EXPLORING THE MODE OF ACTION OF MULTIFUNCTIONAL ORGANOCATALYSTS

by
Sezen Alsancak

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in
Chemical Engineering

Yeditepe University
2023

EXPLORING THE MODE OF ACTION OF MULTIFUNCTIONAL ORGANOCATALYSTS

APPROVED BY:

Prof. Dr. Nihan Çelebi ……
(Thesis Supervisor)
(Yeditepe University)

Prof. Dr. Tuğba Davran Candan ……
(Boğaziçi University)

Assist. Prof. Dr. Cem Levent Altan ……
(Yeditepe University)

Prof. Dr. Viktorya Aviyente ……
(Boğaziçi University)

Prof. Dr. Safiye Erdem ……
(Marmara University)

DATE OF APPROVAL: ..../..../20..

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. If any material used in the thesis requires copyright, the necessary permissions have been obtained. No material from this thesis has been used for the award of another degree.

I accept all kinds of legal liability that may arise in cases contrary to these situations.

Name, Last name Sezen Alsancak

Signature …..…….……..

ACKNOWLEDGEMENTS

I want to express my sincere appreciation to my dear advisor Prof. Nihan Çelebi for her great courage, endless motivation, immense support, continuous help, understanding, and advice. Throughout my educational and academic life, she kept inspiring me in all ways, and it has been a great pleasure for me to be her student as always. She led me with her valuable advice throughout all stages of my academic journey. I also appreciate her inspiring and encouraging me to finish this Ph.D. thesis. I met her just six months before the end of my undergraduate education, and even though I never thought about starting a Ph.D., she encouraged me to pursue it. She gave me the opportunity to work on her TUBITAK project and enabled me to start my Ph.D., which changed the direction of my life. I am so glad that I met her.

I would like to thank Assist. Prof. Semin Funda Oğuz, for allowing me to work on the TUBITAK project and for opening new doors in my career. I am very grateful to her for her endless motivation, guidance, kindness, and friendship.

While my name is the only one visible on the dissertation cover, I want to thank my friend Yeşim Çamlısoy who has contributed significantly to its creation. She is one of my best friends and study partners during my undergraduate and doctoral life, which spanned 14 years. She is also my officemate who made my teaching assistant responsibility more manageable and pleasant. I want to thank Yeşim, especially for her love, support, and guidance that saved me from being too meticulous throughout these difficult years; this dissertation is simply impossible without them. I enjoyed working with her, and she was a valuable collaborator. She is my biggest academic and personal supporter, and I know she is always by my side in every aspect. I am so lucky that we grew up together and will always be side by side. She is one of the people I am grateful to have met in my life.

I wish to appreciate the examining committee members, Prof. Viktorya Aviyente, Prof. Tuğba Davran Candan, Prof. Safiye Erdem, and Assist. Prof. Cem Levent Altan for their supporting remarks and valuable advice.

I want to thank all the Department of Chemical Engineering members at Yeditepe University for sharing their wisdom and various forms of support throughout these years. I want to give

special thanks to Ezgi Baran İlhan for her friendship, encouragement, and support. She always made life easier in the department.

I want to thank TUBITAK (Project No. 114Z791) for the financial support. Although not directly part of the TUBITAK project proposal, the studies included in this thesis were presented in the project report as outputs of the proposed protocol in the project. The numerical calculations reported in this paper were fully performed at TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources).

I also want to thank Umay Topçubaşı, Tuğçe Ütnier, Kaan Öztürk, and Ece Atlan for their endless support, and encouragement and for sharing the most wonderful times at Yeditepe University.

I am thankful to my lifelong friends Candan Aksoy and Ece Kuralay for their support and motivation, and for sharing the most joyful times of my life. I feel fortunate to have friendships with great people who were there whenever I needed them.

Last but not least, I want to express my sincere gratitude to my family for their endless love and patient support throughout my academic life. I am always grateful to my sister Ece Alsancak, my mother, Ülkiye Alsancak, and my father, Yüksel Alsancak. I am very appreciative of my husband, Oğulcan Koca. Without his support, it would have been a more challenging journey. He has been by my side, my biggest supporter since the beginning of my academic career. He has supported me in all situations and given me the courage to advance in my academic career. I also have a wonderful daughter. My dear girl, Pelin Koca (PeKo), you will forever be the source of light and happiness in my life, and I believe that with your support, I can overcome any difficulty.

ABSTRACT

EXPLORING THE MODE OF ACTION OF MULTIFUNCTIONAL ORGANOCATALYSTS

Quantum mechanical calculations have successfully uncovered organocatalytic reaction mechanisms and explained their outcomes. However, understanding the mode of action of complex multifunctional organocatalysts remains challenging due to their numerous conformational degrees of freedom. This work presents a computational protocol called CIPOC (Computational Identification of Potential (Organo)Catalysts) that combines quantum mechanical calculations and drug design tools to address the challenges of exploring the mode of action and active conformations of complex multifunctional organocatalysts. CIPOC was applied (1) to discover novel chiral catalytic structures for the catalysis of the asymmetric conjugate addition reaction of dimethyl malonate to *trans-(β)*-nitrostyrene, which is particularly noteworthy to access gamma-aminobutyric acid (GABA) close analogs, (2) to uncover the mode of action of dimeric quinidine in the reaction of oxindole with nitrosobenzene that yielded unexpectedly aminooxylation products. We discovered a novel 2-aminoDMAP/urea catalyst, which catalyzed the reaction of dimethyl malonate to *trans-(β)*-nitrostyrene eight-fold faster than Takemoto's catalyst, previously the best-known catalyst for this transformation. We showed that quinidine dimer binds nitrosobenzene through its protonated amine, while both hydroxyl groups in the catalyst structure interact with the substrates and orient the enolate resulting in a concerted addition-proton transfer pathway. Additionally, both studies revealed the importance of catalyst distortion in determining the catalytic activity, highlighting the extraordinary performance of urea-based catalysts compared to their thiourea analogs, despite their lower acidity. The impact of catalysts on the mechanism and selectivity has been thoroughly investigated and explained. The essential interactions and binding preferences of catalysts in critical transition states have been explored, and the factors that influence the stereoselectivity of the reaction have been clarified. This work presents a new perspective on the challenges of exploring the mode of action and active conformations of complex multifunctional organocatalysts. As a future work, an automated tool that can identify potential catalyst candidates in a given pool for a target reaction could enormously facilitate an organocatalyst screening process.

ÖZET

ÇOK FONKSİYONLU ORGANOKATALİZÖRLERİN ETKİ ŞEKLİNİN İNCELENMESİ

Kuantum mekanik hesaplamaları, organokatalitik reaksiyon mekanizmalarını başarılı bir şekilde ortaya çıkarmış ve sonuçlarını açıklamıştır. Bununla birlikte, karmaşık çok fonksiyonlu organokatalizörlerin etki mekanizmasını anlamak, sayısız konformasyonel serbestlik derecesi nedeniyle zorlu bir hal almaktadır. Bu çalışma, CIPOC (Potansiyel (Organo)Katalizörlerin Hesaplama ile Tanımlanması) adında bir hesaplama protokolünü sunmaktadır. Bu protokol, kuantum mekanik hesaplamaları ve ilaç tasarım araçlarını birleştirerek karmaşık çok fonksiyonlu organokatalizörlerin etki mekanizmasını ve etkin konformasyonlarını keşfetme zorluklarını ele almayı amaçlamaktadır. CIPOC, belirtilen amaçlarla kullanılmıştır: (1) dimetil malonatın *trans-(β)*-nitrostiren ile asimetrik konjuge eklem reaksiyonunu katalizlemek için yeni kiral katalitik yapılar keşfetmek, bu da gamma-aminobütirik asit (GABA) benzerlerine erişimi mümkün kılmaktadır ve (2) oksindolün nitrosobenzen ile reaksiyonunda dimerik kinidinin etki mekanizmasını ortaya çıkarmak, bu reaksiyonda beklenmedik şekilde aminooksilasyon ürünleri elde edilmiştir. Dimetil malonatın *trans-(β)*-nitrostiren ile reaksiyonunu Takemoto'nun katalizöründen sekiz kat daha hızlı bir şekilde katalizleyen yeni bir 2-aminoDMAP/üre katalizörü keşfettik. Kinidin dimerinin nitrosobenzen ile protonlanmış amin tarafından bağlandığını, katalizör yapısındaki her iki hidroksil grubunun substratlarla etkileşime girdiğini ve enolatı yönlendirdiğini ve bu şekilde birleşik bir eklenme-proton transfer yolu oluşturduğunu gösterdik. Ayrıca, her iki çalışma da katalizör distorsiyonunun katalitik aktiviteyi belirlemedeki önemini ortaya koymuş olup, asitlik düzeyi daha düşük olsa bile üre bazlı katalizörlerin, tiyoüre analoglarına kıyasla olağanüstü performansa sahip olduğunu vurgulamıştır. Katalizörlerin mekanizma ve seçicilik üzerindeki etkisi detaylı bir şekilde incelenmiş ve açıklanmıştır. Önemli geçiş hallerinde katalizörlerin temel etkileşimleri ve bağlanma tercihleri araştırılmış ve reaksiyonun seçiciliğini etkileyen faktörler açıklığa kavuşturulmuştur. Bu çalışma, karmaşık çok fonksiyonlu katalizörlerin etki mekanizmasını ve etkin konformasyonlarını keşfetme zorluklarına yeni bir bakış açısı sunmaktadır. Belirli bir hedef reaksiyon için potansiyel katalizör adaylarını belirleyebilen otomatik bir aracın geliştirilmesi, organokatalizör tarama sürecini büyük ölçüde kolaylaştırabilir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

ABSTRACT vi

ÖZET vii

LIST OF FIGURES x

LIST OF TABLES xiv

LIST OF SYMBOLS/ABBREVIATIONS xv

1. INTRODUCTION 1

2. COMPUTATIONAL DETAILS 4

2.1. QUANTUM MECHANICS 4

2.1.1. *Ab initio* Methods 6

2.1.2. Semi-empirical Methods 7

2.1.3. Density Functional Theory 8

2.1.3.1. Functionals 10

2.1.3.2. Basis Sets 11

2.2. SOLVENT MODELS 12

2.2.1. Explicit Solvation Models 13

2.2.2. Implicit Solvation Models 14

2.3. VIRTUAL SCREENING 15

2.3.1. Pharmacophore Modelling in Virtual Screening 15

3. IDENTIFICATION OF CHIRAL BIFUNCTIONAL ACID-BASE CATALYSTS FOR MICHAEL ADDITION REACTIONS 17

3.1. INTRODUCTION 17

3.2. COMPUTATIONAL APPROACH 20

3.2.1. Theozyme Generation 20

3.2.1.1. Theozymes Based on Different Catalytic Groups 21

3.2.1.2. Theozymes Based on Different Binding Modes 22

3.2.2. Creating Catalytic Group Map & Pharmacophore Search 24

3.2.3. Generating Conformer Library of Organocatalyst 25

3.2.4. Matching the Theozyme 26
3.2.5. Quantum Mechanical Calculations 26
3.3. RESULTS AND DISCUSSION 28
3.3.1. Theozymes Based on Different Catalytic Groups 28
3.3.2. Theozymes Based on Different Binding Modes 30
3.3.3. Matching and Evaluation of the Full Catalysts 30
3.3.4. Catalytic Performance of Proposed Catalysts 37
3.4. CONCLUSION AND FUTURE WORK 38
4. EXPLORING THE MODE OF ACTION OF DIMERIC QUINIDINE CATALYZED ENANTIOSELECTIVE AMINOOXYLATION REACTION 39
4.1. INTRODUCTION 40
4.2. COMPUTATIONAL APPROACH 43
4.2.1. Theozyme Generation 44
4.2.2. Creating Catalytic Group Map & Pharmacophore Search 46
4.2.3. Generating Conformer Library of Organocatalyst 48
4.2.4. Matching the Theozymes 48
4.2.5. Quantum Mechanical Calculations 49
4.3. RESULTS AND DISCUSSION 50
4.3.1. Theozymes Based on Different Binding Modes 50
4.3.2. Matching and Evaluation of the Full Catalysts 52
4.3.2.1. Stepwise Pathway 52
4.3.2.2. Concerted Pathway 56
4.3.2.3. Evaluation of *N*- Selectivity 59
4.4. CONCLUSION AND FUTURE WORK 66
REFERENCES 67
APPENDIX A 79
APPENDIX B 84
APPENDIX C 85

