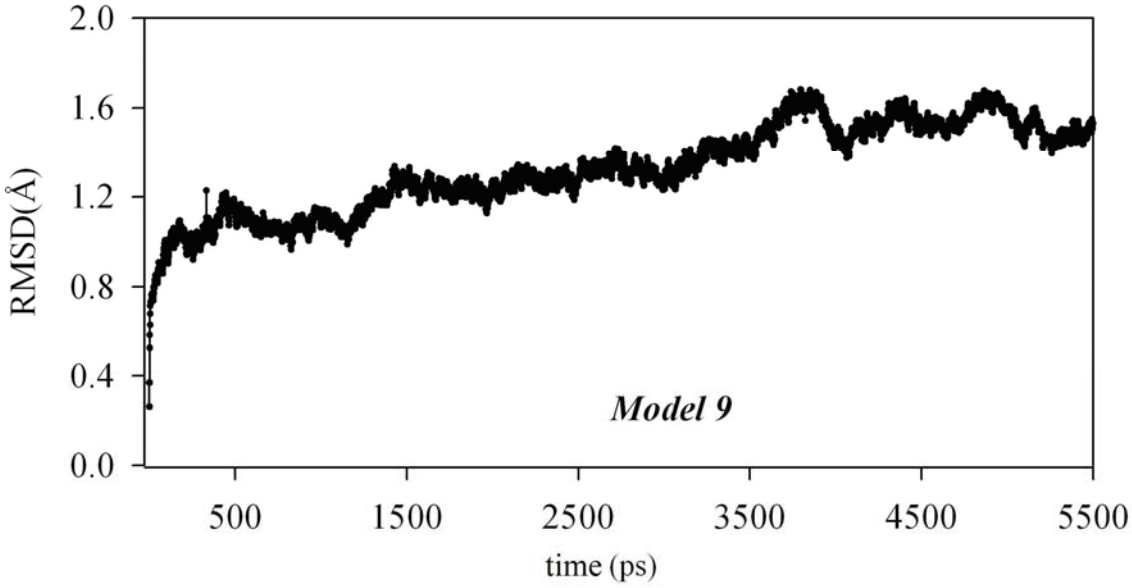


Figure 5.16. RMSD vs. time graphs for ***Model 9***.

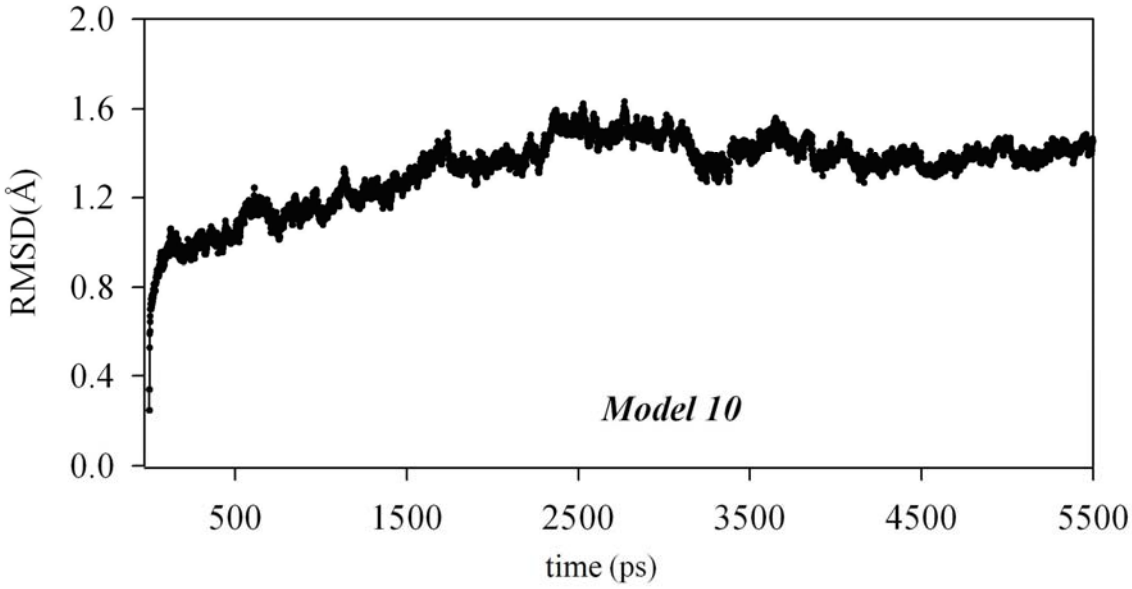


Figure 5.17. RMSD vs. time graphs for Model 10.

The interaction with the surrounding residues confines ThDP in a special "V" shape (Figure 5.18-5,22) which is described by the orientation of pyrimidine and thiazolium rings,

through the torsional angles, ϕ_P and ϕ_T, flanking the bridging atom between the two rings. ϕ_P is the torsional angle on the pyrimidine side and ϕ_T on the other. One of the most important indicators to monitor the stability of MD simulations of targeted systems in this study are the dihedral angles, ϕ_P and ϕ_T, defining the "V" shape of ThDP (Figure 5.1). ϕ_T is expected to be ca. 95.4° and ϕ_P is expected to be ca. -64.7° [35]. We find that the "V" shape of ThDP was conserved in all of the simulations within ± 25˚ (Figures 5,18-5.22) (only in the following models the average of the ϕ_T angles deviates largely from the experimental predictions: 116˚ for ***Model 6***, 120˚ for ***Model 8***, 75˚ ***Model 10***) (Figure 5.20, 5.21, 5.22).

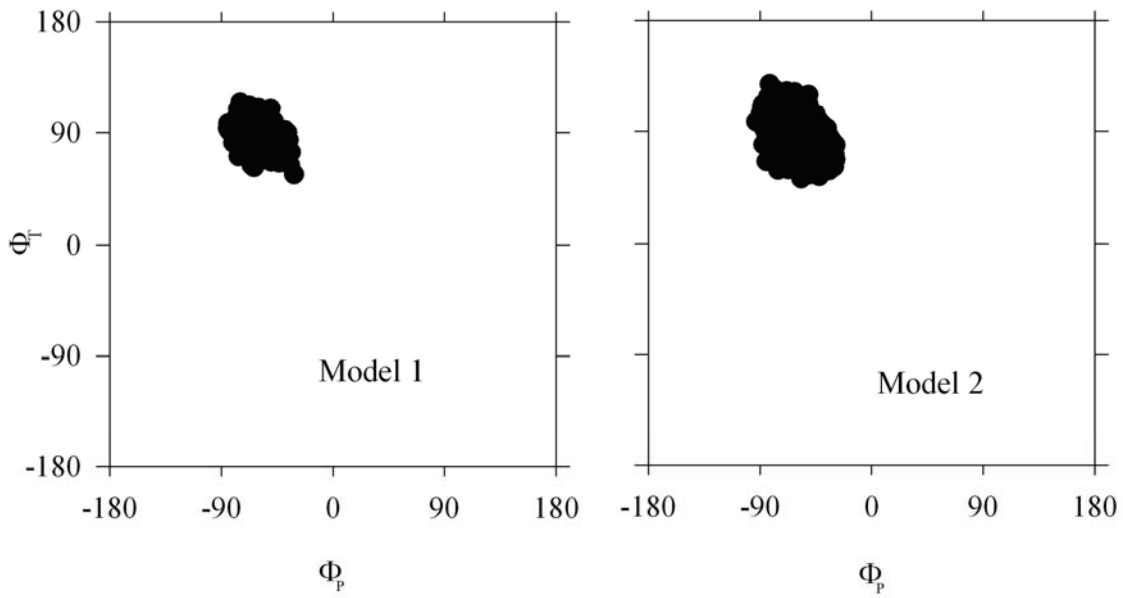


Figure 5.18. Dihedral angles ϕ_P vs. ϕ_T for ***Models 1 and 2***.

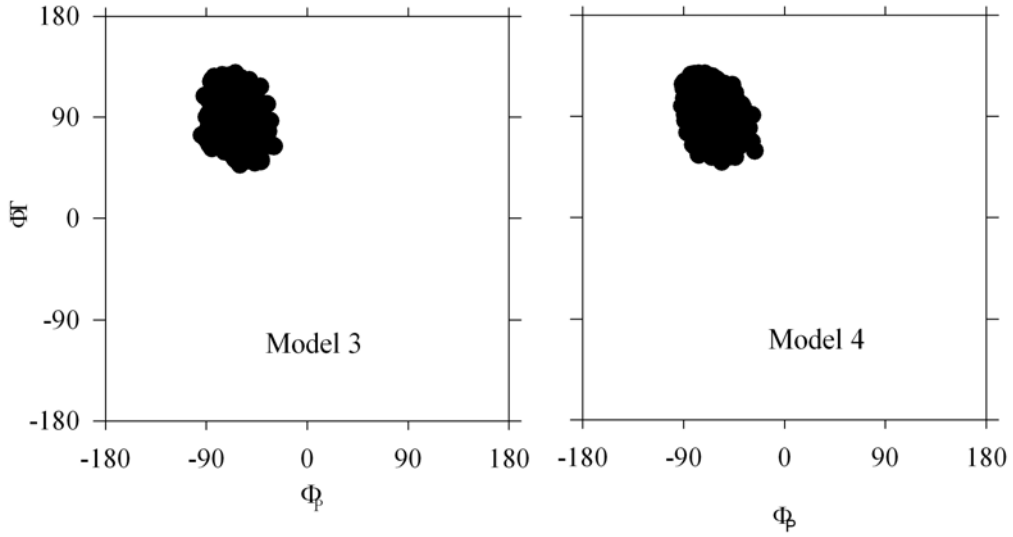


Figure 5.19. Dihedral angles ϕ_P vs. ϕ_T for ***Models 3 and 4***.

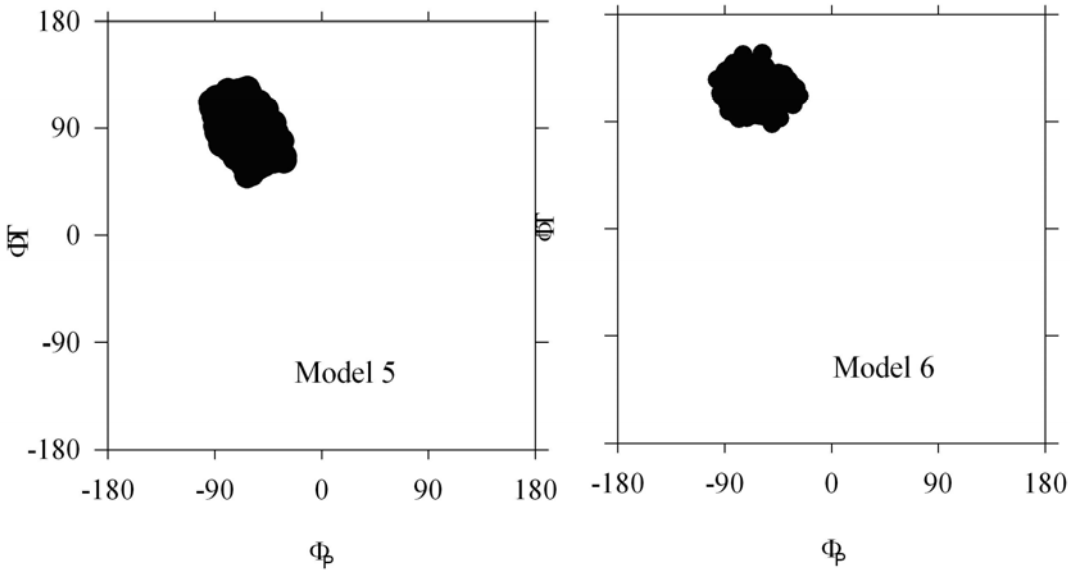


Figure 5.20. Dihedral angles ϕ_P vs. ϕ_T for ***Models 5 and 6***.