LIST OF FIGURES

Figure 1.1. Outline of the developed computational approach 2

Figure 2.1. Schematic representation of solvent models [17] 13

Figure 3.1. Michael addition reaction of malonates to nitroolefins catalyzed by Takemoto's organocatalysts [56] 18

Figure 3.2. Newman projections (viewed along the forming C-C bond) of different binding modes reported (a) Takemoto [57], (b) Pápai [58], and (c) Izzo et al. [60] 19

Figure 3.3. Studied catalyst library 20

Figure 3.4. Newman projections (viewed along the forming C-C bond) of the theozymes with different catalytic groups 21

Figure 3.5. Newman projections (viewed along the forming C-C bond) of the theozymes with different binding modes 22

Figure 3.6. Nucleophilic effect of the enolate to the *Si* face or *Re* face of *trans-β*-nitrostyrene 22

Figure 3.7. Relative orientations of representative catalytic functional groups 23

Figure 3.8. Relative orientations of the 2-amino-pyridine group for *Si* and *Re* face model 23

Figure 3.9. Pharmacophore-like features in MOE [64] 25

Figure 3.10. Generation of catalytic group map (CGM) using theozyme and matched conformer of bifunctional catalyst, orange sphere=aromatic center, pink sphere= hydrogen-bond donors 26

Figure 3.11. Matching the catalytic active site model, orange sphere=aromatic center, pink sphere= hydrogen-bond donors 27

Figure 3.12. Newman projections (viewed along the forming C-C bond) of all theozymes for BMA and their activation free energies relative to the starting materials 29

Figure 3.13. Newman projections (viewed along the forming C-C bond) of Theozyme 2S with different binding modes and their activation free energies relative to the starting materials 30

Figure 3.14. Newman projections (viewed along the forming C-C bond) of the transition states for *Re* and *Si* face attack in the presence of 2O, 2S, and 1S catalysts and their activation energies 32

Figure 3.15. Distortion/Interaction Analysis of *(R)*-TS_{2O} & *(R)*-TS_{2S} and their theozyme transition states 33

Figure 3.16. Newman projections (viewed along the forming C-C bond) of the transition states for *Re* face in the inclusion of 3S, 2S, and 4S and their activation energies 35

Figure 3.17. Newman projections (viewed along the forming C-C bond) of the transition states for *Re* face in the presence of 2S with different binding modes and their activation free energies 36

Figure 4.1. Reaction of *N*-protected 3-methyl-2-oxindole with nitrosobenzene [91] 40

Figure 4.2. Reactions of 3-prochiral 2-oxindoles with nitrosobenzene 41

Figure 4.3. The proposed catalytic mechanism of the reaction between *N*-benzyl-protected 3-methyl-2-oxindole and nitrosobenzene in the presence of cinchona alkaloid-derived catalyst by (a) Barbas and co-workers [91] and (b) Çamlısoy's Thesis [98] 42

Figure 4.4. Barbas' dimerized quinidine derivative catalyst [91] 44

Figure 4.5. Theozymes of aminooxylation reaction for each binding mode 45

Figure 4.6. Proposed binding modes for cinchona alkaloid derived catalysts for (a) catalytic cycle I (b) catalytic cycle II 45

Figure 4.7. Construction of CGM using modelTS and matched conformer of dimeric catalyst 49

Figure 4.8. Matching the catalytic active site model for hydroxyamination 50

Figure 4.9. Theozymes and their relative Gibbs free energies relative to the starting materials for aminooxylation reaction 51

Figure 4.10. Different arrangement of benzyl group of substrate for aminooxylation model and their Gibbs free energies relative to starting materials 52

Figure 4.11. Transition states of aminooxygenation for the catalytic cycle I in the presence of dimeric catalyst with different binding modes and their Gibbs free energies relative to starting materials 54

Figure 4.12. Different arrangement of benzyl group of substrate for *(R)*-$TS_{A1.1}$ and their Gibbs free energies relative to starting materials 55

Figure 4.13. Transition state alternatives for *(R)*-$TS_{A1.2}$ and their Gibbs free energies relative to starting materials 56

Figure 4.14. Transition states of aminooxygenation for catalytic cycle II in the presence of a dimeric catalyst with different binding modes and their Gibbs free energies relative to starting materials 57

Figure 4.15. Lowest energy transition state geometry, *(R)*-$TS_{A2.2}$ from two different views and its Gibbs free energies relative to starting materials 58

Figure 4.16. Transition state alternatives for *(S)*-$TS_{A2.1}$ and their Gibbs free energies relative to starting materials 58

Figure 4.17. Theozymes and their relative Gibbs free energies relative to the starting materials for hydroxyamination reaction 60

Figure 4.18. Theozyme comparison Model A2 vs Model H2 and their relative Gibbs free energies relative to the lowest energy Model H2 61

Figure 4.19. *(R)*-$TS_{H2.2}$ and its Gibbs free energies relative to starting materials 62

Figure 4.20. Distortion/Interaction Analysis of *(R)*-$TS_{A2.2}$ & *(R)*-$TS_{H2.2}$ 63

Figure 4.21. Free energy profile in kcal/mol for the reaction of *N*-protected 3-methyl-2-oxindole with nitrosobenzene 65

Figure C.1. Free energy profile in kcal/mol for the addition reaction of dimethylmalonate to *trans*-*β*-nitrostyrene [B3LYP-D3/6-311+G(d,p)-SMD(Toluene)//B3LYP-D3/6-31G(d)-SMD(Toluene)].....85

LIST OF TABLES

Table 3.1. Selected pharmacophore groups for CGM (3D layout is shown in Figure 3.10)24

Table3.2. Experimental comparison of the catalytic performances of 2-aminoDMAP/(thio)urea (2S and 2O) and those of Takemoto's (1S).....37

Table 4.1. Selected pharmacophore groups for CGM (3D layout is shown in Table 4.2) ..46

Table 4.2. Representation of 3D layout of CGMs for aminooxylation modelTS.....47

Table A.1. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of bifunctional acid-base catalyzed model transition states of *trans-β*-nitrostyrene and dimethyl malonate.....79

Table A.2. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of bifunctional acid-base catalyzed transition states of *trans-β*-nitrostyrene and dimethyl malonate.....80

Table A.3. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of energy profile for bifunctional thiourea catalyzed transition states of *trans-β*-nitrostyrene and dimethyl malonate.....81

Table A.4. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of model transition states of *N*-benzyl-protected 3-substituted-2-oxindole with nitrosobenzene 82

Table A.5. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of dimeric quinidine catalyzed transition states of *N*-benzyl-protected 3-substituted-2-oxindole with nitrosobenzene 83

Table B.1. Distortion/Interaction Analysis of *(R)*-TS_{2O} & *(R)*-TS_{2S} * 84

Table B.2. Distortion/Interaction Analysis of $TS_{Theozyme\ 2O}$ & $TS_{Theozyme\ 2S}$ *.....84

Table B.3. Distortion/Interaction Analysis of *(R)*-$TS_{A2.2}$ & *(R)*-$TS_{H2.2}$ *.....84

LIST OF SYMBOLS/ABBREVIATIONS

Å	Angstrom
E	Electronic energy
e	Charge on the electron
$ee\%$	Enantiomeric excess
$E_{XC}[\rho(\boldsymbol{r})]$	Exchange-correlation energy
$\widehat{H}$	Hamiltonian operator
$J[\rho(\boldsymbol{r})]$	Classical Coulomb energy
m	Mass of particle
ΔG_{elec}	Electrostatic component of the solute-solvent interaction energy
ΔG_{cav}	Energetic cost of placing the solute in the medium through a cavity
ΔG_{disp}	Dispersion interactions between solvent and solute
$\Delta G^{\neq}$	Activation Gibbs free energy
$\Delta H^{\neq}$	Activation enthalpy
R	Universal gas constant
r	Distance between the particles
T	Temperature
TS	Transition state
$T_s[\rho(\boldsymbol{r})]$	Kinetic energy of non-interacting electrons
V_{XC}	Exchange-correlation potential
$V(\boldsymbol{r})$	External potential
$V_{HF}(i)$	Hartree-Fock potential
Z	Atomic number
ψ	Wave function
∇^2	Laplacian operator
$\hat{f}_i$	Fock operator
x_i	molecular orbital of the electron (i)
χ_i	Spin orbital
ε_i	energy of the orbital of the electron (i)
$\rho(\boldsymbol{r})$	Electron density

$\hbar$	Planck's constant
AM1	Austin model 1
B3LYP	Becke-3-parameter Lee-Yang-Parr functional
B	Becke
BM	Binding mode
CC	Coupled cluster theory
CGM	Catalytic group map
CIPOC	Computational identification of (organo)catalysts
COSMO	Conductor-like screening model
CSM	Continuum solvent models
DFT	Density functional theory
DMAP	Dimethylaminopyridine
DMSO	Dimethyl sulfoxide
EDA	Energy decomposition analysis
GGA	Generalized gradient approximation
GTO	Gaussian-type orbital
HBA	H-bond acceptor
HBD	H-bond donor
HF	Hartree-Fock
IGM	Independent gradient model
IRC	Intrinsic reaction coordinate
K	Kelvin
LDA	Local density approximation
LSDA	Local spin density approximation
LYP	Lee-Yang-Parr
MC	Monte Carlo
MD	Molecular dynamic
mdb	Molecular database
MGGA	Meta generalized gradient approximation
MM	Molecular mechanic
MNDO	Modified neglect of diatomic overlap
MOE	Molecular operating environment

MP	Møller-Plesset perturbation theory
NDDO	Neglect of the diatomic differential overlap method
PCM	Polarizable continuum model
PDDG/PM3	Pairwize distance directed Gaussian/Parametric method 3
pdb	Protein data bank
PM3	Parametric method 3
RMSD	root mean square deviation
QM	Quantum mechanic
QTAIM	Quantum theory of atoms in molecules
SCF	Self-consistent field
SMD	Solvation model based on density
STO	Slater-type orbital
THF	Tetrahydrofuran
TS	Transition state
VS	Virtual screening
3D	Three dimensional

1. INTRODUCTION

Understanding the activation mechanism and mode of action of known catalysts is crucial for rational and systematic design of novel catalysts for important transformations commonly employed in industry. Computational approaches have emerged as excellent tools to explain reaction mechanisms and selectivity [1]. Quantum mechanical calculations have successfully uncovered numerous organocatalytic reaction mechanisms and explained the observed reaction outcomes; yet, understanding the mode of action of highly complex multifunctional organocatalysts is still a challenging task due to their large number of conformational degrees of freedom. The existing computational strategies for exploring organocatalytic reactions are limited by the manual construction of a few structures based on chemical intuition and imagination. There are still no approaches that can adequately capture the full conformational space of the reaction and important binding modes of multifunctional organocatalysts. An automated tool that can identify potential catalyst candidates in a given pool for a target reaction could enormously facilitate an organocatalyst screening process.