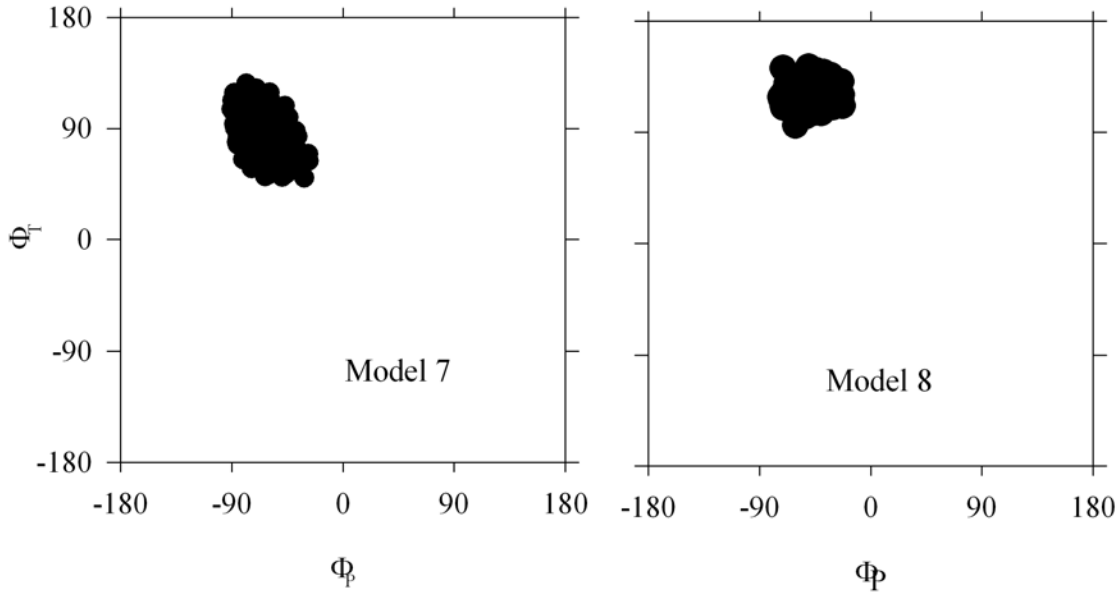


Figure 5.21. Dihedral angles ϕ_P vs. ϕ_T for ***Models 7 and 8***.

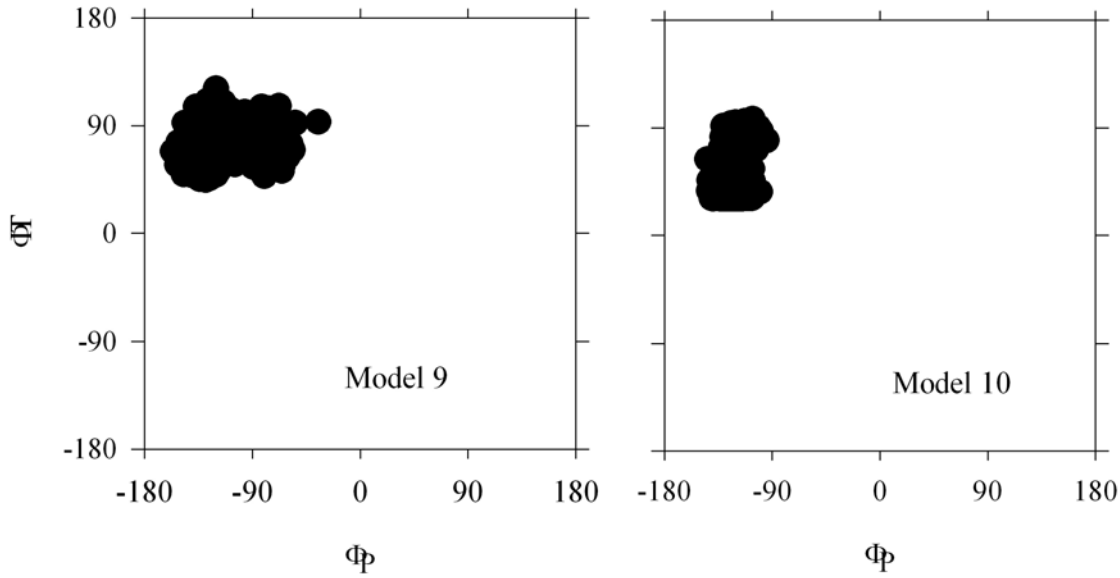


Figure 5.22. Dihedral angles ϕ_P vs. ϕ_T for ***Models 9 and 10***

Another characteristic property of ThDP is the octahedral arrangement of some residues and ThDP diphosphate oxygen atoms (Figure 5.23). Octahedral geometry around the Mg (II) cation is established by two coordination covalent bonds with a distance of ca. 1.8 ± 0.2 Å to the two oxygen atoms of the diphosphate tail. The other corners of the octahedron are occupied by D428, N455, T457 and a water molecule. The interaction distances between the surrounding molecules and Mg (II) are 1.8 Å ± 0.7 Å, 2.0 Å ± 0.3 Å, 2.1 Å ± 0.4 Å and 2.0 Å ± 0.3 Å , respectively, in ***Model 1***. Octahedral geometry around Mg (II) is strictly conserved in all model simulations (Table 5.10). Correlation of the dihedral angles with experimental data along with the confirmed equilibrium properties justifies the usage of our model systems and the parameters used in representing their interactions.

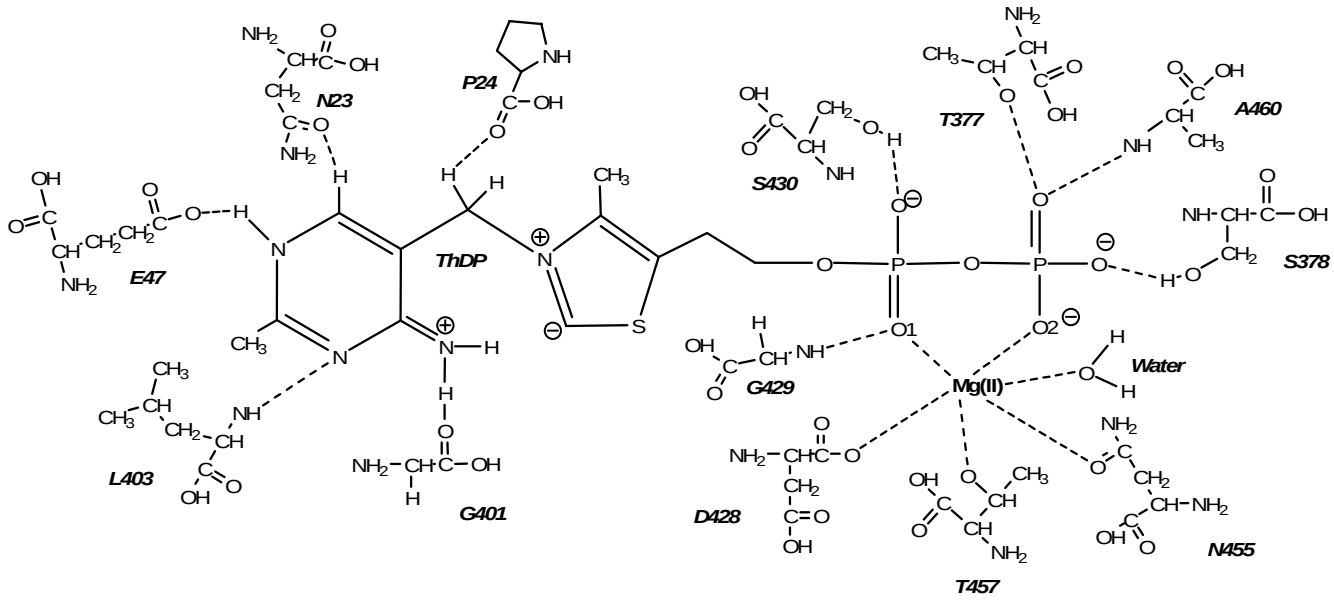


Figure 5.23. Binding of ThDP and the octahedral geometry around Mg (II) cation.

Table 5.12. Geometric parameters representing the octahedral geometry of the Mg (II) cation and the interactions of ThDP with surrounding residues. Average values are calculated for the conformers produced during the MD simulations.[†]

	Mg(II)						ThDP							
	N455	**D428**	**T457**	**Water**	**O1**	**O2**	**E47**	**T377**	**G401**	**L403**	**S430**	**G429**	**A460**	**S378**
Model 1	2.05	1.82	2.15	2.05	1.85	1.78	1.73(161)	2.80	2.26(144)	2.30(163)	1.65(167)	2.05(158)	1.83(165)	1.63(156)
Model 2	2.03	1.82	2.09	2.04	1.85	1.81	1.73(163)	5.38	1.92(159)	2.28(160)	1.65(168)	2.02(160)	1.79(166)	1.68(168)
Model 3	2.07	1.82	2.06	2.02	1.85	1.80	1.74(160)	2.51	1.96(155)	2.37(161)	1.66(166)	2.06(156)	1.86(164)	1.64(168)
Model 4	2.06	1.82	2.07	2.05	1.84	1.80	1.73(163)	4.30	1.97(159)	2.36(164)	1.68(167)	2.06(157)	1.83(165)	1.67(167)
Model 6	2.06	1.82	2.08	2.03	1.84	1.80	1.75(163)	4.57	1.91(156)	2.25(162)	1.63(166)	2.07(156)	1.76(165)	1.62(168)
Model 8	2.07	1.83	2.06	2.00	1.82	1.83	1.86(161)	1.71(157)	4.21127)	2.97147)	1.64(168)	1.94(161)	3.55(131)	1.68(168)
Model 9	2.11	1.83	2.05	2.02	1.82	1.82	2.82(133)	1.71(159)	4.04(123)	2.11(152)	1.64(167)	1.20(156)	3.79(132)	1.68(167)
Model 10	2.12	1.83	2.04	2.03	1.82	1.83	2.65(111)	1.75(156)	2.84(130)	2.95(152)	1.64(168)	1.96(159)	3.34(137)	1.69(168)
X-Ray	2.32	2.34	2.29	2.10	2.76	2.73	2.65	2.59	2.65	3.19	2.76	2.51	2.82	2.77

[†] All distances are in Å. Interaction angles (°) for hydrogen bonds are presented in parentheses. Models representing mutation are not included, since they create artificial changes in the enzyme environment

5.3.1. Proton Donors/Acceptors in the First Nucleophilic Attack

Since MD simulations do not directly model the reaction step, the suitability of a model is assessed by monitoring the conformations attained by the reacting partners during the course of the run [95]. Termed as near-attack conformation (NAC), those reactants that sustain a favorable geometry for the reaction to actually take place are analyzed in detail. The NAC percentages of the models proposed are compared to elucidate the protonation states of amino acids in the reaction path. In the NAC percentage of a nucleophilic attack, we have considered the C2-C_{donor} bond distance, the C2-C_{donor}-O_{donor} angle and the N3-C2-C_{donor}-O_{donor} dihedral angle in the conformers generated during the simulations. First, we have chosen the subset of conformers that have the C2-C_{donor} bond distance less than 3.4 Å. Among these conformers, we have picked those that have the C2-C_{donor}-O_{donor} angle within ± 20° of 105° [96] and finally concentrated on the conformers satisfying dihedral criteria for nucleophilic attack, the N3-C2-C_{donor}-O_{donor} dihedral within ±20° of 25°. Furthermore, in the NAC percentage of a developing hydrogen bond, we have considered the hydrogen bond distance to be less than 2.8 Å (sum of van der Waals radii of an oxygen and a hydrogen atom). The NAC percentages displayed in Tables 5.11 and 5.12 are the ratio of conformers satisfying the corresponding NAC criteria to the total number of conformers produced in the simulation.

Table 5.13. NAC percentages for nucleophilic attacks.

	C2-C$_{donor}$ [a]	***C2-C$_{donor}$-O$_{donor}$*** [b]	***C7′-N3-C2-C$_{donor}$*** [c]
Model 3	49.7	37.8	22.5
Model 4	20.0	12.5	8.0
Model 5	24.0	12.9	0.3
Model 7	66.1	55.1	29.6

[a] < 3.4 Å , [b] 105± 20°, [c] 25± 20°.

Table 5.14. H_{bond} NAC percentages of hydrogen bonds in model systems. Refer to Figure 5.28 for the definition of bonds.

	N4′- O_{donor}	**H70 (bond a)**	**S26 (bonds b/c)**	**H281 (bonds d/e)**
Model 6	99.6	—	—	—
Model 8	—	100	50.9 /97.4	19.1/72.1
Model 9	—	78.8	76.4*	—
Model 10	—	100	39.9/39.4	—**

* In this case S26 directly interacts with $O_{acceptor}$ without the mediating water molecule.

** H281 interacts with $O_{acceptor}$ with a bridge composed of two water molecules. Percentages for ThDP-water1, water1-water2, water2-H281 are 45.3%, 5.8% and 7.0% respectively.

With this definition of NACs for nucleophilic attack, 49.7 % of the conformers in ***Model 3*** were in close van der Waals contact between C2 and C_{donor} and 37.8 % of these conformers were located within ± 20° of the Bürgi-Dunitz angle of 105° (Table 5.11). In ***Model 3***, 22.5 % of the conformers have the thiazolium ring, containing C2 and benzaldehyde lying in the same plane (25° ± 20°). High NAC percentage of this model shows that the first nucleophilic attack probably proceeds via the proposed proton donor/acceptor mechanism in ***Model 3***. According to this proposal, H70 should be protonated and H281 should be deprotonated (Figure 5.24).

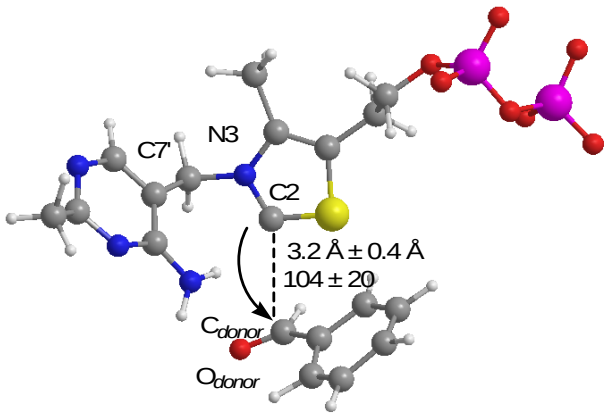


Figure 5.24. Nucleophilic attack of Ylide-ThDP to benzaldehyde (*Model 3)*. Average values for the C2-C_{donor} bond distance and the C2-C_{donor}-O_{donor} angle of the conformers satisfying NAC criteria are depicted.

In this model, the alignment of H281, S26 and a water molecule as described in Figure 5.25 is important to notice. This is a long hydrogen transfer bridge, implying that H_{donor} is transferred to H281 via this proton bridge. In 28.7 % of the conformers S26-water hydrogen bond, in 19.9 % of the conformers H281-water hydrogen bond and in 19.8 % of the conformers H_{donor}-S26 hydrogen bond length is less than 2.8 Å. The protonation of O_{donor} is claimed to have already taken place before the nucleophilic attack [97].

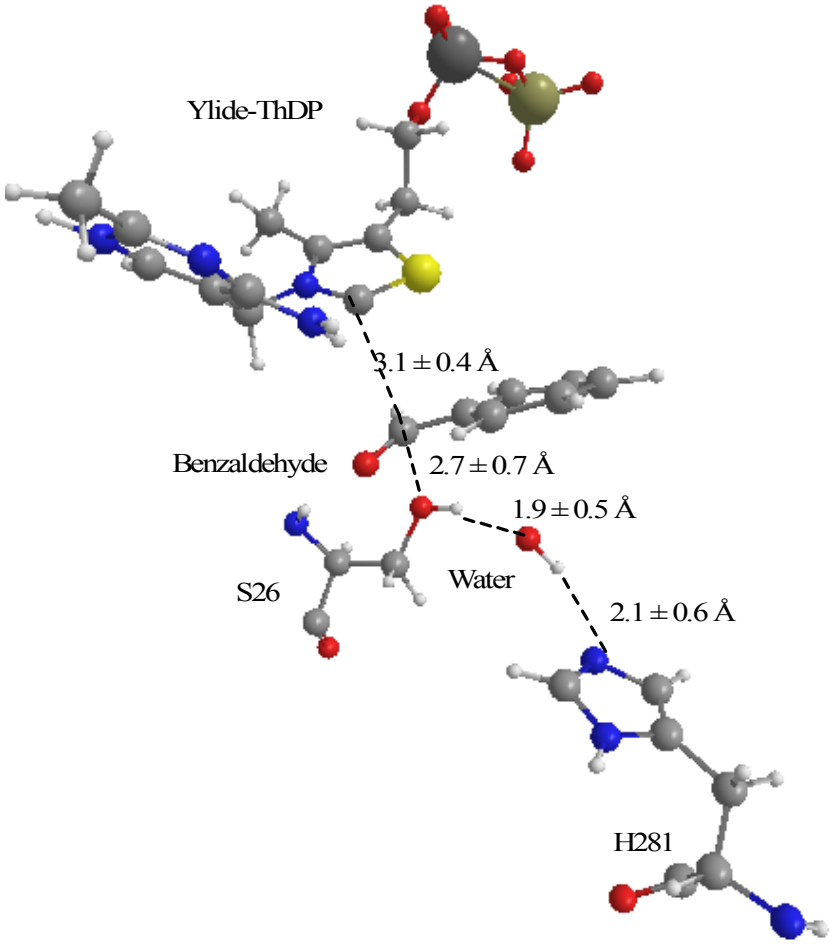


Figure 5.25. H_{donor} transfer bridge in *Model 3*. Average bond distances within the NAC limits for the first nucleophilic attack on the trajectory of the model.

In ***Model 3***, H70 points one of its protons towards O_{donor}. N4′ also directs a proton towards the same oxygen. This three-membered interaction aligns benzaldehyde towards the C2 atom before the reaction. The interactions of N4′ hydrogen and H70 hydrogen with O_{donor} are not within the NAC limits of a regular hydrogen bond. The average bond length among the conformations within the NAC limits is 3.7 ± 1.3 Å for N4′ hydrogen and 3.7 ± 1.8 Å for H70 hydrogen. Therefore, the function of H70 may be to electrostatically interact with the donor aldehyde and align it into the correct nucleophilic attacking position. The proton donor at this

stage is most probably N4′, since it is the intimate neighbor of O_{donor}. Jordan *et al.* have experimentally identified intermediates where N4′ is the proton donating species [84, 98]. The functions of **HB1** and **B3** in the catalytic cycle of the enzyme (Figure 5.2) might thus be carried out by the same group, the 4′-amino group of ThDP.

As it is a charged amino acid, histidine can have two protonation states. To check the protonation state defined in ***Model 3***, we use ***Model 4*** as an alternative where H281 is the charged residue around the active center that has the potential to transfer hydrogen. The NAC percentage of conformers satisfying all of the three NAC requirements in ***Model 4*** is 8.0%, which is approximately one third of nucleophilic attack the NAC percentages of ***Model 3***. Low NAC percentage of ***Model 4*** indicates a lower probability for the model to represent the real system. In ***Models 1*** and ***3***, H281 is modeled to be deprotonated to behave as a proton acceptor, **B2**. Hasson and coworkers assigned a catalytic function to this histidine residue [35]. In ***Model 4***, H281 is in its protonated form. Based on the high NAC percentages observed with ***Model 3*** and the proton transfer bridge (S26-water-H281), we conclude that H281 should be present in its deprotonated form and function as **B2/HB4**. Although H281 is located far from the nucleophilic center, the aldehyde proton can be transferred to H281 with the assistance of S26 and a water molecule (Figure 5.25).

In ***Model 5***, H70 is mutated to an alanine residue which is incapable of hydrogen bonding. The overall NAC percentage drops to 0.3%, implying that H70 is a catalytically important amino acid for the nucleophilic attack to take place, in agreement with the experimental predictions. The decrease in the NAC percentages and experimental yield with the H70A mutation was previously attributed to either it being a proton donor or having an electrostatic effect to align the benzaldehyde molecule towards ThDP. ***Model 3*** and ***Model 5*** corroborate the former explanation.

In addition to the two alternative models to ***Model 3***, we have also investigated post reactive interactions to elucidate the correct protonation states of the amino acids. In ***Model 6***, the intermediate product: enamine/carbanion, **IV** is modeled in enzyme environment as a

continuation of protonation states proposed in ***Model 3***. 99.6% of the conformers produced by ***Model 6*** have a hydrogen bond between the N4′ nitrogen and the proton of O_{donor} with bond length less than 2.8 Å. In these conformers, the hydroxyl proton of enamine is directed towards N4′. Although it can rotate, the hydroxyl proton in the majority of the conformers stays at N4′ side of the molecule by conserving the hydrogen bond distance. The nucleophilic attack starting with ***Model 3*** yields ***Model 6*** with considerably high NAC percentages.

In ***Model 7***, we study the effect of the S26A mutation in the enzymatic activity. It is known that S26A mutation decreases the yield of the acyloin reaction [35]. Serine is known to function as a proton carrier even though it is not a charged residue [99]. It transfers the proton via its side chain hydroxyl group. We find that the NAC percentage for the C2-C_{donor} bond distance throughout the simulation and the C2-C_{donor}-O_{donor} bond angle are greater than that of ***Model 3*** (66.1% and 55.1%, respectively; Table 5.10). None of the conformations within the NAC limits were able to interact with H281, since alanine can not allocate a proton. The nucleophilic attack NAC percentages for this model are high, since the relatively bulky side group of S26 is removed from the active center, which leaves more space for the donor aldehyde to approach the active site of the enzyme. The results obtained with this model indicate that S26 is not an active residue before the first nucleophilic attack takes place. S26 is removing the H_{donor} after the attachment of donor aldehyde to ylide-ThDP.

5.3.2. Proton Donors/Acceptors in the Second Nucleophilic Attack

The second nucleophilic attack in the acyloin formation is the nucleophilic attack of enamine/carbanion, **IV** to the acceptor aldehyde, which ends up with a bond formation between C_{donor} and $C_{acceptor}$. We first report a QM model for this step of the reaction. We have located a transition state and modeled it with the B3LYP/6-31+G* methodology (Table 5.13). In this transition state, the *re face* of the enamine attacks the *si face* of the acceptor aldehyde (Figure 5.26). Our QM models also form a basis for the related MD models in the selection of atomic parameters. We have also investigated the reactivity differences between benzaldehyde and acetaldehyde as substrates towards enamine/carbanion **IV**. The activation barrier for

benzaldehyde was 1.6 kcal/mol higher than that of acetaldehyde. The free energy barrier of the nucleophilic attack to benzaldehyde is 5.2 kcal/mol higher than that of acetaldehyde. Thus, we conclude that the nucleophilic attack to acetaldehyde is more favored than the one for benzaldehyde. This finding is consistent with the experimental predictions with the enzyme mediated acyloin condensation reaction with acetaldehyde and benzaldehyde as the respective acceptor aldehyde [100], where the former gives higher yields.