In this thesis, we describe a computational approach, CIPOC (Computational Identification of Potential (Organo)Catalysts) [2] that combines the quantitative power of quantum mechanical calculations with drug design tools to shed light on the ambiguous active conformations and mode of action of multifunctional organocatalysts with complex structures, and to discover new catalysts.

This approach (Figure 1.1) is based on the construction of a theoretical active site model including functional groups that are involved in the structures of the multifunctional organocatalysts. It basically is a quantum mechanically optimized transition state of the rate-determining step of the target reaction with optimal 3D arrangement of catalytic functional groups around the substrates. These models, so-called theozymes, have been successfully used in the “inside-out” design of enzymes for reactions not catalyzed by nature [3–5] and multifunctional spirocyclic oligomers as transesterification catalysts [6]. Catalytic group maps (CGM) are constructed with pharmacophore-like features such as an H-bond acceptor and donor, aromatic ring centers, etc. Then, the catalyst conformer library is screened against CGM in order to identify the conformers that can stabilize the optimized transition structure. On the quantum mechanically optimized theozyme, replacing the model catalytic groups

with the matching conformer of the catalyst gives the transition structure in the presence of the chiral multifunctional catalyst. Finally, the rate and selectivity of the leads are evaluated using quantum mechanical calculations.

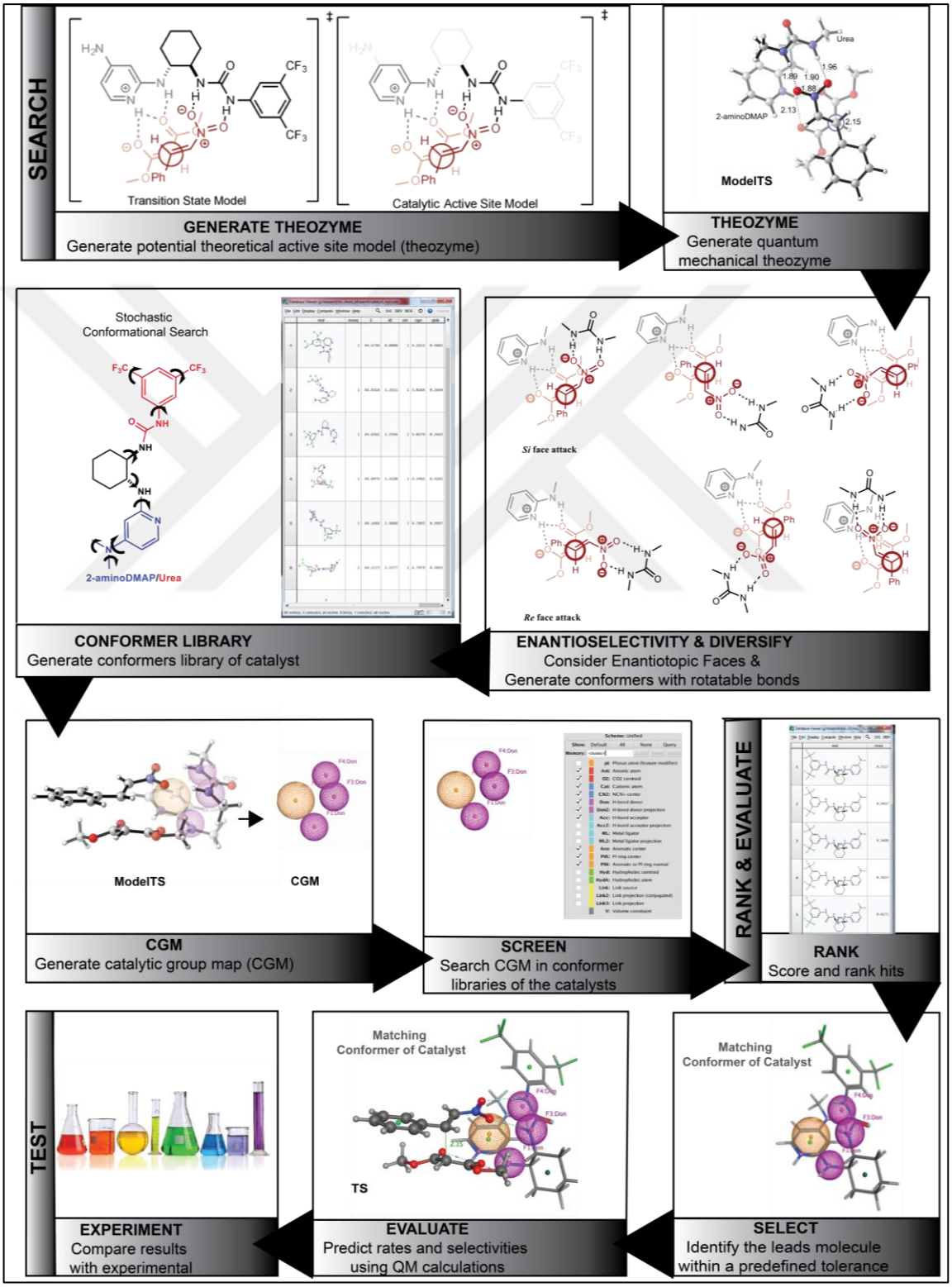


Figure 1.1. Outline of the developed computational approach

The dissertation starts with Computational Details (Chapter 2); the fundamental principles underlying the computational techniques employed in this study are presented. The next two chapters describe the application of CIPOC for two target reactions. The first reaction of interest is the asymmetric conjugate addition reaction of dimethyl malonate to *trans-(β)*-nitrostyrene in the presence of chiral bifunctional acid-base organocatalysts (Chapter 3). Second, the mode of action of dimeric quinidine-catalyzed enantioselective reaction of *N*-benzyl-protected oxindole with nitrosobenzene is explored (Chapter 4). Both chapters (Chapters 3 & 4) are designed with their introduction, details of computational approach, results and discussion, and conclusion sections. In the introduction parts, the importance of the subject has been emphasized, and relevant investigations in this field are presented. The CIPOC protocol is explained in detail in the computational approach sections, focusing on the example catalyst studied for the target reaction in that part. The following part of this thesis presents the outcomes of our computations, which are analyzed, discussed, and compared with experimental findings. Lastly, the concluding remarks are presented. Additionally, recommendations for a future study are noted for each section.

2. COMPUTATIONAL DETAILS

Computational chemistry is applied to predict molecular structure and reactivity on computers with well-developed mathematical methods. It has become an increasingly important tool in modern chemical research and is considered indispensable by many chemists [7].

Computational chemistry uses mathematical techniques to mimic chemical structures and reactions based on fundamental physical laws. Instead of conducting experiments, chemical phenomena can be examined through computer-based calculations by chemists.

Through theoretical chemistry, researchers can gather knowledge about the property of molecules and chemical reactions that cannot be obtained through experimental observations alone. This makes computational chemistry an essential part of modern chemical research, as it can provide insights into chemical systems that are difficult to obtain through traditional experimental methods. As such, computational chemistry can be considered a complementary research area to experimental chemistry [8].

Generally, methods of computational chemistry are investigated in two categories:

- Classical (or Newtonian) mechanical methods
- Quantum mechanical methods

2.1. QUANTUM MECHANICS

Quantum mechanics (QM) has become an essential tool for studying chemical reactions. The Schrödinger equation lies at the core of QM. It gives exact solutions for the particle in a box, the particle on a ring, the harmonic oscillator, the particle on a sphere, and the hydrogen atom [9].

The field of quantum mechanics has been employed to explain the behaviour of small-scale systems. The fundamental principle of QM is the Schrödinger equation, which defines the electronic properties of molecules.

The form of time-independent Schrödinger equation is:

$$\hat{H}\psi = E\psi \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian operator (involves kinetic and potential energy), E is the electronic energy of the system, and ψ is the wave function. A system is described by ψ, a function that depends on the locations of nuclei and electrons in the system. The energy of the system, involving both the potential and kinetic energies of the nuclei and electrons, can be represented as a sum by the Hamiltonian operator.

$$\hat{H} = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i{}^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k{}^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} \tag{2.2}$$

In this equation, m_e and m_k represent the mass of the nucleus and electron, ħ is Planck's constant divided by 2π, ∇^2 is a mathematical operator called Laplacian, e is the charge carried by an electron, r represents the distance between the two particles, and Z is the atomic number. The given expression consists of five terms: the first term expressed the kinetic energy of the electrons, the second term expressed the kinetic energy of the nuclei, and the third term expressed the attraction between the electrons and the nuclei. The remaining two terms represent the interelectronic and internuclear repulsions, respectively [7].

The Schrödinger equation can be precisely calculated for hydrogen atoms only, but for many-particle systems, exact solutions cannot be found due to the correlated motions of the particles. However, some approximations can be made to overcome this issue. Additionally, since the nucleus of the mass of the molecule is significantly greater than that of the electrons, its motion can be assumed negligible compared to that of the electrons. They can be considered stationary, i.e., their kinetic energy can be supposed to be zero. The Born-Oppenheimer approximation suggests that the attractive potential energy term between nuclei and electrons can be neglected. The term for nuclear kinetic energy is considered independent of electrons [10]. The potential energy term between nuclei is assumed to be constant, and hence, the Schrödinger equation for Born-Oppenheimer approximation becomes:

$$\left(\hat{H}_{el} + \hat{V}_{nn}\right)\psi_{el} = E_{el}\psi_{el} \tag{2.3}$$

Various QM methods based on the Schrödinger equation with approximate solutions were developed to solve Equation 2.3. Two main categories exist for QM methods.

- Semi-empirical
- Non-empirical

2.1.1. *Ab initio* Methods

Ab initio methods utilize QM calculations to solve the molecular Schrödinger equation related to the molecular Hamiltonian. Due to the electron-electron interaction terms, finding an exact analytical solution to the Schrödinger equation is challenging. Therefore, numerical iterative methods solve the equation with a predefined error margin. The available computational methods provide precise and trustworthy numerical solutions for wave functions and energies [10].

The Hartree-Fock (HF) approach is the simplest *ab initio* electronic structure calculation, where the instantaneous electron-electron repulsion is neglected and its average effect taken into account. To correct for electron-electron repulsion, post-Hartree-Fock methods such as coupled cluster theory (CC) and Møller-Plesset perturbation theory (MPn) are commonly used. However, the Hartree-Fock method may not be sufficient for some cases, especially bond-breaking processes.

Ultimately, the variational theorem offers a means to determine the best solution for a problem, particularly in the case of many-body systems where no definitive solution exists. The theorem asserts that the energy calculated from an approximate wave function will always be higher than the actual energy. This means a more accurate wave function will give a lower energy value. The "best" wavefunction is consequently attained when the energy reaches its minimum point. At this juncture, the first derivative of energy will be equal to zero. To reach the energy minimum, an iterative procedure called the self-consistent field (SCF) theory is employed. This iterative process enables the adjustment of the wave function to approach the most stable energy state successively. SCF method considers both Coulomb and exchange integrals and utilizes HF equation to iteratively determine the optimal wavefunction.

$$\hat{f}_i x_i = \varepsilon_i x_i \tag{2.4}$$

In equation 2.4, x_i is the molecular orbital of the electron (i), $\hat{f}_i$ is the Fock operator, and ε_i is the corresponding energy of the orbital. The one-electron Fock-operator isrepresented as shown in equation 2.5.

$$\hat{f}_i = -\frac{1}{2}\nabla_i{}^2 - \sum_{k}^{nuclei} \frac{Z_k}{\boldsymbol{r}_{ik}} + V_{HF}(i) \tag{2.5}$$

where $V_{HF}(i)$ is the HF potential which is defined the average repulsive potential experienced by each electron due to the other electrons. The SCF method begins by initializing an initial guess for the molecular orbitals based on a specific molecular geometry. The calculated wavefunction x_i is then used to generate a new Hartree-Fock potential, $V_{HF}(i)$. This iterative procedure is repeated until convergence is accomplished [10].