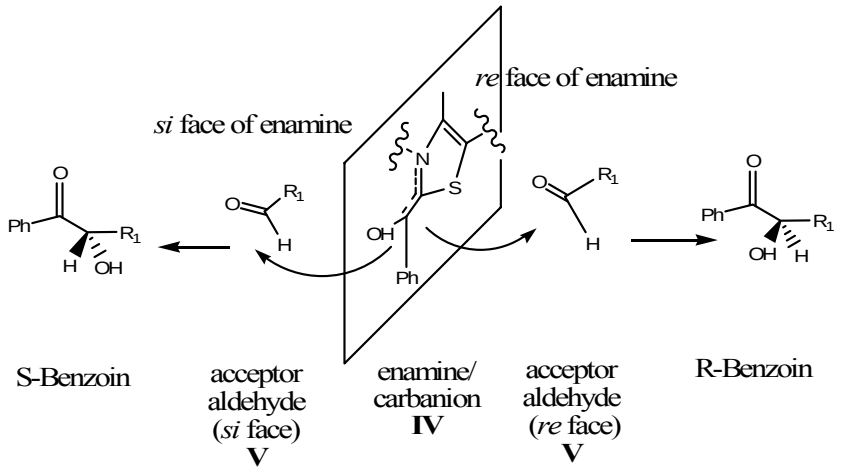
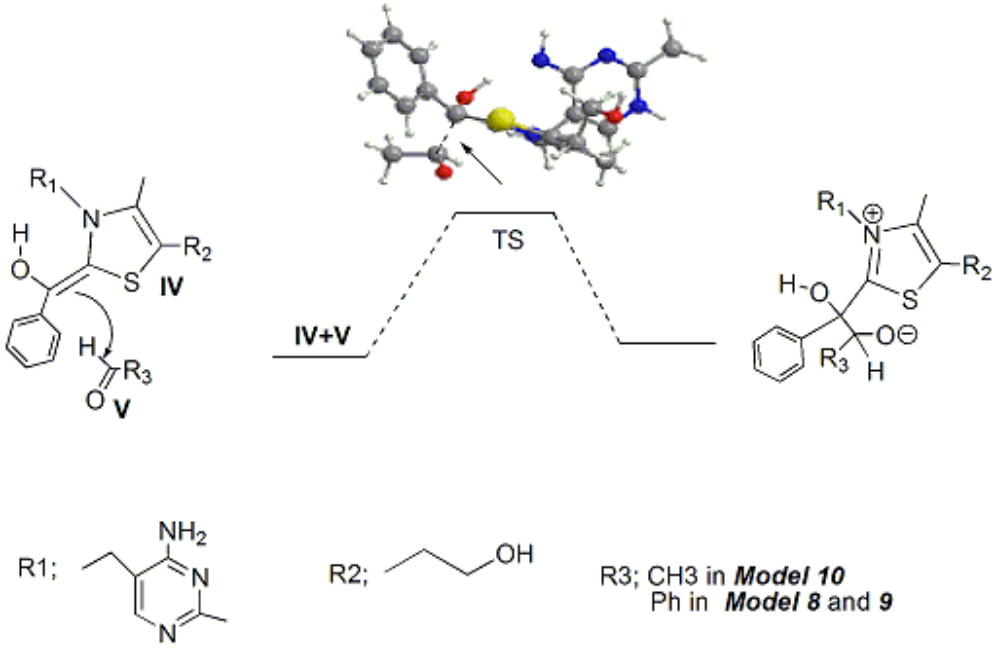


Figure 5.26. Enantioselectivity in the nucleophilic attack of enamine/carbanion, IV to the acceptor aldehyde, V. R_1= CH_3 for *S*-HPP and R_1= Ph for *R*-benzoin.

Table 5.15. Energetics of the second nucleophilic attack. *



R_3	Total Energy **IV** + **V** (au)	ΔE ‡a	$(\Delta E + ZPE)$ ‡b	ΔG ‡ c	ΔH ‡ d
CH_3	-1658.6864 (-1658.7429)	29.9 (29.6)	32.0 (31.7)	41.6 (45.0)	30.9 (30.6)
Ph	-1850.4297 (-1850.4920)	31.5 (32.0)	32.6 (33.0)	46.8 (47.2)	32.0 (32.5)

* B3LYP/6-31G*energies are given in the table and energies with B3LYP/6-31+G* are in parentheses.

[a] Relative energy with respect to reactants (**IV** and **V**).

[b] Relative energy with respect to reactants with zero-point correction.

[c] Relative free energy with respect to reactants.

[d] Relative enthalpy with respect to reactants

The free energy difference between acetaldehyde and benzaldehyde is triggered by the difference in Mulliken charge distributions of the reactants (the charge distribution of reactants and transition states of the two models were collected in Figure 5.27 and Table 5.14). The charge on $C_{acceptor}$ of acetaldehyde is more positive than the one on benzaldehyde (+0.26 vs. +0.19, respectively). The more positively charged $C_{acceptor}$ suggests an increased electrophilic character on this atom. The $C_{acceptor}$ of acetaldehyde facilitates the nucleophilic attack.

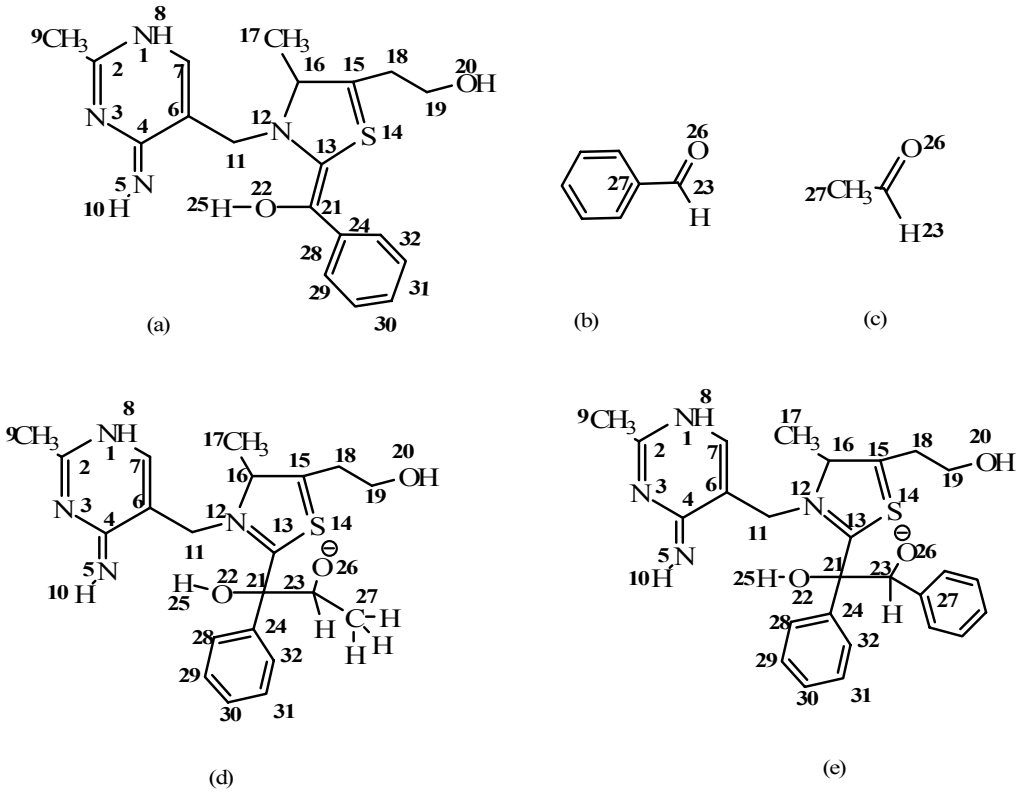


Figure 5.27. Atom numbers for charge distribution in QM models, (a) BFD-BA, (b) BA, (c) AA, (d) BFD-BA+AA, (e) BFD-BA+BA

Table 5.16. Charge Distribution in QM Models.

		Charges				
Atom Number	**Atom**	**BFD-BA+AA**	**BFD-BA+BA**	**AA**	**BA**	**BFD-BA**
1	N	-0.61	-0.61	-	-	-0.61
2	C	0.50	0.50	-	-	0.49
3	N	-0.51	-0.51	-	-	-0.51
4	C	0.43	0.43	-	-	0.45
5	N	-0.69	-0.70	-	-	-0.69
6	C	0.16	0.16	-	-	0.16
7	C	-0.01	-0.01	-	-	-0.02
8	H	0.34	0.34	-	-	0.34
9	C	-0.53	-0.53	-	-	-0.53
10	H	0.31	0.31	-	-	0.31
11	C	-0.29	-0.29	-	-	-0.25
12	N	-0.41	-0.41	-	-	-0.50
13	C	0.17	0.19	-	-	-0.02
14	S	0.33	0.34	-	-	0.31
15	C	-0.21	-0.21	-	-	-0.16
16	C	0.39	0.39	-	-	0.31
17	C	-0.54	-0.54	-	-	-0.54
18	C	-0.33	-0.33	-	-	-0.35
19	C	-0.03	-0.03	-	-	-0.03
20	O	-0.61	-0.61	-	-	-0.61
21	C	0.13	0.11	-	-	0.19
22	O	-0.67	-0.67	-	-	-0.66
23	C	0.28	0.20	0.26	0.19	-
24	C	0.10	0.09	-	-	0.08
25	H	0.43	0.45	-	-	0.47
26	O	-0.68	-0.70	-0.38	-0.41	-
27	C	-0.47	0.11	-0.52	0.10	-
28	C	-0.18	-0.17	-	-	-0.14
29	C	-0.13	-0.14	-	-	-0.14
30	C	-0.13	-0.13	-	-	-0.12
31	C	-0.14	-0.14	-	-	-0.13
32	C	-0.20	-0.16	-	-	-0.14

Although the QM model on the small molecules gives the propensity of the reaction to take place, in the crowded enzyme environment the reactions are expected to be modified. In particular, its effect on enantioselectivity is of interest. To probe the effect of the protein electrostatic environment on the second nucleophilic attack, we have therefore performed MD calculations with ***Models 8***, ***9*** and ***10***. All of these are modeled as a continuation of ***Model 3*** which provided the most realistic model of the enzyme environment, whereby the nucleophilic attack generates a chiral center at the end. We have considered both *re* and *si* faces of enamine/carbanion while modeling this stage (Figure 5.26). ***Model 8*** represents the attack of the enamine/carbanion, **IV** to the acceptor aldehyde, **V** from the *re* face. This model does not satisfy any NAC requirements previously defined for a nucleophilic attack. Instead, prereactive interactions, which occur before the nucleophilic attack, are dentrimental for setting up the stage. We have observed a hydrogen bond between $O_{acceptor}$ and H70 in all of the conformers of ***Model 8*** (average hydrogen bond length is 1.62 ± 0.41 Å). We deduce that H70 is a catalytically important residue, since there is also a weak interaction of O_{donor} and H70 before the first nucleophilic attack for the donor aldehyde to approach. It aligns both donor and acceptor aldehydes to the correct positions for the nucleophilic attack to take place.

The S26 participates in proton transfer in ***Model 3*** via a bridge to H281 as described in Figures 5.25. A continuation of this interaction is observed in ***Model 8*** (Figure 5.28) with the difference of S26 also interacting with $O_{acceptor}$ via a hydrogen bond. In 50.9 % of the conformers produced by this model, the hydrogen bond length between $O_{acceptor}$ and water is less than 2.8 Å (average hydrogen bond length is 2.53 ± 1.63 Å). 97.4 % of the conformers have a hydrogen bond between S26 and the same water molecule (average hydrogen bond length is 1.88 ± 1.11 Å). These results explain the importance of S26 as a catalytically active residue as was also described in the literature [99]. It bridges a proton transfer in the first nucleophilic attack and is hydrogen bonded to $O_{acceptor}$ in the second.

H281 also interacts with $O_{acceptor}$ via a water molecule directing one of its protons towards $O_{acceptor}$ and its oxygen towards H281 (Figure 5.28). The first interaction (water-

$O_{acceptor}$) is present in 72.1 % of the conformers, where the average bond length was 1.76 ± 0.70 Å and the second interaction (water-H281) is present in 19.1 % of the conformers (average hydrogen bond length is 2.38 ± 0.56 Å). This interaction supports the **B2/HB4** proton transfer function of H281.