The Hartree-Fock method, also known as SCF theory is a valuable tool for initial predictions about various systems. It can generate accurate outcomes when determining the configurations and vibrational frequencies of steady molecules as well as some transitional states. But, this method fails to consider electron correlation, which arises from the immediate repulsion between electrons. As a result, it is not suitable for accurately modelling the energetics of reactions and bond dissociation [11].

2.1.2. Semi-empirical Methods

Semi-empirical methods perform calculations by neglecting or parameterizing certain integrals based on the experimental information. Therefore, this method is high-speed and suitable for large molecules. While semi-empirical methods may provide accurate results for molecules similar to those used for parameterization, their accuracy is only sometimes reliable for various systems.

Modern semi-empirical models are based on the Neglect of the Diatomic Differential Overlap (NDDO) method, which makes it simple. There are four different NDDO models. These are Austin Model 1 (AM1), Parametric Method 3 (PM3), Pairwize Distance Directed Gaussian/Parametric Method 3 (PDDG/PM3), and Modified Neglect of Diatomic Overlap (MNDO) [12].

Semi-empirical calculations have the advantage of being much faster than ab-initio calculations. It could be used to find an initial prediction for ab-initio calculations. However, semi-empirical methods are that the outcomes can be doubtful, and limited properties reliably can be expected. In addition, if a molecule analyzed has similarities to the molecules used in the method's database for parameterization, then the outcomes of the calculations may be very accurate. On the other hand, if a molecule that differs greatly from the molecules in the parameterization database may result in less accurate answers [10].

2.1.3. Density Functional Theory

Density Functional Theory (DFT) is based on electron density rather than wave functions, in contrast to Hartree-Fock Theory. The fundamental principles of DFT are the Hohenberg-Kohn theorems [11], which were proposed in 1964. According to the first theorem, the electron density $\rho(\boldsymbol{r})$ defines the external potential $V(\boldsymbol{r})$, while the second theorem introduces the variational principle. This represents that for a specified electron density, there is only one specific distribution of nuclei with their presented charges, and the energy is minimized with respect to electron density. It can be expressed using Equation 2.6.

$$\rho(\boldsymbol{r}) = N \int \dots \int |\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_n)|^2 d_{\boldsymbol{r}_1} d_{\boldsymbol{r}_2} \dots d_{\boldsymbol{r}_n} \tag{2.6}$$

where $\boldsymbol{r}$ describes coordinates of electrons. The electronic energy of the system in its ground state is a function of the electron density. It is the essence of the "functional" aspect of DFT and can be defined as:

$$E[\rho(\boldsymbol{r})] = \int V_{ext}(\boldsymbol{r})\rho(\boldsymbol{r})d\boldsymbol{r} + T[\rho(\boldsymbol{r})] + V_{ee}[\rho(\boldsymbol{r})] \tag{2.7}$$

The equation can be described as the sum of three terms, where the term $V_{ext}(\boldsymbol{r})$ represents the interactions of the electrons with an external potential. The second term, $T[\rho(\boldsymbol{r})]$, is the kinetic energy of the interacting electrons, and the third term, $V_{ee}[\rho(\boldsymbol{r})]$, defines the interelectronic interaction energy. The electronic energy is changed as:

$$E[\rho(\boldsymbol{r})] = \int V_{ext}(\boldsymbol{r})\rho(\boldsymbol{r})d\boldsymbol{r} + T_s[\rho(\boldsymbol{r})] + J[\rho(\boldsymbol{r})] + E_{XC}[\rho(\boldsymbol{r})] \tag{2.8}$$

where is the classical Coulomb energy $J[\rho(\boldsymbol{r})]$ between electrons, which is commonly referred as the Hartree electrostatic energy:

$$J[\rho(\boldsymbol{r})] = \frac{1}{2}\int\int \frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} d_{\boldsymbol{r}_1} d_{\boldsymbol{r}_2} \tag{2.9}$$

$T_s[\rho(\boldsymbol{r})]$ describes the kinetic energy of the non-interacting electrons:

$$T_s[\rho(\boldsymbol{r})] = \sum_{i=1}^{N} \int \psi_i(\boldsymbol{r}) \left(-\frac{\nabla^2}{2}\right) \psi_i(\boldsymbol{r}) d\boldsymbol{r} \tag{2.10}$$

and the exchange-correlation energy functional, $E_{XC}[\rho(\boldsymbol{r})]$ represents the correction to the kinetic energy resulting from the interactions among electrons and nonclassical adjustments to the electron-electron repulsion energy. It includes contributions from exchange $E_X[\rho(\boldsymbol{r})]$ and correlation $E_C[\rho(\boldsymbol{r})]$[10].

The Kohn-Sham scheme provides the full statement for the energy of an *N*-electron system, which can be written as Equation 2.11 [11].

$$\begin{aligned} E[\rho(\boldsymbol{r})] = &- \sum_{A=1}^{M} \int \frac{Z_A}{|\boldsymbol{r} - \boldsymbol{R}_A|} \rho(\boldsymbol{r}) d\boldsymbol{r} + \sum_{i=1}^{N} \int \psi_i(\boldsymbol{r}) \left(-\frac{\nabla^2}{2}\right) \psi_i(\boldsymbol{r}) d\boldsymbol{r} \\ &+ \frac{1}{2}\int\int \frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} d_{\boldsymbol{r}_1} d_{\boldsymbol{r}_2} + E_{XC}[\rho(\boldsymbol{r})] \end{aligned} \tag{2.11}$$

The wave function of this system is represented by using a Slater determinant of one-electron functions X_i, and the electron density is represented as:

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} \langle X_i | X_i \rangle \tag{2.12}$$

The solution to the Kohn-Sham equations determines the Kohn-Sham orbitals [11].

$$h_i{}^{KS} X_i = \varepsilon_i X_i \tag{2.13}$$

where the Kohn-Sham Hamiltonian, h_i, is expressed as:

$$h_i{}^{KS} = -\frac{\nabla^2}{2} - \sum_{k}^{M} \int \frac{Z_k}{|\boldsymbol{r}_i - \boldsymbol{r}_k|} + \frac{\rho(\boldsymbol{r})}{\boldsymbol{r}_{ij}} d\boldsymbol{r} + V_{XC} \tag{2.14}$$

In Equation 2.14, V_{XC} is the exchange-correlation potential. It is related to exchange-correlation energy by Equation 2.15 [11].

$$V_{XC} = \frac{\delta E_{XC}}{\delta \rho} \tag{2.15}$$

2.1.3.1. Functionals

The term $E_{XC}[\rho(\boldsymbol{r})]$ accounts for more than just exchange and correlation effects, as it also includes the discrepancy in the kinetic energy of the actual system and a theoretical non-interacting system. The specific mathematical expression for this exchange-correlation function is currently unknown, and the functionals described here are only various approximations of this function.

Local Density Approximation (LDA) is commonly used, which assumes a uniform electron gas where the electron density is the same everywhere to approximate the exchange-correlation function. However, this assumption is inappropriate for molecules where electron density varies fast. To overcome this, Generalized Gradient Approximation (GGA) methods take into account the inhomogeneity of electron density by incorporating its gradient in addition to the density itself in the exchange and correlation energies [12].

In its original form, a meta-GGA (MGGA) DFT functional includes the second derivative of the electron density (the Laplacian). This progression after the GGA incorporates only the density and its first derivative in the exchange-correlation potential. In contemporary usage, the term MGGA functional more commonly refers to a function that combines the kinetic energy density, specifically the Laplacian of the orbitals. Several functionals belonging to this category include tHCTH, TPSS, VSXC, and M06L [11].

The hybrid methods combines mixing various amounts of the HF nonlocal exchange operator with local DFT exchange-correlation functionals and gradient-corrected density functions. The parameters of the hybrid functional are optimized by comparing the results to a database of thermochemical data.

Empirical three-parameter hybrids refer to a mixing approach that combines LDA, HF exchange, and gradient-corrected exchange and correlation functionals.

$$E_{XC} = (1-a){E_X}^{LDA} + a{E_X}^{HF} + b\Delta{E_X}^{GC} + {E_C}^{LDA} + c\Delta{E_C}^{GC} \tag{2.16}$$

where ${E_X}^{HF}$ is the nonlocal HF exchange energy, ${E_X}^{LDA}$ and ${E_C}^{LDA}$ are exchange and correlation energies within LDA formalism, $\Delta{E_X}^{GC}$ and $\Delta{E_C}^{GC}$ are the gradient corrections for the exchange and correlation functionals.

Hybrid density functional methods refer to alternative approaches that include a percentage of HF exchange with the exchange correlation of the GGA method. Three-parameter functional B3LYP [13, 14] which is used in this thesis, is the most famous example of hybrid density functional. It is a combination method of LDA and GGA functionals taken from the DFT and HF methods, as shown in Equation 2.17.

$$E_{XC}{}^{B3LYP} = (1-a)E_X{}^{LSDA} + aE_X{}^{HF} + bE_X{}^{B} + (1-c)E_C{}^{LSDA} + aE_C{}^{LYP} \quad (2.17)$$

where a, b, and c were optimized to 0.20, 0.72, and 0.81, respectively, using a set of atomization energies, proton affinities, ionization potentials and total atomic energies, and LSDA (Local Spin Density Approximation), HF (Hartree-Fock), B (Becke) and LYP (Lee-Yang-Parr) are the functionals [11].

Several one-parameter hybrid methods have been suggested, which modify the percentage of HF exchange incorporated into the functional. The one-parameter hybrid exchange-correlation energy can be expressed as follows:

$$E_{XC} = (1-a)E_X{}^{DFT} + aE_X{}^{HF} + E_C{}^{DFT} \quad (2.18)$$

In equation 2.18, a is the percentage of HF exchange in hybrid functional, $E_X{}^{HF}$ is the nonlocal HF exchange energy, $E_C{}^{DFT}$ is the local DFT correlation energy ($E_C{}^{DFT} = E_C{}^{LDA} + E_C{}^{GC}$), and $E_X{}^{DFT}$ is the local DFT exchange energy ($E_X{}^{DFT} = E_X{}^{LDA} + E_X{}^{GC}$)[11].

2.1.3.2. Basis Sets

A basis set is a mathematical representation of the orbitals that create a system and are used to perform computational chemistry to describe the electronic wave function in the HF method or DFT. The aim is to convert the partial differential equations of the model into algebraic equations that can be easily implemented on a computer.

A basis set is a collection of functions used to describe the orbitals in a system. Roothaan proposed the Hartree-Fock orbitals in 1951 as a linear combination of atomic orbitals, and the coefficients are determined through calculation. There exist two primary categories of basis sets: Gaussian-type orbitals (GTOs) and Slater-type orbitals (STOs), with the latter being more commonly used in modern DFT.

The basic Hartree-Fock method uses the smallest possible set of basis functions on each atom to contain all the electrons and maintain spherical symmetry, resulting in molecular

orbitals with limited flexibility. Using larger basis sets increases the number of adjustable coefficients and results in a more accurate description of the molecular orbitals. Extremely large basis sets can provide nearly complete flexibility, and the limit of this approach is called the HF limit, which provides the best possible definition of the system using a single electron configuration.

Many Gaussian basis sets are available for performing SCF calculations, with the ones developed by Pople and co-workers being the most commonly used. The STO-3G basis set is the simplest and smallest one, where three Gaussian functions represent the Slater-type orbitals. It is considered a minimal basis set as it contains only the number of orbitals required to hold the electrons of the neutral atom.

Pople and co-workers developed different levels of basis sets for carrying out self-consistent field (SCF) calculations, and one of the most commonly used is the split-valence basis set. A minimal basis set (STO-3G) has limitations regarding its ability to adjust its orbitals to fit the molecular environment. Split-valence basis sets solve this problem by dividing the basis into two parts: an inner, compact orbital, and an outer, more diffuse one. This allows for greater variation in the size of the atomic orbitals that contribute to the molecular orbital within the range set by the inner and outer basis functions.

Adding polarization functions to increase flexibility can improve the split-valence basis set. In polarization basis sets, d orbitals are added to all heavy atoms, designated with a * or (d). It is also possible to add polarization to hydrogen atoms by using **.

To better represent the electronic wave function in specific systems, larger-sized functions called diffuse functions can be added to the standard valence-size functions. These diffuse functions allow orbitals flexibility. They are significant in systems where electrons are relatively far from the nucleus, such as those with lone pairs, anions, and excited states. The 6-31+G (d) basis set includes diffuse functions added to heavy atoms, while the 6-31++G (d) also adds them to hydrogen atoms. The 6-31+G* and 6-31+G** basis sets, which use split valence approximation, have been utilized in gas-phase calculations. [10].

2.2. SOLVENT MODELS

A solvent model is an approach in computational chemistry used to investigate the behavior of solvated condensed phases, including reactions and processes occurring in biological, chemical, and environmental solutions. By accounting for the behavior of solvents, these

models allow simulations and thermodynamic calculations that can provide new insights and predictions about physical processes [15, 16].

Solvent models are used in computational chemistry to simulate and calculate reactions and processes in solutions, such as biological, chemical, and environmental systems. Theoretical studies related to reaction mechanisms using QM can approximate the thermodynamic properties of liquids based on three main types of solvent models: implicit (or continuum), explicit solvation models, and cluster/continuum (or hybrid implicit/explicit) (Figure 2.1) [17]. To model the solvation effect, implicit solvent models assume the solvent as a homogeneous medium that surrounds the solute, whereas explicit solvent models consider the behavior and interactions of individual solvent molecules in a specific region around the solute. Hybrid models combine aspects of both explicit and implicit models to minimize computational costs while maintaining some spatial resolution of the solvent [18].

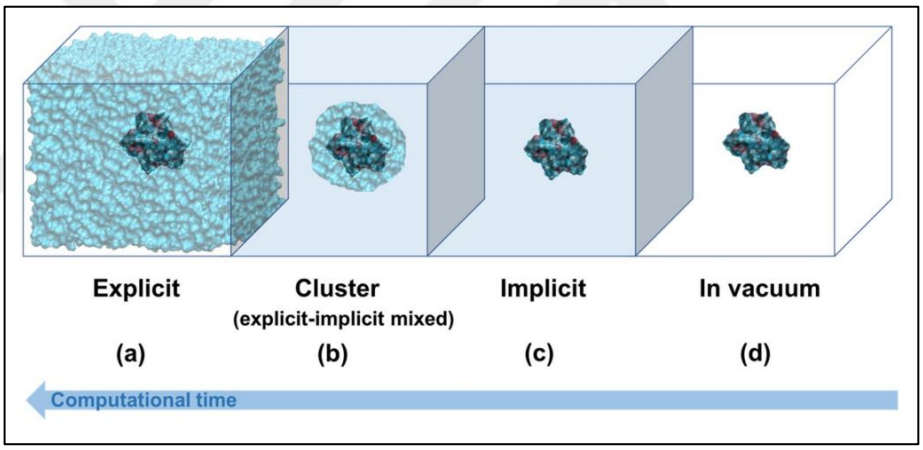


Figure 2.1. Schematic representation of solvent models [17]

2.2.1. Explicit Solvation Models

Explicit solvent models directly include the solvent molecules in the computation by considering their coordinates and molecular degrees of freedom. This method is considered more realistic because it takes into account direct and specific interactions between the solvent and solute, in contrast to implicit continuum models. Fully explicit solvent models can provide a realistic representation of the environment in which catalytic reactions occur in the condensed phase. In such models, the solute is situated inside a box surrounded by sufficiently solvent molecules to replicate the system's density. These models are typically

utilized in molecular mechanics (MM), Monte Carlo (MC), and molecular dynamics (MD) simulations. QM calculations may use solvent clusters. MD simulations are useful for studying a chemical system's time evolution. They often rely on empirical force fields to effectively determine the properties and motions of large systems [11].

2.2.2. Implicit Solvation Models

Continuum solvent models (CSM) are a type of implicit solvent model used in computational chemistry where no explicit solvent molecules are included in the studied system. The homogeneous polarizable medium in this approach refers to the solvent, and it is described by its dielectric constant ε. The solute is positioned within a cavity formed within the solvent medium in continuum solvent models,. The solute's electron density within the cavity induces the solvent medium's polarization, which modifies its electron distribution. This method evaluates the electrostatic part to the solvation Gibbs energy. However, it is required to consider non-electrostatic contributions and electrostatic ones to obtain more precise solvation energies [19]. Different theoretical assumptions, cavity definitions, and boundary conditions create various implicit solvation models. These are the Solvation Model Based on Density (SMD), the Polarizable Continuum Model (PCM), and the Conductor-like Screening Model (COSMO). These solvent models are computationally efficient and maintain a system size similar to that in the gas phase while avoiding sampling over solvent degrees of freedom. Therefore, they have become the most widely used method in homogeneous catalysis [20]. Nonetheless, continuum solvent models have some limitations. They describe the solvent as a uniform medium and thus cannot account for specific interactions between solute and solvent molecules, which is significant in particular scenarios. Therefore, if solvent molecules can directly include in the reaction, this solvent model may not be suitable.

In the implicit solvation model, different physical effects contribute to the overall solvation process and solvation of a free energy of a system is expressed as:

$$\Delta G_{sol} = \Delta G_{elec} + \Delta G_{cav} + \Delta G_{disp} \tag{2.19}$$

In equation 2.19, ΔG_{elec} is the electrostatic component of the solute-solvent interaction energy and has a stabilization effect. ΔG_{cav} is the energetic cost of placing the solute in the

medium through a cavity. Dispersion interactions between solvent and solute are defined as ΔG_{disp} which add stabilization to the solvation free energy.

Various well-established solvent models are employed effectively in different systems. PCM is a widely used implicit model that has led to the development of various versions. It is based on the Poisson-Boltzmann equation, a modified form of the original Poisson's equation. SMD has been commonly worked in addition to the PCM. SMD model solves the Poisson-Boltzmann equation in a way similar to PCM but uses a set of precisely parameterized radii to construct the cavity [20]. The SMD method, which is used in quantum-chemical research [16], is more sophisticated than other models in that it includes specific parameterizations for various solvents using several parameters.

2.3. VIRTUAL SCREENING

In drug discovery, virtual screening (VS) is a computer-based approach to search large databases of small molecules to discover potential drug candidates that could bind to a specific protein receptor or enzyme target [21].

Virtual screening is organized into two primary categories: structure-based (docking) and ligand-based screening. Docking is a computational method used in structure-based drug design that predicts the geometry of protein-ligand interactions and estimates binding affinities. On the other hand, ligand-based screening uses active compounds as templates [22].

2.3.1. Pharmacophore Modelling in Virtual Screening

Pharmacophore modeling is commonly utilized in virtual screening to discover molecules that can prompt the requested biological activity. To achieve this, scientists create a pharmacophore model or a set of criteria that describes the required 3D arrangement for the desired interaction pattern. This model is then used to search databases of chemical compounds for molecules that fit the criteria and are likely to show the desired biological activity [23].

To determine molecules that exhibit the desired biological activity in drug discovery, researchers construct a pharmacophore model, which includes various features positioned in a specific 3D pattern. Each feature is described by a sphere determining the allowed deviation from its exact position. This method represents various molecular recognition patterns such as hydrogen bond donors, hydrogen bond acceptors, anionic, cationic, hydrophobic, and aromatic rings using spheres with a defined radius as a geometric constraint. The size of the sphere defines the level of strictness in the constraint. A second feature can be added to indicate the direction of specific interactions like hydrogen bonds and aromatic rings. The pharmacophore query can combine multiple features in any manner [24]

In order to screen a library of catalysts, pharmacophore features are commonly benefited as queries. The library comprises different conformations, and each is compared to the query by aligning the molecule's pharmacophore features with those of the query. A molecule is considered to be a hit if it fits within the spheres representing the features of the query. However, in cases where the query is too complex to find matching molecules from the library, partial matching can be permitted, in which only the crucial features are matched [23].

3. IDENTIFICATION OF CHIRAL BIFUNCTIONAL ACID-BASE CATALYSTS FOR MICHAEL ADDITION REACTIONS

Asymmetric organocatalysis is an attractive approach for synthesizing organic molecules with high enantiomeric purity [25–27]. Significant advancements in organic synthesis have been reported in recent years, in the field of organocatalysis, with a particular focus on thiourea bifunctional organocatalysts [25, 28]. Although organocatalysis is commonly regarded as an environmentally friendly and cost-effective catalytic alternative, it often requires high concentrations of catalysts (10 mol% or higher) [29]. As a result, searching for more active and stereoselective catalysts becomes a fundamental challenge in this type of catalysis. Thus, an understanding of the reaction mechanism is crucial not only for discovering new catalysts but also for the development of more efficient low-loading catalytic systems.

In this chapter of the thesis, the proposed protocol (CIPOC) was applied to discover new multifunctional organocatalysts and understand the origins of catalysis and enantioselectivity for the addition of dimethyl malonate with *trans-β*-nitrostyrene in collaboration with Prof. Cihangir Tanyeli's Research Group at Middle East Technical University (METU).

3.1. INTRODUCTION

Asymmetric organocatalysis by bifunctional acid/base type small organic molecules enjoys widespread use as catalysts in stereoselective transformations [30–32]. The design of such catalysts essentially relies on the suitable placement of an acid (H-bond donor, HBD) and a base (H-bond acceptor, HBA) together around some stereogenic elements. The simplicity of this guiding design principle has given inspiration to many hundreds to develop their own with much of the interest centered on HBDs (for example, urea and thiourea) [33–37]. Although relatively less explored, Brønsted basic HBAs (for example, amine), have been shown to be game-changers in these endeavors [38]. Guanidines [39–43], iminophosphoranes [44, 45], cyclopropenimines [46, 47], and amidines [48–52] are among some influential, emerging *organo-superbase* motifs developed to this end. Inspired by these organobases, in particular, with that of Wulf et al. [53], Tanyeli and co-workers have found

2-aminoDMAP superbase a fruitful precatalyst readily yielding sulfonamides [54] and squaramides [55] catalyze Michael additions of 1,3-diketones to *trans*-nitroolefins. These catalysts, however, failed to give any addition products with malonates, that gives *gamma*-aminobutyric acid (GABA) close analogues. The first efficient organocatalytic process for this reaction was made possible by the *tert*-amine/thiourea catalyst (**1S**), a.k.a. Takemoto's catalyst [56, 57]. This catalyst typically required 10 mol% catalyst loadings and day-long reaction times to afford Michael adducts in good yields (Figure 3.1).