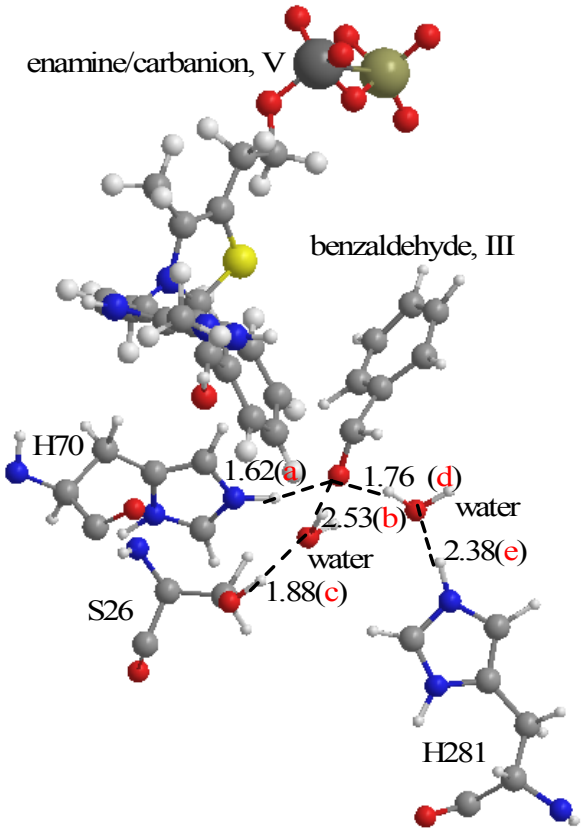


Figure 5.28. Prereactive interactions before the second nucleophilic attack (*Model 8*).

S26 and H70 also interact with $O_{acceptor}$ in *Model 9*. In this model 78.8 % of the conformers have H70-$O_{acceptor}$ interaction (average hydrogen bond length is 1.68 ± 0.39 Å) and 76.4% of them have the S26-$O_{acceptor}$ interaction (average hydrogen bond length is 1.66 ± 0.68 Å). However, H281 remains on the other side of the acceptor aldehyde and cannot interact with the $O_{acceptor}$ in any of the conformers. There are also no charged residues around $O_{acceptor}$

to transfer a proton to this atom. We therefore find that the second nucleophilic attack cannot proceed with benzaldehyde approaching from the inner side (or *si* face) of the "V" shaped ThDP. The results with ***Model 8*** and ***9*** give an insight for the *R* stereoselectivity, with benzaldehyde as the donor aldehyde.

In order to mimic the prereactive interactions at the second nucleophilic attack to acetaldehyde, we have prepared ***Model 10***, which is similar to ***Model 9.*** In ***Model 10*** all of the conformers have a hydrogen bond between H70 and $O_{acceptor}$. The latter atom is also hydrogen bonded to S26 with a bridging water. 39.9 % of the conformers have hydrogen between $O_{acceptor}$ and the water molecule, and 39.4 % of the conformers have a hydrogen bond between S26 and the same water molecule. Different from the other models, this model has two bridging water molecules combining H281 to $O_{acceptor}$. H281 always protonates $O_{acceptor}$ but, in different ways in different models. In ***Model 10***, at least 5.8 % of the conformers have the correct interaction to transfer a proton.

We have prepared a model similar to ***Model 9*** with acetaldehyde as the acceptor aldehyde depicting the approach of acetaldehyde from the inner face of "V" shaped enamine/carbanion, **IV** (*re* face). None of the conformers have acetaldehyde at an interaction distance to the nucleophilic center or hydrogen bonding distances to a catalytically active residue in this model, therefore, we have not presented the results of this model. The results of this model give an insight for the *S* stereoselectivity in BFD, when acetaldehyde is the acceptor.

5.3.3. Interaction of ThDP with the Surroundings in Model Systems

The changes in the interactions of ThDP with the surrounding amino acids provide a means by which to understand how the cofactor is stabilized in the presence (or absence) of the aldehydes that partake in the reaction. All nitrogen atoms in ThDP interact with a proton donor or acceptor amino acid and participate in forcing ThDP into its "V" shape, except N3. The interactions tabulated in Table 5.10 indicate the rigidity of the octahedral geometry

around the Mg (II) cation. In the model systems representing the interactions before and after the first nucleophilic attack the interactions with the pyrimidyl part of ThDP are conserved. The main effects of both approaching aldehydes are observed on the phosphate tail of ThDP. By the approach of the acceptor aldehyde, however, some of them are disturbed. The hydrogen bond between N1′ and the hydroxyl proton of E47 are known to play a switching function in the mechanism of the reaction [83]. In the ylide form ThDP E47 is known to transfer a proton to N1′, following which a positive charge develops on the pyrimidyl group. This strain is relaxed in the imino and ylide form of ThDP by releasing one of the N4′ protons. The proton transfer from N1′ to E47 was proposed to be responsible for the communication of two neighboring active centers in the dimer [83] and plays an important role in the acyloin formation as well as interacting with L403 to keep the "V" shape of ThDP. By the approach of the second aldehyde, this hydrogen bond length increases based on the same reasons as in case of E47. The interaction of nitrogen atoms with a proton donor or acceptor assist the localization of any charge developed or depleted on the system. N3 is located in the inner part of ThDP and stabilizes the fluctuations on the total charge of the intermediate structures. The sulfur atom, bonded to C2, is expected to act as an electron donor since it is an electron rich soft atom with a large atomic radius (Figure 5.1). In ***Model 8***, the hydrogen bond length between G401 and the N4′ proton (ca.1.96 Å) increases when compared to ***Model 3*** (ca. 4.21 Å), because of the conformational strain developed on ThDP by the approach of the second aldehyde. "V" shape of ThDP is disturbed, forced by first the approach and then the attachment of the aldehydes. Approach and attachment of the first substrate creates a double bond on the 4′-imino group, which inhibits the free rotation of this group in ***Models 1*** and ***3,*** causing the increase of hydrogen bond length between the N4′ proton and G401.

There are two types of bonding interactions at the active center: Either coordination covalent bonds or hydrogen bonds between the enzyme, ThDP and the donor aldehyde. Coordination covalent bonds anchor ThDP to the enzyme and are rigid throughout the simulations in all the models. On the other hand, the amino acid residues hydrogen bonded to ThDP open a corridor for the entrance of the substrates.

The hydrogen bond length between the side chain carboxylate group of E47 and N1′ proton is ca. 1.73 Å and the angle is 161° in ***Models 1*** and ***3***. However, by the attachment of the donor aldehyde, this bond length increases to 1.86 Å in ***Model 8*** and to 2.65 Å in ***Model 10*** indicating that the proton spends much of its time with N1′ as the substrates approach ThDP. This is a consequence of the removal of a proton of N4′ by the docked donor aldehyde, after which the pyrimidine ring rearranges itself to compensate for the developed positive charge by resonance stabilization. The length of the hydrogen bond between E47 and N1′ is also affected by the formation of the C2-C_{donor} bond. Factors affecting the nature of E47-N1′ bond also influence the other hydrogen bonding interactions between the surrounding residues and ThDP.

The hydrogen bond between the S430 side chain hydroxyl group proton and ThDP oxygen is not affected by the approach of the donor aldehyde and is firm in all of the simulations. The most dramatic change is observed in the bond distance between T377 and ThDP (Figure 5.23 and Table 5.10). By the approach of the first aldehyde (***Model 3*** versus ***Model 6*** are compared), T377 moves away from ThDP, creating a space for the entrance of the substrate.

5.4. CONCLUSIONS

In this work, we have presented a comprehensive theoretical investigation on the mechanism of nucleophilic attacks in the catalytic cycle of BFD. Different pre and postreactive QM and MD model systems are constructed in order to understand the involvement of the catalytic amino acid residues and substrates along the acyloin condensation reaction. Additionally, the contributions of certain residues are estimated from MD simulations on model systems, where they are mutated. Catalytic functions of the related residues are checked also by assigning opposing functions to the residues in different models. The NAC percentages have been used to select the most probable and realistic models.

The mode of action of ThDP in the catalytic cycle is assisted by the surrounding residues, mainly H70, H281, S26 and E47 which are the catalytically active amino acids whose presence are compulsory for ThDP function. We note that other ThDP dependent enzymes carry histidine residues in locations equivalent to that of H70 and H281 that perform similar functions as in BFD [82, 83].

In the reaction considered, there are two proton transferring species as described by ***HB1/B4*** and ***B2/HB3***. Their protonation states are restored back to their original position and the active site residues gain their original protonation states at the end of the reactions. This proton transfer progresses with respect to the active center communication patterns as described by Luisi *et al.* [83]. However, there are two competing trends on the identity of proton donor/acceptor pairs in the reaction cycle of BFD in the literature. In the first trend, the first proton donation is assigned to an unknown general acid [29]. In the second, this role is assigned to N4′ in ThDP, which is a natural continuation of the first proton removal from C2 of ThDP to generate a ylide [101-103]. Our models corroborate the roles of these proton transfer centers as follows: the 4′-amino group of ThDP functions as ***HB1*** in the first nucleophilic attack and as ***B4*** in the second. This nitrogen atom protonates O_{donor}, then the first nucleophilic attack takes place and at the end H_{donor} is removed by H281. Thus, H281

functions as ***B2*** to accept H_{donor} in the first nucleophilic attack and as ***HB3*** in the second to donate it back to $O_{acceptor}$. S26 is another catalytically active residue. It assists H281 in deprotonation of the donor aldehyde and protonation of the acceptor aldehyde. In both first and second nucleophilic attacks, H70 interacts with $O_{aldehyde}$ and aligns it towards the nucleophilic center. The results of our study indicate that H70 has an electrostatic effect on the approaching aldehyde whose absence would block the initiation of the reaction.

In this study, we have further elaborated on the reactivity and the enantioselectivity differences in the products with acetaldehyde or benzaldehyde. A QM study of the second nucleophilic attack reveals that acetaldehyde is energetically more reactive than benzaldehyde due to the high electrophilic character of its $C_{acceptor}$. Moreover, the enantioselectivity difference in the products is shown to arise from the face selectivity of the enamine, by modeling the alternative approach directions of the aldehydes using MD. Acetaldeyde can only approach from the *si* face of the enamine to produce *S*-2-HPP, whereas benzaldehyde approaches form the *re* face to produce R-benzoin. Approach of acetaldehyde from *re* face and approach of benzaldehyde from *si* face are not favored. Protein electrostatic environment therefore forces the substrates for a special alignment.

Finally, we emphasize that, although we have made significant progress in understanding the catalytic mechanism and enantioselectivity of BFD at a molecular level, it is important to point out that MD tools only describe the interactions at the atomic level, and cannot model the actual reaction step. The latter requires quantum mechanics based approaches to elucidate the electronic interactions. In particular, hybrid QM/MD approaches must be utilized to further probe the details of these reactions, where the findings of the current study provide a solid basis for the models.

6. GENERAL CONCLUSIONS

The mechanism of cyanide ion catalyzed benzoin condensation reaction from benzil was elucidated with QM tools at the B3LYP/6-31+G*//B3LYP/6-31G* level of theory. This method is known to well represent the behavior of nucleophilic addition type reactions of anionic substances. The reaction mechanism suggested by Lapworth was the basis for our calculations [12].

The effect of substituents on the rate determinig step were investigated on two benzil derivatives; 1,2-bis(2-chlorophenyl)ethane-1,2-dione and 1,2-bis(2-fluorophenyl)ethane-1,2-dione and three benzaldehyde derivatives; o-fluoro benzaldehyde, o-methyl benzaldehyde and 2-pyridinecarboxaldehyde. The product is the ester of the corresponding unsymmetrical benzoin derivative.

First, the reaction path has been modeled with PM3, and then the rate determining steps were modeled at the DFT level. The energy barriers of the rate determining steps have been considered in order to explain the experimentally observed reactivity differences. The formation and the attack of the cyanohydrin intermediate have been proposed as the rate determining steps depending on the nature on the substituents. The differences at the activation barriers of cyanohydrin formation have been used to explain the differences on the reaction barriers of substituted benzil molecules. On the other hand, the activation barriers for the nucleophilic attack of cyanohydrin have reproduced qualitatively the experimental yields of the various benzaldehyde derivatives subjected to condensation with O-benzoylated cyanohydrin. In the formation of O-benzoylated cyanohydrin intermeadiate the presence of electron withdrawing substituents at the ortho position decreases the reactivity of the species, and in the nucleophilic attack of the O-benzoylated cyanohydrin the reactivity increases with the o-fluoro substitution due to its the electrophilic character. However, in the case of the o-

methyl substituent, the reaction is comparatively less favored due to the weak electron donation.

ThDP in enzyme enviroment is another catalyst for benzoin condensation, it specifically catalyzes the enantioselective benzoin formation. There are two competing trends on the identity of proton donor/acceptor pairs in the reaction cycle of BFD in the literature. In the first trend, the first proton donation is assigned to an unknown general acid. In the second, this role is assigned to neighboring atoms of the nucleophilic carbon atom of ThDP. Several QM and MD models representing the two nucleophilic half reactions in the reaction path of the enzyme for the pre and post reactive stages of the reaction have been prepared. The attack of ylide-ThDP on a donor aldehyde -first nucleophilic half reaction- has been modeled with six different models, five of them being related to the pre-reactive intermediates and one of them to the post reactive stage of the first nucleophilic attack. In the pre-reactive models, mutations of amino acids have been considered. It has shown that H70, H281, S26 and E47 are the important amino acids assisting the catalytic function of ThDP in BFD.

The next stage in the course of the reaction was the attack of enamine/carbanion on the acceptor aldehyde. Benzaldeyde and acetaldehyde were modeled as acceptor aldehydes. In order to produce MD parameters and explain the reactivity differences between benzaldehyde and acetaldehyde, QM analysis of the molecules has been carried out. This part of the reaction has been modeled with three different models. It has been concluded that in the same protein environment acetaldehyde is more reactive towards the enamine/carbanion complex. The nucleophilic attack of enamine/carbanion to the acceptor aldehyde has been discussed with four MD models. All of the models at this stage were pre-reactive models two of them representing the attack on benzaldehyde from two different faces of the enamine/carbanion; the other two being related to acetaldehyde. The picture was different with benzaldehyde and acetaldehyde. The difference was on the attacking face of the enamine/carbanion and aldehydes. In the case of benzaldehyde, the aldehyde can only approach to enamine/carbanion from the inner part of the V shaped ThDP. The NAC percentage was very small for the other face. BFD produces (*R*)-Benzoin stereoselectively, as shown experimentally. The attack of enamine/carbanion on benzaldehyde with two different models, representing the attack on two

different faces of enamine has been realized. In the case of acetaldehyde only one model depicting the attack from the outer face for enamine/carbanion has been prepared. Based on NAC percentages, acetaldehyde has been found to be approached only from the outer part of the enamine/carbanion expected to yield (*S*)-enantiomer as experimentally suggested. Benzaldehyde approaches from the inner face of enamine/carbanion to yield (*R*) -benzoin.

7. FUTURE WORK

Different solvents used in the conventional benzoin condensation reactions are known to affect the product distributions. Polar protic solvents are known to promote the carbon-carbon ligation. Modeling the conventional benzoin condensation in dimethylsulfoxide (DMSO) would allow one to rationalize the solvent effects in these reactions.

Some of the amino acids in the reaction path of the ThDP-dependent enzymes are known to be responsible of a communication between the active centers: if a ylide forms in one active center, a second nucleophilic attack takes place at the second. This active center communication is known to be triggered by the transfer of a proton between the active centers. A research on the active center communication for ThDP-dependent enzymes with can be carried out.

In the second part of this thesis the roles of the catalytically important residues near the active center of BFD in benzoin condensation are identified however the second post-nucleophilic attack situation has not been analyzed: it can be handled with the methodologies used so far.

Benzoin condensation or more generally acyloin condensation is catalyzed also by the enzymes benzaldehyde lyase (BAL) and pyruvate decarboxylase (PDC). These two enzymes have similar catalytically active residues as BFD, they also both have ThDP at the active center. It would be interesting to model the catalytic effect of these enzymes since BAL is known to produce (*R*)-HPP and (*R*)-benzoin, in contrast to BFD. PDC is known to select short aliphatic substrates as opposed to BFD which catalyzes aromatic substrates, even though both of them can decarboxylase to benzoylformate. The challenge would be to detect these preferences with hybrid computational tools.

Finally, a detailed study can be conducted with QM/MM tools on the two nucleophilic half reactions taking place in the enzyme environment in order to elucidate better the functions of the surrounding residues.

REFERENCES

1. Demir, A. S., Pohl, M.,Janzen, E. and M. Muller, "Enantioselective Synthesis of Hydroxy Ketones Through Cleavage and Formation of Acyloin Linkage. Enzymatic Kinetic Resolution via C-C Bond Cleavage ", *Journal of the Chemical Society, Perkin Transactions 1*, pp. 633-635, 2001.

2. Pirrung, M. C. and S. W. Shuey, "Photoremovable Protecting Groups for Phosphorylation of Chiral Alcohols. Asymmetric Synthesis of Phosphotriesters of (-)-3',5'-Dimethoxybenzoin", *Journal Organic Chemistry*., Vol. 59, pp. 3890-3897, 1994.

3. Rock, R. S. and S. I. Chan, "Preparation of a Water-Soluble "Cage" Based on 3',5'-Dimethoxybenzoin", *Journal of the American Chemical Society*, Vol. 120, pp. 10766-10767, 1998.

4. Papageorgiou, G. and J. E. T. Corrie, "Synthesis and Properties of Carbamoyl Derivatives of Photolabile Benzoins", *Tetrahedron*, Vol. 53, pp. 3917-3932, 1997.

5. Pappas, S. P., *UV Curing*: *Science and Technology*, Technology Marketing Corp., Norwalk, 1978.

6. Wöhler, F. and J. Liebig, *Ann Pharm*, Vol. 3, pp.249-282, 1832.

7. Adam, W., Lazarus, M., Saha-Möller, C. R. and Schreier, P., "Biocatalytic Synthesis of Optically Active α-Oxyfunctionalized Carbonyl Compounds", *Accounts of Chemical Research.*, Vol. 32, pp. 837-845, 1999.

8. Wallace, O. B., Smith D. W., Deshpande, M. S., Polson, C. and Felsenstein, K. M., "Inhibitors of Aβ production: Solid-Phase Synthesis and SAR of α-hydroxycarbonyl

Derivatives", *Bioorganic and Medicinal Chemistry Letters*, Vol. 13, pp. 1203–1206. 2003.

9. Fang, Q. K., Han, Z., Grover, P., Kessler, D., Senanayake, C. H. and S. A., Wald, "Rapid Access to Enantiopure Bupropion and Its Major Metabolite by Stereospecific Nucleophilic Substitution on an α-ketotriflate", *Tetrahedron:Asymmetry*, Vol. 11, pp. 3659-3663, 2000.

10. a) Müller, C. E., Diekmann, M., Thorand, M. and V., Ozola, "[3H]8-Ethyl-4-methyl-2-phenyl-(8*R*)-4,5,7,8-tetrahydro-1H-imidazo[2,1-i]-purin-5-one ([3H]PSB-11), a Novel High-Affinity Antagonist Radioligand for Human A3 Adenosine Receptors", *Bioorganic and Medicinal Chemistry*, Vol. 12*,* pp. 501-503, 2002. b) Adam, W., Mueller, M. and F. Prechtl "Dimethyldioxirane Oxidation of Titanium Enolates: Diastereoselective .alpha.-Hydroxylations", *Journal of Organic Chemistry,* Vol. 59, pp. 2358-2364, 1994.

11. Corrie, J.E.T. and Trentham, D. R., "Synthetic, Mechanistic and Photochemical Studies of Phosphate Esters of Substituted Benzoins", *Journal of Chemical Society, Perkin Transactions. 1*, pp. 2409-2417, 1992.

12. Lapworth, A.., "XCVI.—Reactions Involving the Addition of Hydrogen Cyanide to Carbon Compounds", *Journal of the Chemical Society*, Vol. 83, pp. 995-1005, 1903.

13. Kuebrich, J. P., Schowen, R. L., Wang, M. and M. E. Lupes,"Mechanism of the Benzoin Condensation", *Journal of the American Chemical Society*, Vol. 93, pp. 1214-1220, 1971.

14. Kwart, H., and M. Baevsky, "The Cyanide Ion Catalyzed Cleavage of Aromatic α-Diketones", *Journal of the American Chemical Society,* Vol. 80, pp. 580-588, 1958.

15. Staudinger, H., *Berichte der Deutschen Chemischen Gesellschaft*, Vol. 46, pp. 3535-3538, 1913.

16. Kuebrich, J. P. and R. L., Schowen, "Chemistry of the Active-Aldehyde Intermediate", *Journal of the American Chemical Society*, Vol. 93, pp. 1220-1223, 1971.

17. a) Stewart, T. D. and B. J. Fontana, "The Effect of Solvation upon the Dissociation of Acetone Cyanohydrin" *Journal of the American Chemical Society*, Vol. 62, pp. 3282-3285, 1940. b) Svirbely, W. J. and J. F. Roth, "Carbonyl Reactions, "I. The Kinetics of Cyanohydrin Formation in Aqueous Solution", *Journal of the American Chemical Society*, Vol. 75, pp. 3106-3111, 1953.

18. a) Demir, A. S., Sesenoglu, O., Eren, E., Hosrik, B., Pohl, M., Jansen, E., Kolter, D., Feldmann, R., Dunkelmann, P.and M. Muller, "Enantioselective Synthesis of alpha-Hydroxy Ketones via Benzaldehyde Lyase-Catalyzed C-C Bond Formation Reaction ", *Advanced Synthesis and Catalysis*, Vol. 344, pp. 96-103, 2002. b) Dunkelmann, P., Kolter-jung, D., Nitsche, A., Demir, A. S., Siegert, P., Lingen, B., Baumann, M., Pohl, M. and M. Muller, "Development of a Donor-Acceptor Concept for Enzymatic Cross-Coupling Reactions of Aldehydes: The First Asymmetric Cross-Benzoin Condensation", *Journal of the American Chemical*, Vol. 124, pp. 12084-12085, 2002.

19. Clerici, A., and O. Porta, "Reductive coupling of benzoyl cyanide and carbonyl compounds by aqueous titanium (III) ions. A New Convenient and Selective Access to the Less Stable Mixed Benzoins", *Journal of Organic Chemistry*, Vol. 58, pp. 2889-2893, 1993.

20. a) Ugai, T. and S. Tanaka, *Journal of Pharmaceutical Society of Japan*, Vol. 63, pp. 296-300, 1943, b) Breslow, R. "On the Mechanism of Thiamine Action .4. Evidence From Studies on Model Systems", *Journal of the American Chemical*,Vol. 80, pp. 3719-3726, 1958, c) Breslow, R. and E. Kool, "Dichotomous Salt Effects in the Hydrodhobic Acceleration of the Benzoin Condensation", *Journal of the American*

Chemical, Vol. 110, pp. 1596-1597, 1988, d) Castells, J., Lopez-Calahorra, F. and L. Domingo, "Postulation of Bis(Thiazolin-2-Ylidene)s as The Catalytic Species in the Benzoin Condensation Catalyzed by a Thiazolium Salt Plus Base", *Journal of Organic Chemistry*, Vol. 53, pp. 4433-4436, 1988. e) Diederich, F. and H. Lutter, "Catalytic Cyclophanes .4. Supramolecular Catalysis of Benzoin Condensations by a Thiazolium Cyclophane", *Journal of the American Chemical*, Vol. 111, pp. 8438-8446, 1989.

21. Corrie, J. E. T. "Preparation and properties of unsymmetrical benzoins and related compounds"**,** *Tetrahedron*, Vol. 54. pp. 5407-5416, 1998.

22. Demir, A. S., Dunnwald, T., Iding, H., Pohl, M. and M. Muller, *Tetrahedron:Asymmetry*, Vol. 10, pp. 4769-4774, 1999 and the references therein.