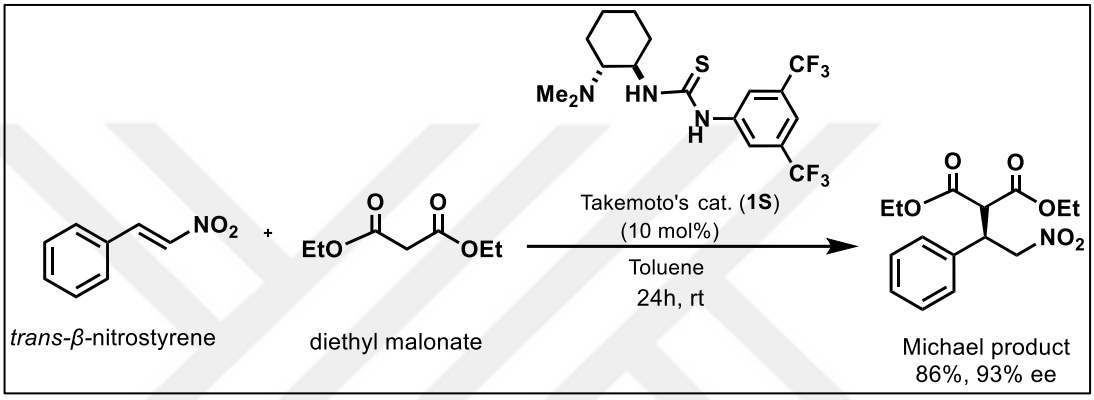


Figure 3.1. Michael addition reaction of malonates to nitroolefins catalyzed by Takemoto's organocatalysts [56]

To develop more active and stereoselective catalysts, an understanding of the reaction mechanism is important. Several groups computationally explored the origin of enantioselectivity and the rate-determining step of the reaction. Pápai and co-workers [58] reported a low energy binding mode in which the malonate interacts with the thiourea and nitroolefin interacts with a protonated amine, in contrast to the binding mode proposed by Takemoto, in which malonate H-bond with the protonated amine, whereas thiourea stabilizes the nitroolefin (Figure 3.2 (a) & (b)). Theoretical investigations by Liu et al. [59] supported the binding mode proposed by Pápai and co-workers [58]. Even though Hamza et al. [58] and Liu et al. [59] provide valuable insights into the enantioselectivity of the reaction, the lack of overall free energy barrier ($\Delta G^{\neq}$) limits the interpretation of the experimentally observed kinetics. Liu et al. identified the final proton transfer step as the rate-determining step in the gas phase reaction. However, when the solvation effects are included in the calculations [60], C-C bond formation is identified as the rate-determining step. In this study, Izzo et al. [60] reported a new binding mode that was not considered in the aforementioned

studies (Figure 3.2 (c)). This new pathway involves the formation of a H-bond between the nitrogen of the thiourea catalyst and the dimethyl malonate. The proposed transition structure was calculated to be higher in energy than the transition state proposed by Pápai and co-workers [58]. Pliego [61] employed theoretical calculations that consider both gas-phase and solvent effects to determine the complete mechanism and free energy profile. The results suggest that the C-C bond formation step is rate-determining. Apart from these studies aimed at elucidating the reaction mechanism and enantioselectivity, there are no efforts specifically focusing on the design of a more active and selective catalyst for this reaction. Therefore, our catalyst database which includes 1565 cinchona, 21 β-isocupredine derivatives, and 10 cyclohexyl derivatives (Figure 3.3) is screened using CIPOC to discover more active and stereoselective catalysts for the target reaction.

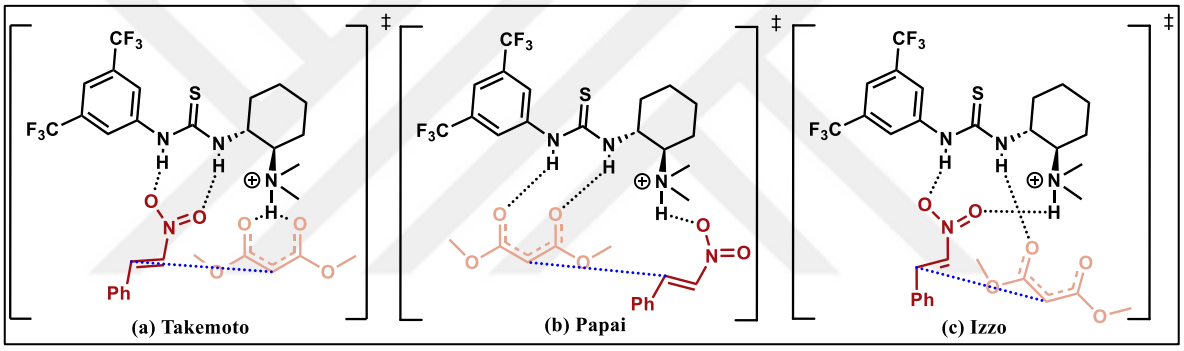


Figure 3.2. Newman projections (viewed along the forming C-C bond) of different binding modes reported (a) Takemoto [57], (b) Pápai [58], and (c) Izzo et al. [60]

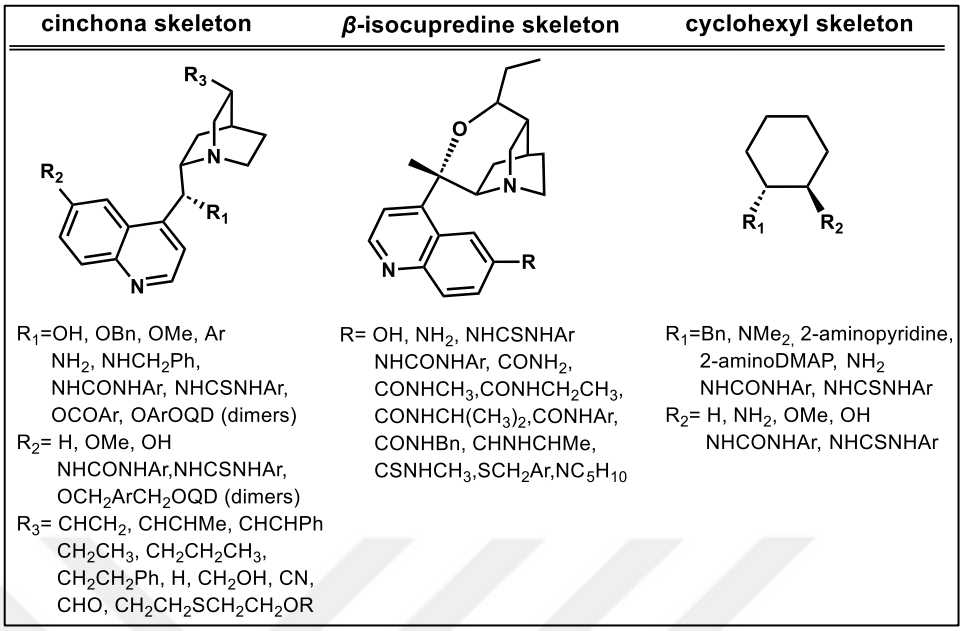


Figure 3.3. Studied catalyst library

3.2. COMPUTATIONAL APPROACH

CIPOC summarized in Figure 1.1 was employed to identify novel catalytic structures for the addition reaction of dimethyl malonate to *trans-β*-nitrostyrene. Initially, theoretical active site models (theozymes) including representative functional groups involved in the catalyst library (Figure 3.3) were generated at the B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31G(d) level to evaluate their catalytic performances, optimal 3D arrangements for maximum TS stabilization taking into account their binding mode preferences. Based on the lowest energy theozymes catalytic group maps (CGMs) were generated and diversified followed by their screening in the conformer library of the catalyst database. Final step involves the evaluation of hit catalyst candidates using QM calculations. The details of each step is described in the following subsections. Complete methodology for all QM calculations is given in subsection 3.2.5.

3.2.1. Theozyme Generation

Theozymes created with QM transition state optimization in the presence of catalytic functional groups provide valuable quantitative information on the degree of individual

atomic interactions' contribution to catalysis and transition state stabilization [62]. It basically is used to explore the optimal 3D arrangement of catalytic functional groups for maximum transition state stabilization. Theozymes have shed light on the factors that contribute to catalysis for many different reactions, and the insights gained have paved the way for the development of computationally designed enzymes [63] and spirocyclic organocatalysts [6].

3.2.1.1. Theozymes Based on Different Catalytic Groups

We generated theozymes using different catalytic groups to understand their effect on the activation barrier of the rate determining C-C bond formation step. Three main theozymes which are shown in Figure 3.4 were constructed using different catalytic groups for the most inclusive catalytic functional groups.

First model (Figure 3.4 **Theozyme 1S**) consists of *tert*-amine and thiourea groups and is created based on the cinchona/*β*-isocupredine derivatives and Takemoto's catalyst, which is the most efficient organocatalyst for the target reaction identified so far [56, 57]. Other model (Figure 3.4 **Theozyme 2S**) consists of 2-aminoDMAP and thiourea groups and is constructed based on the studies of Tanyeli and co-workers showing the efficiency of 2-aminoDMAP superbase motif to access sulfonamides [54] and squaramides [55]. Another model consisting of 2-aminoDMAP and urea groups (Figure 3.4 **Theozyme 2O**), is urea analogue of **Theozyme 2S**.

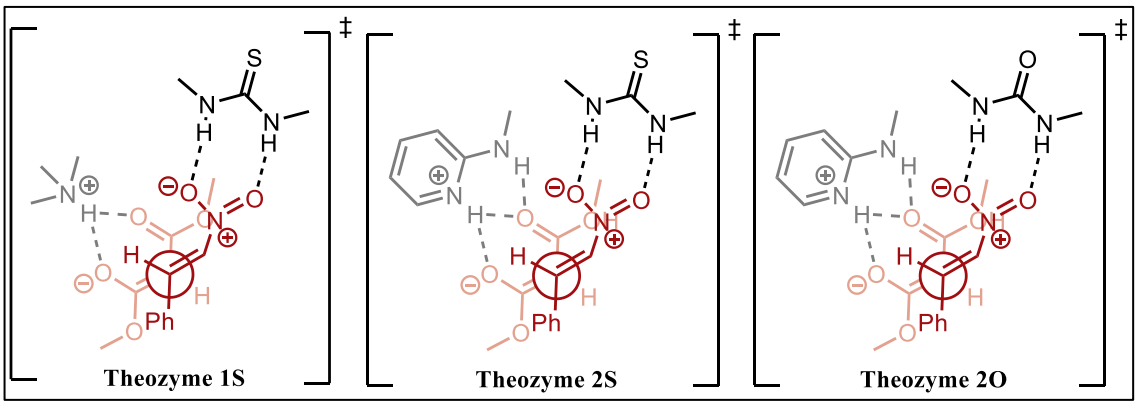


Figure 3.4. Newman projections (viewed along the forming C-C bond) of the theozymes with different catalytic groups

3.2.1.2. Theozymes Based on Different Binding Modes

We generated theozymes using different binding modes to understand their effect on the activation barrier of the rate determining C-C bond formation step. Two main theozymes which are shown in Figure 3.5 were constructed based on proposal by Takemoto (**Binding mode A** (**BMA**)) and Pápai (**Binding mode B** (**BMB**)). Although Izzo [60] proposed an alternative binding mode, it has not been considered in this study as it is shown to be energetically unfavorable.

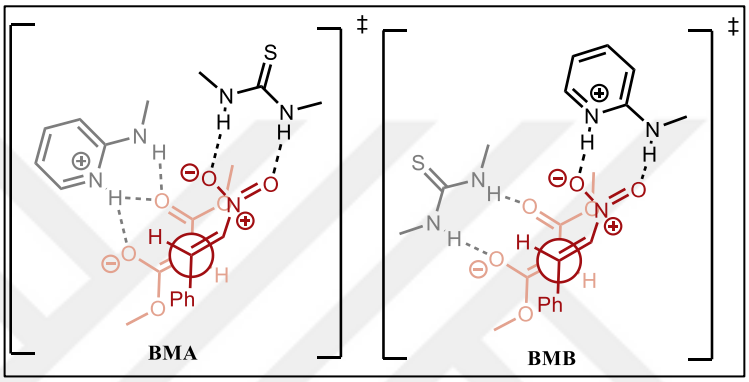


Figure 3.5. Newman projections (viewed along the forming C-C bond) of the theozymes with different binding modes

Enantiotopic faces are considered to examine selectivity. The enantioselectivity was explored at the C-C bond formation step through the nucleophilic attack of the enolate to the *Si* face or *Re* face of *trans-β*-nitrostyrene (Figure 3.6).