23. a) Tachibana, Y, Kihara, N., Takata, T. "Asymmetric Benzoin Condensation Catalyzed by Chiral Rotaxanes Tethering a Thiazolium Salt Moiety via the Cooperation of the Component: Can Rotaxane be an Effective Reaction Field? ", *Journal of the American Chemical*, Vol. 126, pp. 3438-3439, 2004. b) Linghu, X., Potnick, J. R. and Johnson, J. S. "Metallophosphites as Umpolung Catalysts: the Enantioselective Cross Silyl Benzoin Reaction"*, Journal of the American Chemical*, Vol. 126, pp. 3070-3071, 2004, c) Dijkstra, G.D.H., Kellogg, R. M., Wynberg, H., Svendsen, I. M. and K. B. Sharpless, "Conformal study of Cinchona Alkaloids. A Combined NMR, Molecular Mechanics and X-Ray Approach", *Journal of the American Chemical*, Vol. 111, pp. 8069-8076, 1989.

24. Jordan, F., "Current Mechanistic Understanding of ThDP-Dependent Enzymatic Reactions", *National. Productivity. Reports*,.Vol. 20, pp. 184-201. 2003.

25. Mosbacher, T. G., Mueller, M. and Schulz, G., E. "Structure and mechanism of the ThDP-dependent benzaldehyde lyase from *Pseudomonas fluorescens*" *FEBS Journal* Vol. 272, pp. 6067–6076, 2005.

26. Dakin, H. D. and C. R. Harrington, "The Action of Ammonium Cyanide on Diketones", *Journal of Biological Chemistry*, Vol. 55, pp. 487-494, 1923.

27. Trisler, J. C. and J. L. Frye, "Reaction of Benzine with Cyanide in Dimethyl Sulfoxide", *Journal of Organic Chemistry*, Vol. 30, pp. 306-307, 1965.

28. Demir, A. S. and O. Reis, "A Convenient and Selective Synthesis of Unsymmetrical Benzoins via the Cyanide Ion Catalyzed Cleavage of Benzils", *Tetrahedron*, Vol. 60, pp. 3803-3811, 2004.

29. Kluger, R. "Thiamin Diphosphate: a Mechanistic Update on Enzymic and Nonenzymic Catalysis of Decarboxylation", *Chemical Reviews*, Vol. 87, pp. 863-876, 1987.

30. Kluger, R. "Lessons from Thiamin Watching", *Pure and Applied Chemistry*, Vol. 69, pp. 1957-1967, 1997.

31. Breslow, R. "On the mechanism of thiamine action: 4. Evidence from studies on model Systems", *Journal of the American Chemical*, Vol. 80, pp. 3719-3726, 1958.

32. Wilcocks, R., Ward, O. P., Collins, S., Dewdney, N. J., Hong, Y. and Prosen, E. "Acyloin Formation by Benzoylformate Decarboxylase from *Pseudomonas putida*", *Applied and Environment Microbiology*, Vol. 58, pp. 1699-1704, 1992.

33. Hasson, M. S., A. Muscate, M. J. McLeish, L. S. Polovnikova, J. A.Gerlt, G. L. Kenyon, G. A. Petsko, and D. Ringe, "The Crystal Structure of Benzoylformate Decarboxylase at 1.6 Å Resolution: Diversity of Catalytic Residues in Thiamin diphosphate-dependent Enzymes", *Biochemistry*, Vol. 37, pp. 9918-9930, 1998.

34. Tsou, A. Y., S. C. Ransom, J. A. Gerlt, D. D. Buechter, P. C. Babbitt, and G. L Kenyon, "Mandelate Pathway of *Pseudomonas putida*: Sequence Relationships Involving Mandelate Racemase, (*S*)-Mandelate Dehydrogenase, and Benzoylformate

Decarboxylase and Expression Of Benzoylformate Decarboxylase in *Escherichia coli*.", *Biochemistry*, Vol. 29, pp. 9856-9862, 1990.

35. Polovnikova, E.S., McLeish, M.J., Sergienko, E.A., Burgner, J.T., Anderson, N.L., Bera, A.K., Jordan, F., Kenyon, G.L. and Hasson, M.S., "Structural and Kinetic Analysis of Catalysis by a Thiamin Diphosphate-Dependent Enzyme, Benzoylformate Decarboxylase", *Biochemistry*, Vol. 42, pp. 1820-1830, 2003.

36. a) Dunnwald, T., Demir, A.S., Siegert, P., Pohl M. and Muller, M. "Enantioselective Synthesis of (*S*)-2-hydroxypropanone Derivatives by Benzoylformate Decarboxylase Catalyzed C-C Bond Formation", *European Journal of Organic Chemistry*, Vol. 11, pp. 2161-2170, 2000. b) Demir, A. S., Dünnwald, T., Iding, H., Pohl M. and Müller M., "Asymmetric Benzoin Reaction Catalyzed by Benzoylformate Decarboxylase", *Tetrahedron:Asymmetry*, Vol. 10, pp. 4769-4774, 1999. c) Dunkelmann, P., Kolter-Jung, D., Nitsche, A., Demir, A., S., Siegert, P., Lingen, B., Baumann, M., Pohl, M. and Muller M. "Development of a Donor-Acceptor Concept For Enzymatic Cross-Coupling Reactions of Aldehydes: The First Asymmetric Cross-Benzoin Condensation", *Journal of the American Chemical Society*, Vol. 124, pp. 12084-12085, 2002. d) Dunnwald, T. and Muller, M. "Stereoselective Formation of Bis(Alpha-Hydroxy Ketones) via Enzymatic Carboligation", *Journal of Organic Chemistry*, Vol. 65, pp. 8608-8612, 2000.

37. Hegeman, G.D. "Benzoylformate Decarboxylase (*Pseudomonas putida*)", Methods in Enzymology, Vol. 17, pp. 674-678, 1970.

38. Barrowman, M. M., W. A. Harnett, J. Scott, A., Fewson, and Kusel, J.R.C. "Immunological Comparison of Microbial TPP-Dependent Nonoxidative α-Keto Acid Decarboxylases", *FEMS Microbiology Letters.*, Vol. 34, pp. 57-60, 1986.

39. Barrowman, M. M. and Fewson, C. A. “Phenylglyoxylate Decarboxylase and Phenylpyruvate Decarboxylase from *Acinetobacter-Calcoaceticus”*, *Current Microbiology*, Vol. 12, pp. 235-239, 1985.

40. Muller, Y. A, and G. E. Schulz, “Structure of the Thiamine- and Flavin-Dependent Enzyme Pyruvate Oxidase”, *Science*, Vol. 25, pp. 965-967, 1993.

41. D. Dobritzsch, S. König, G. Schneider and G. G. Lu, “High Resolution Crystal Structure of Pyruvate Decarboxylase from Zymomonas mobilis”, *Journal of Biological Chemistry*, Vol. 273, pp. 20196-20204, 1998.

42. a) Schrodinger, E., “Quantisierung als Eigenwertproblem (Erste Mitteilung)”, *Annalen der Physik*, Vol. 79, pp. 361-376, 1926. b) Schrodinger, E., “Quantisierung als Eigenwertproblem (Zweite Mitteilung)”, *Annalen der Physik*, Vol. 79, pp. 489-527, 1926. c) Schrodinger, E., “Quantisierung als Eigenwertproblem (Dritte Mitteilung. Storungstheorie mit Anwendung auf den Starkeekt der Balmerlinien) ”, *Annalen der Physik*, Vol. 80, pp. 437-490, 1926. d) Schrodinger, E., “Quantisierung als Eigenwertproblem (Vierte Mitteilung)”, *Annalen der Physik*, Vol. 81, pp. 109-139, 1926. e) Schrodinger, E., “Der stetige Ubergang von der Mikro- zur Makromechanik”, *Die Naturwissenschaften*, Vol. 14, pp. 664-666, 1926.

43. M. Born and R. Oppenheimer, “Zur Quantentheorie der Molekeln”, *Annals of Physics*, Vol. 84, pp. 457-483, 1927.

44. Pople, J. A., Santry, D. P. and D. A. Segal, “Approximate Self-Consistent Molecular Orbital Theory. I. Invariant Procedures”, *Journal of Chemical Physics*,Vol. 43, pp. 129-135, 1965.

45. Dewar, M. J. S. and W. Thiel, “The MNDO Method. Approximations and Parameters”, *Journal of the American Chemical Society*, Vol. 99, pp. 4899-4907, 1977,

46. Dewar, M. J. S., Zoebisch, E. G. Healy, E. F. and J. J. P. Stewart, "AM1: A New General Purpose Quantum Mechanical Molecular Model", *Journal of the American Chemical Society*, Vol. 107, pp. 3902-3909, 1985.

47. Stewart, J. J. P., "Optimization of Parameters For Semi-Empirical Methods I", *Journal.of Computational Chemistry*., Vol. 10, pp. 209-220, 1989.

48. Handy, N. C., *Density Functional Theory*, *in: B. O. Roos (ed.), Lecture Notes in Quantum Chemistry*, Vol. 2, pp. 91-123, Springer-Verlag, Berlin, 1994

49. Parr, R., Yang, W. *Density Functional Theory of Atoms and Molecules*, Oxford, New York 1989.

50. Andzelm, J. and E. Wimmer, "Density Functional Gaussian-Type-Orbital Approach to Molecular Geometries, Vibrations, and Reaction Energies", *Journal of Chemical Physics*. Vol. 96, pp.1280-1303, 1992.

51. Dickson, R.M., and A. D. Becke, "Basis-Set-Free Local Density Functional Calculations of Geometries of Polyatomic Molecules", *Journal of Chemical Physics*, Vol. 99*,* pp. 3898-3905, 1993.

52. Johnson B.G., Gill P.M.W. and J. A. Pople, "The Performance Of A Family Of Density Functional Methods", *Journal of Chemical Physics*, Vol. 98, pp. 5612-5626, 1993.

53. Zhou, X,. Wheelees, C. J. M. and R. Liu, "Density Functional Theory Study of Vibrational Spectra. 1. Performance of Several Density Functional Methods in Predicting Vibrational Frequencies"*, Vibrational Spectroscopy*, Vol. 12, pp. 53–63, 1996.

54. Miehlich, B., Savin, A., Stoll, H., and H. Preuss, "Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr", *Chemical Physics Letters*, Vol. 157, pp. 200-206, 1989.

55. Lee, C., W. Yang and R. G. Parr, "Development of Colle-Salvetti Correlation Energy Formula into a Functional of the Electron Density", *Physical Review B: Condensed Matter*, Vol. 37, pp. 785-789, 1988.

56. a) Becke, A. D., "Density Functional Thermochemistry. III. The Role of Exact Exchange", *Journal of Chemical Physics*, Vol. 98, pp. 5648-5652, 1993. b) Stephens, P. J., Devlin, F. J., Chabalowski, C. F. and M. J. Frisch, *Journal of Physical Chemistry*, Vol. 98, pp. 11623-11627, 1994

57. J. B. Foresman and A. Frisch., *Exploring Chemistry with Electronic Structure Methods,* Gaussian, Inc. Pittsburgh, PA. 1996

58. F. Jensen., *Introduction to Computational Chemistry,* John Wiley & Sons, Ltd. England. 1999.

59. Foresman, J.B., Keith, T.A., Wieberg, K.B., Snoonian, J. and M. J. Frisch, "Solvent Effects. 5. Influence of Cavity Shape, Truncation of Electrostatics, and Electron Correlation on Ab Initio Reaction Field Calculations", *Journal of Physical Chemistry*, Vol. 100, pp. 16098-16104, 1996.

60. Parr, R. G. and W. Yang, "Density-Functional Theory of the Electronic Structure of Molecules", *Annual Review of Physical Chemistry*, Vol. 46, pp. 701-728, 1995.

61. a) Chermette, H., "Chemical Reactivity Indexes In Density Functional Theory", *Journal of Computational Chemistry*, Vol. 20, pp. 129-154, 1999. b) De Proft, F. and

P. Geerlings, "Conceptual and Computational DFT in the Study of Aromaticity", *Chemical Reviews*, Vol. 101, pp. 1451-1464, 2001.

62. Geerlings, P., De Proft, F. and W. Langenaeker, "Conceptual Density Functional Theory", *Chemical Reviews*, Vol. 103, pp. 1793-1873, 2003.

63. De Proft, F., Van Alsenoy, C., Peeters, A., Langenaeker, W. and P. Geerlings, "Atomic Charges, Dipole Moments, and Fukui Functions Using the Hirshfeld Partitioning of the Electron Density", *Journal of Computational Chemistry*, Vol. 23, pp. 1198-1209, 2002.

64. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, John Wiley & Sons Ltd., USA, 2004.

65. Leach, A. R., *Molecular Modelling Principles and Applications,* Prentice Hall, England, 2001.

66. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta Jr., S. and P. Weiner, "A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins", *Journal of the American Chemical Society*, Vol. 106, pp. 765-784, 1984.

67. Bernard R. Brooks, Robert E. Bruccoleri, Barry D. Olafson, David J. States, Swaminathan, S. and M. Karplus, "CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations", *Journal of Computational Chemistry*, Vol. 4, pp. 187-217, 1983.

68. van Gunsteren W. F., Billeter S.R., Eising A.A., Hünenberger, P.H., Krüger P., Mark A.E., Scott, W.R.P. and Tironi I.G, *Biomolecular Simulation: The GROMOS96*

Manual and Userguide, vdf Hochschulverlag AG an der ETH Zurich, Groningen, 1996.

69. Rapaport, D. C., *The Art of Molecular Dynamics Simulation*. 1995, Cambridge: University Press.

70. Gunsteren, W. and H. J. C. Berendsen, "Algorithms for Macromolecular Dynamics and Constraint Dynamics", *Molecular Physics*,Vol. 34, pp. 1311-1327, 1977.

71. Andersen, H. C., "Rattle - A Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations", *Journal of Computational Physics*, Vol. 52, pp. 24-34, 1983.

72. Allen, M. P. and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford: Clarendon Press, 1987.

73. York, D. M., Darden, T. A. and Pedersen, L. G. "The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals - A Comparison of the Ewald and Truncated List Methods", *Journal of Chemical Physics*,.Vol. 99, pp. 8345-8348, 1993.

74. a) Cheatham III, T.E., Miller, J.L., Fox, T., Darden, T.A. and P.A Kollman, "Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins", *Journal of the American Chemical Society*, Vol. 117, pp. 4193-4194, 1995. b) Louise-May, S., Auffinger, P., Westhof, E. "Calculations of Nucleic Acid Conformations**"**, *Current Opinion in Structural Biology* , Vol. 6*,* pp. 289-298, 1996.

75. Topal, K. G., Unaleroglu, C. and V. Aviyente , "Computational Study of The Synthesis of Benzoin Derivatives from Benzil", *International Journal of Quantum Chemistry*, Vol. 106, pp. 1596-1610, 2006.

76. *Spartan*, version 5.1.1. Wavefunction, Inc., Irvine, CA, 2002.

77. Koch, W., Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, Wiley-VCH: Weinheim, Germany, 2001.

78. Gaussian 98, Revision A.7, Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, Jr., J. A., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson,G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Baboul, A. G., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., B. Johnson,, Chen, W., Wong, M. W., Andres, J. L., Gonzalez, C., Head-Gordon, M., Replogle, E. S., Pople, J. A. Gaussian, Inc., Pittsburgh PA, 1998.

79. Gonzalez, C. and H. B. Schlegel, "Ab Initio Study of the Addition Reaction of the Methyl Radical to Ethylene and Formaldehyde", *Journal of Chemical Physics*, Vol. 90, pp. 2435-2440, 1989.

80. Reis, O. Private communications.

81. Hammond, G. S. "A Correlation of Reaction Rates", *Journal of the American Chemical Society*, Vol. 77, pp. 334-338, 1955.

82. Schutz, A., Golbik, R., Konig, S., Hubner, G. and Tittmann, K., "Intermediates and Transition States in Thiamin Diphosphate-Dependent Decarboxylases. A Kinetic and NMR Study on Wild-Type Indolepyruvate Decarboxylase and Variants Using

Indolepyruvate, Benzoylformate, and Pyruvate as Substrates", *Biochemistry*, Vol. 44, pp. 6164-6179, 2005.

83. Frank, R. A. W., Leeper, F. J. and Luisi, B. F., "Structure, Mechanism and Catalytic Duality of Thiamine-Dependent Enzymes, *Cellular and Molecular Life Sciences*, Vol. 64, pp. 892-905, 2007.

84. Nemeria, N., Baykal, A., Joseph, E., Zhang, S., Yan, Y., Furey, W. and Jordan, F. "Molecular Dynamics Simulations on the Coenzyme Thiamin Diphosphate in Apoenzyme Environment", *Biochemistry*, Vol. 43, pp. 6565-6575, 2004.

85. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L. and Schulten, K. "Scalable Molecular Dynamics with NAMD", *Journal of Computational. Chemistry,* Vol. 26, pp. 1781-1802, 2005.

86. a) MacKerell, A. D., Jr., et al. "All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins", *Journal of Physical Chemistry B*, Vol. 102, pp. 3586-3616, 1998. b) Pavelites, J. J., Gao, J., Bash, P. A. and A. D. M., Jr. Mackerell, "A Molecular Mechanics Force Field for NAD+ NADH, and The Pyrophosphate Groups of Nucleotides", *Journal of Computational. Chemistry*, Vol. 18, pp. 221-239, 1997.

87. a) Foloppe, N. and A. D., MacKerell, Jr. "All-atom Empirical Force-Field for Nucleic Acids. 1) Parameter Optimization Based On Small Molecule and Condensed Phase Macromolecular Target Data", *Journal of Computational. Chemistry*, Vol.21, pp. 86-104, 2000. b) MacKerell, A. D., Jr., and N. Banavali, "Protein-Facilitated Base Flipping in DNA by cytosine-5-methyltransferase", *Journal of Computational. Chemistry*, Vol. 21, pp. 105-120, 2000.

88. Verlet, L., "Computer "Experiment" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules", *Physical Review.*, Vol. 159, pp. 98-103, 1967.

89. Humphrey, W., Dalke, A. and K., Schulten, "VMD: Visual Molecular Dynamics", *Journal of Molecular Graphics,* Vol. 14, pp. 33-38, 1996.

90. Lie, M. A., Celik, L., Jørgensen, K. A. and B Schiøtt, "Cofactor Activation and Substrate Binding in Pyruvate Decarboxylase. Insights into the Reaction Mechanism from Molecular Dynamics Simulations", *Biochemistry*, Vol. 44, pp. 14792-14806, 2005.

91. Gilson, M.K., Gilson, H.S.R. and M. J. Potter, "Fast Assignment of Accurate Partial Atomic Charges. an Electronegativity Equalization Method that Accounts for Alternate Resonance Forms", *Journal of Chemical Information and Computer Science*, Vol. 43, pp. 1982-1997, 2003.

92. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impley, R. W. and Klein, M.-L., "Comparison of Simple Potential Functions for Simulating Liquid Water", *Journal of Chemical Physics*, Vol. 79, pp. 925-935, 1983.

93. Li, H., Robertson, A. D. and, J. H Jensen, "Very Fast Empirical Prediction and Interpretation of Protein pKa Values", *Proteins*, Vol. 61, pp. 704-721, 2005.

94. Bas, D.C., Rogers, D. M. and J. H. Jensen, "Very Fast Prediction and Rationalization of pKa Values for Protein-Ligand Complexes", *Proteins*, in press. 2008.

95. a) Luo, J., and T. C. Bruice, "Dynamic Structures of Horse Liver Alcohol Dehydrogenase (HLADH): Results of Molecular Dynamics Simulations of HLADH-NAD+-PhCH2OH, HLADHNAD+-PhCH2O-, and HLADH-NADH-PhCHO", *Journal of the American Chemical Society*, Vol. 123, pp. 11952-11959, 2001. b) Valiňa, A. L. B., Mazumder-Shivakumar, D., and T. C. Bruice, "Probing the Ser-Ser-Lys Catalytic Triad Mechanism of Peptide Amidase: Computational Studies of the Ground State, Transition State, and Intermediate", *Biochemistry*, Vol. 43, pp. 15657-15672, 2004. c) Schiøtt, B., and Bruice, T. C., "Reaction Mechanism of Soluble Epoxide Hydrolase:

Insights from Molecular Dynamics Simulations", *Journal of the American Chemical Society*, Vol. 124, pp. 14558-14570, 2002.

96. a) Burgi, H. B., Dunitz, J. D. and E. Shefter, "Geometrical Reaction Coordinates. II. Nucleophilic Addition to a Carbonyl Group", *Journal of the American Chemical Society*, Vol. 95, pp. 5065-5067, 1973. b) Burgi, H. B., Lehn, J. M. and G. Wipff, "Ab Initio Study of Nucleophilic Addition to a Carbonyl Group", *Journal of the American Chemical Society,* Vol. 96, pp. 1956-1957, 1974.

97. Wang, J., Dong, H., Li, S. and H. He, "Theoretical Study toward Understanding the Catalytic Mechanism of Pyruvate Decarboxylase", *Journal of Physical Chemistry B*, Vol. 109, pp. 18664-18672, 2005.

98. a) Jordan, F., Zhang, Z. and E. Sergienko, "Spectroscopic Evidence for Participation of the 1',4'-imino Tautomer of Thiamin Diphosphate in Catalysis by Yeast Pyruvate Decarboxylase", *Bioorganic Chemistry*, Vol. 30, pp.188-198, 2002. b) Jordan, F., Nemeria, N. S., Zhang, S., Yan, Y., Arjunan, P. and W. Furey, "Dual Catalytic Apparatus of the Thiamin Diphosphate Coenzyme: Acid-Base via the 1',4'-iminopyrimidine Tautomer Along with its Electrophilic Role", *Journal of the American Chemical Society*, Vol. 125, pp. 12732-12738, 2003.

99. a) Nabedryk, E., Paddock, M. L., Okamura, M. Y. and J. Breton, "An Isotope-Edited FTIR Investigation of the Role of Ser-L223 in Binding Quinone (QB) and Semiquinone (QB-) in the Reaction Center from Rhodobacter sphaeroides", *Biochemistry*, Vol. 44, pp. 14519-14527, 2005. b) Fujimoto, K, Ichinose, H., Nagatsu, T., Nonaka, T., Mitsui, Y. and S. Katoh, "Functionally Important Residues Tyrosine-171 and Serine-158 in Sepiapterin Reductase", *Biochimica et Biophysica Acta*, Vol.1431, pp. 306-314, 1999.

100. Iding, H., Dünnwald, T., Greiner, L., Liese, A., Müller, M., Siegert, P., Grötzinger, J., Demir, A.S. and M. Pohl, "Benzoylformate Decarboxylase from Pseudomonas Putida

as Stable Catalyst for the Synthesis of Chiral 2-hydroxy Ketones", *Chemistry A European Journal*, Vol. 6*,* pp. 1483-1496, 2000.

101. Halpin, R. A., Hegeman, G. D., and G. L. Kenyon, "Carbon-13 Nuclear Magnetic Resonance Studies of Mandelate Metabolism in Whole Bacterial Cells and in Isolated in vivo Cross-Linked Enzyme Complexes", *Biochemistry*, Vol. 20, 1525-1533, 1981.

102. Neidhart, D.J., Howell, P.L., Petsko, G.A., Powers, V.M., Li, R., Kenyon, G.L. and J.A. Gerlt, "Mechanism of the Reaction Catalyzed by Mandelate Racemase: Crystal Structure of Mandelate Racemase at 2.5 Å Resolution, Identification of the Active Site and Possible Catalytic Residues", *Biochemistry*, Vol. 30, pp. 9264-9273, 1991.

103. Wille, G., Meyer, D., Steinmetz, A., Hinze, E., Golbik, R. and K.Tittmann, "The Catalytic Cycle of a Thiamin Diphosphate Enzyme Examined by Cryocrystallography", *Nature Chemical Biology* Vol. 2, pp. 324-328, 2006.