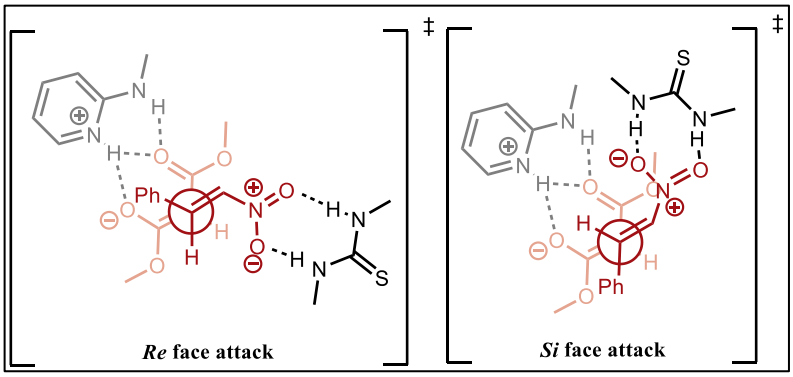


Figure 3.6. Nucleophilic effect of the enolate to the *Si* face or *Re* face of *trans-β*-nitrostyrene

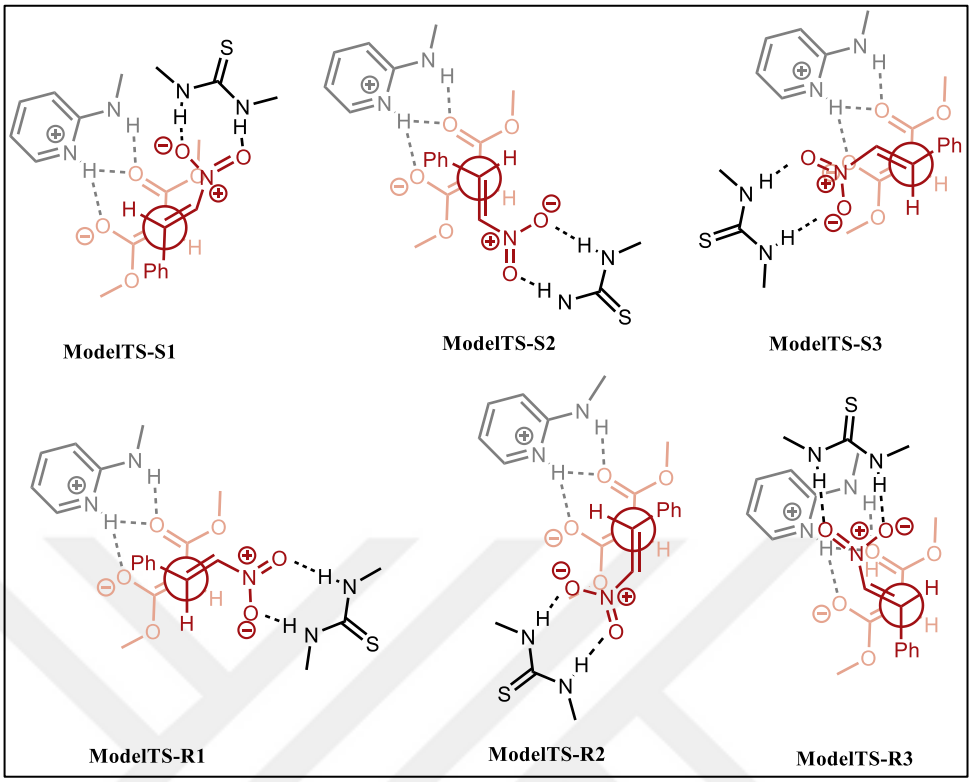


Figure 3.7. Relative orientations of representative catalytic functional groups

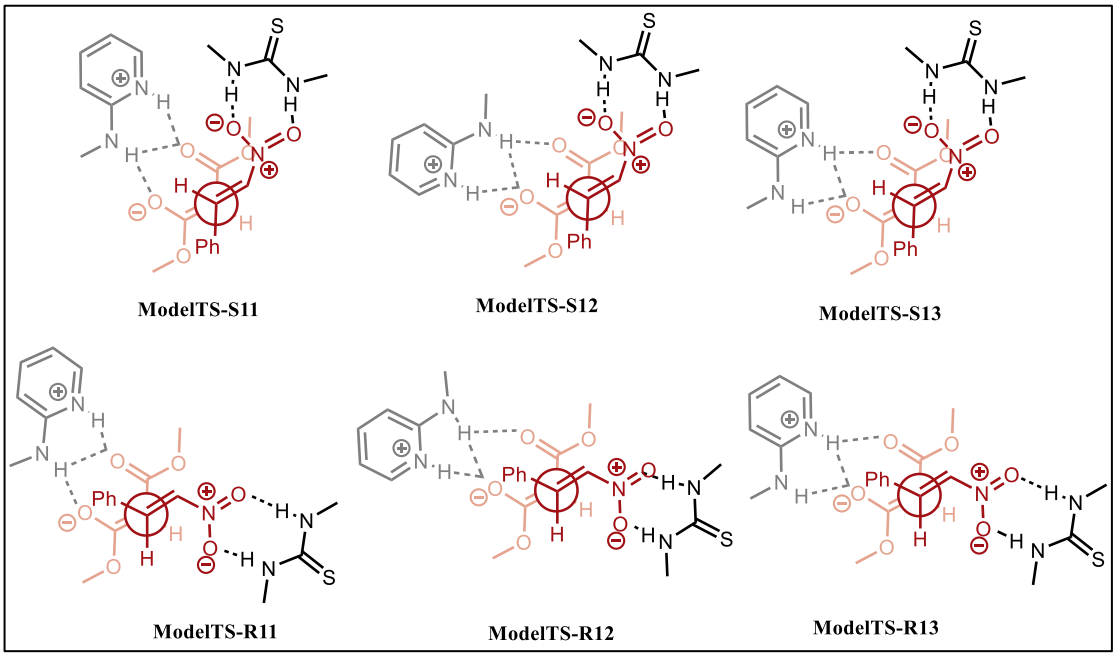


Figure 3.8. Relative orientations of the 2-amino-pyridine group for *Si* and *Re* face model

The next step involved the diversification of transition states via rotation around the forming C-C bond, which determines the relative orientations of catalytic functional groups (Figure 3.7).

For example, diversification for **Theozyme 2S BMA** gives a total of 24 TSs, 12 of which are enantiomers. Located TS structures with promising activation barriers for the C-C bond formation step were used to generate catalytic groups maps (Figure 3.7 & 3.8).

3.2.2. Creating Catalytic Group Map & Pharmacophore Search

A catalytic group map (CGM) was constructed using MOE [64] with pharmacophore-like features (Figure 3.9), such as H-bond acceptor and donor, aromatic ring centers, etc., based on the quantum mechanically optimized theozymes. An inclusive catalytic group map is necessary to ensure matching with three different catalyst skeletons and to avoid missing any matching possibilities. For screening, a partial matching mode was applied to include all our catalyst database (Figure 3.2) with three different skeletons. Partial matching mode is defined as a match of at least two, but not all of pharmacophoric features. Two nitrogens of HBD (alcohol/amine/urea/thiourea) were represented as H-bond donors. One nitrogen of HBA (amine) was represented as H-bond donor/acceptor to include both cyclohexyl and cinchona skeletons. The aromatic ring of HBA was chosen as an aromatic center. Selected features for CGMs (Table 3.1) were used to build the 3D pharmacophore models (Figure 3.10).

Table 3.1. Selected pharmacophore groups for CGM (3D layout is shown in Figure 3.10)

Catalytic Group Map	**HBA Group**			**HBD Group**		
		Radius	**Feature**		**Radius**	**Feature**
CGM	F1	1.0	Don / Acc	F3	1.0	Don
	F2	1.4	Aromatic	F4	1.0	Don

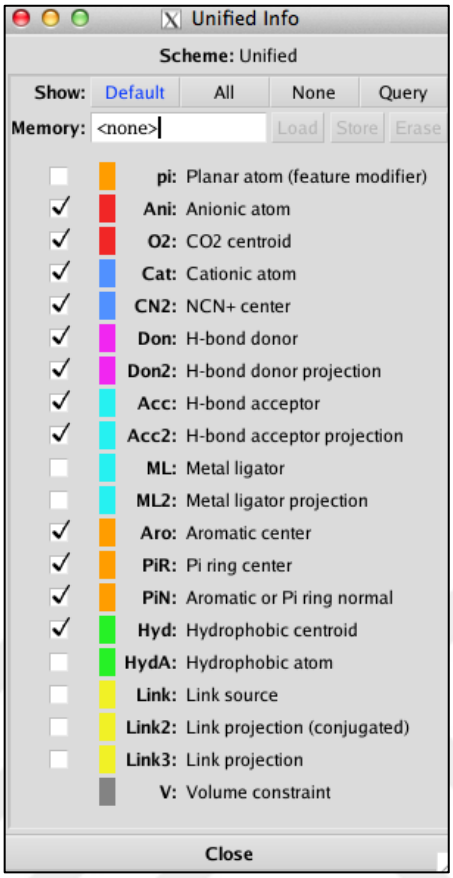


Figure 3.9. Pharmacophore-like features in MOE [64]

3.2.3. Generating Conformer Library of Organocatalyst

A conformer library for each catalyst included in the database (Figure 3.3) is generated to identify the catalyst that could provide the catalytic functional groups in the optimal 3D arrangements depicted in the low-energy theozyme.

A conformational search was performed for each explored catalyst with MOE (Molecular Operating Environment)[64] software with MMFF94x. Because including chirality in the explored space makes a stochastic search well-suited for organic molecules [65], stochastic search was preferred. The rejection and iteration limit was used as 1000 to achieve a sufficient conformational sampling for an initial assessment in a reasonable amount of computational time. Root mean square deviation (RMSD) of the two conformations are less than 0.25, these were considered duplicates. All of the catalysts and their conformations were saved in *.mdb* format as database and later employed for pharmacophore screening. All the conformations of matched catalysts were ranked according to the MM force field

calculations, and the lowest 20 of them were optimized with QM to calculate the lowest energy conformer of each studied catalyst.

3.2.4. Matching the Theozyme

The catalyst conformer library was screened against the CGM to identify the conformers that can stabilize the determine theozyme structure (Figure 3.10). The resulting hit conformers were primarily sorted by their matching scores (RMSD) and their deviations from their optimum catalytic geometries. Partial matches were allowed to obtain all possible catalytic candidates.

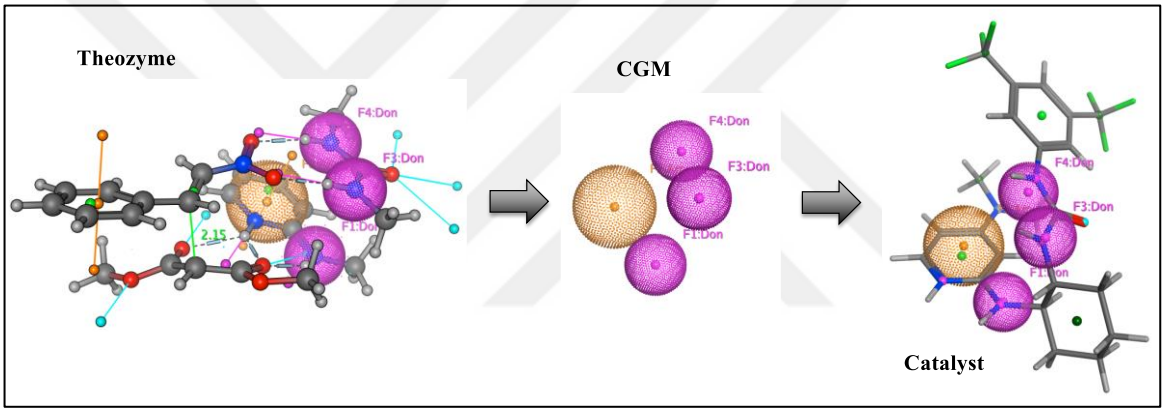


Figure 3.10. Generation of catalytic group map (CGM) using theozyme and matched conformer of bifunctional catalyst, orange sphere=aromatic center, pink sphere= hydrogen-bond donors

3.2.5. Quantum Mechanical Calculations

On the quantum mechanically optimized theozyme, replacing the model catalytic groups with the matching conformer of the catalyst gives the transition structure in the presence of the chiral bifunctional catalyst (Figure 3.11). Finally, the rates and selectivity of the leads were evaluated using QM calculations.

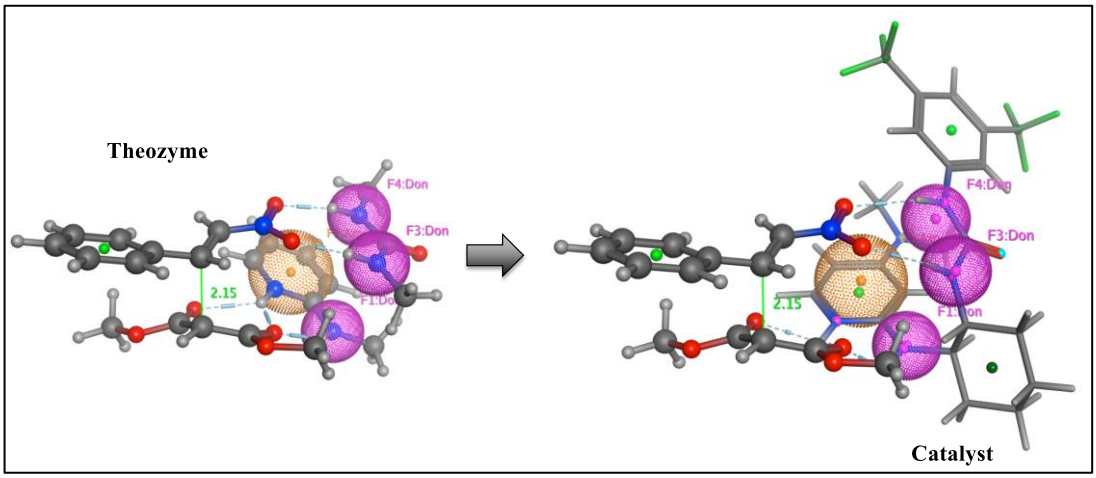


Figure 3.11. Matching the catalytic active site model, orange sphere=aromatic center, pink sphere= hydrogen-bond donors

All QM calculations were carried out using Gaussian16 [66]. Due to the large number of atoms typically involved, the standard B3LYP/6-31G(d) level is often used to create theozyme. Theozymes calculated at this level for different catalytic reactions and unnatural substrates were compared with protein crystal structures, and it was confirmed that the quantum mechanical active site successfully represented the placement of catalytic groups in the enzyme active site [67, 68]. Therefore, it is expected that theozymes should also successfully identify the placement of catalytic groups targeted in organocatalysts. In addition, Houk and his group reviewed over 300 studies conducted between 2000 and 2010 on the mechanisms of organocatalytic reactions and the fundamentals of chiral catalyst selectivity [1], showing that the DFT's standard level, B3LYP/6-31G(d), successfully explains experimental activity and selectivity. So, B3LYP/6−31G(d) was used to optimize all reactants, transition states, and products [13, 14, 69, 70] including Grimme's correction for better treatment of dispersion interactions. Due to the presence of aromatic motifs, where dispersion interactions are particularly important, a dispersion correction energy was necessary for energy calculations to better account for π-π interactions, and the molecular electronic energies were determined using dispersion-corrected DFT. This involved combining the self-consistent field (SCF) energy obtained from KS-DFT calculations with the standard atom pairwise London dispersion energy derived from D3 theory [71]. Single-point energy calculations were computed using the B3LYP-D3/6-311+G(d,p) level of theory. The intrinsic reaction coordinate (IRC) [72, 73] method was used to determine

minimum energy paths from the transition states to verify the reactants and products. Since the system being studied contains zwitterionic intermediates, the effect of the solvent was taken into account both in the geometry optimization and in the calculation of single-point energies. Solvent effects were considered with the SMD model [74] as implemented in Gaussian16. Toluene was used as the solvent to represent the experimental conditions (Figure 3.1). Frequency calculations were used to verify all stationary points as minima or saddle points and to calculate thermal corrections at 298.15 K. The energy corrections to enthalpy and free energy, which accounted for zero-point vibrational energies, were applied to the molecular electronic energies. The calculations were performed at 298.15 K using the harmonic oscillator approximation. Truhlar's quasi-harmonic correction was utilized, and all frequencies below 100 cm^{-1} were adjusted to 100 cm^{-1} to minimize errors in entropy estimation associated with low-frequency vibrational modes [75]. All energies reported in terms of Gibbs free energies with quasi-harmonic in kcal/mol. All distances are given in Å. Carbon atoms are shown in light grey, the hydrogens in white, oxygens in red, chlorine in green, sulfur in yellow, and nitrogen in blue in the figures.

Enantiomeric excess (ee %) was calculated by assuming that the final ratio between the enantiomers is determined by the Boltzmann distribution of TSs leading to each enantiomer. The ee % was calculated using the equation 3.1 [76].

$$ee\ \% = \frac{1 - e^{-\frac{\Delta G^{\neq}{}_{R/S}}{RT}}}{1 + e^{-\frac{\Delta G^{\neq}{}_{R/S}}{RT}}} \times 100 \qquad (3.1)$$

$\Delta G^{\neq}{}_{R/S}$ shows the Gibbs free energy difference between *R* and *S* enantiomeric transition states. *T* is temperature in Kelvin and *R* is universal gas constant 1.987 $cal.K^{-1}.mol^{-1}$.

3.3. RESULTS AND DISCUSSION

3.3.1. Theozymes Based on Different Catalytic Groups

We generated theozymes using different functional groups to understand their effect on the activation barrier of the rate-determining C-C bond formation step. **Theozyme 1S** consists of *tert*-amine (monodentate HBD) and thiourea groups (bidentate HBA). **Theozyme 1S**

particularly represents the functional groups involved in the cinchona/*β*-isocupredine derivatives and Takemoto's catalyst. **Theozyme 2S** consists of 2-aminoDMAP (bidentate HBD) and thiourea groups (bidentate HBA).

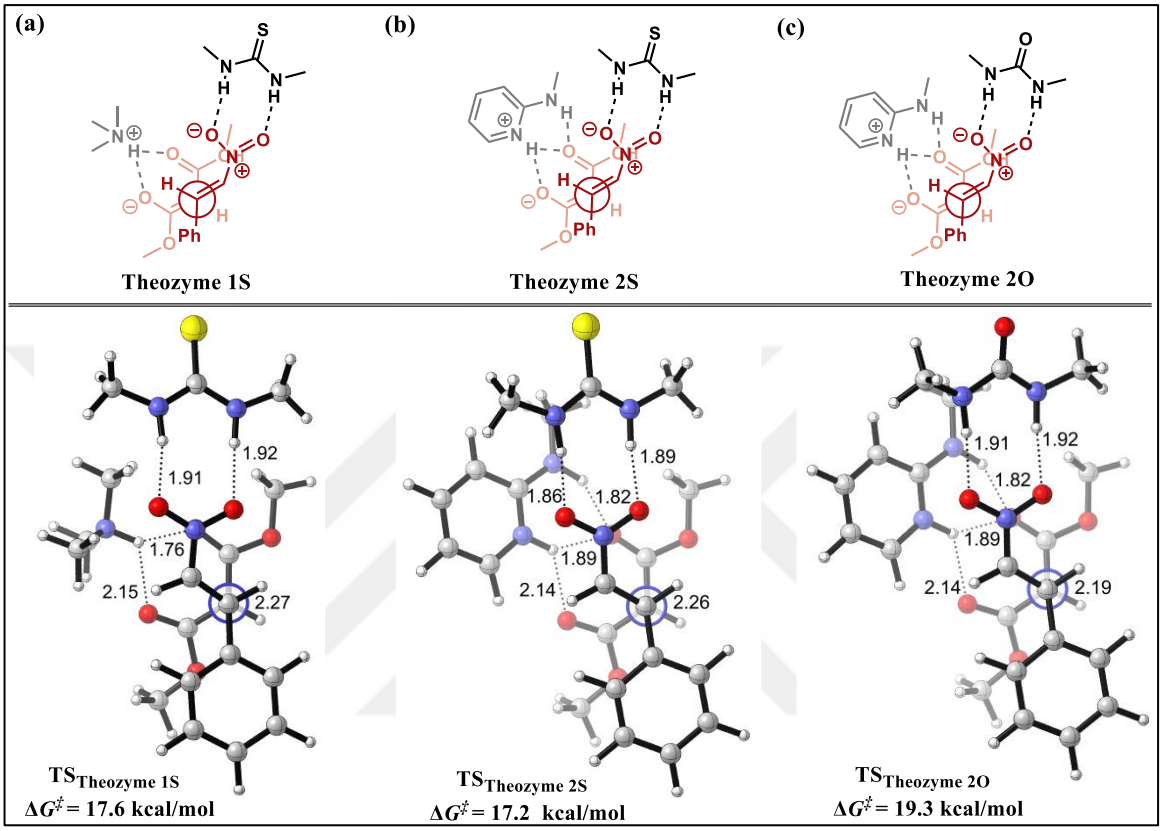


Figure 3.12. Newman projections (viewed along the forming C-C bond) of all theozymes for **BMA** and their activation free energies relative to the starting materials

Figure 3.12 shows the transition state geometries of the corresponding theozymes and their activation energies relative to the starting materials. Calculations show that **TS**$_{\textbf{Theozyme 2S}}$ is 0.4 kcal/mol lower in energy than **TS**$_{\textbf{Theozyme 1S}}$. These results suggest that bidentate H-bonding is only slightly favored over a monodentate H-bonding interaction for **BMA**.

TS$_{\textbf{Theozyme 2S}}$ is found to be 2.1 kcal/mol lower in energy than **TS**$_{\textbf{Theozyme 2O}}$. Although both of the theozymes include bidentate H-bonding, the thiourea group (Figure 3.12(b)) provides better TS stabilization than urea (Figure 3.12(c)). This result is in line with higher acidity of thiourea compared to urea (pKa(thiourea) = 21.1 and pKa(urea) = 26.9, in DMSO)[77, 78] indicating its better H-bonding ability.

3.3.2. Theozymes Based on Different Binding Modes

We generated active site models using different binding modes to understand their effect on the activation barrier of the rate determining C-C bond formation step. Two different binding modes of the lowest energy **Theozyme 2S** was considered (Figure 3.5). Results show that **Theozyme 2S BMB** as proposed by Pápai, is 2.2 kcal/mol higher in energy than **Theozyme 2S BMA**, which is the original proposal of Takemoto and co-workers (Figure 3.13). Therefore, further studies were conducted based on the low-energy transition state model (**TS**$_{\textbf{Theozyme 2S}}$) via **BMA**.

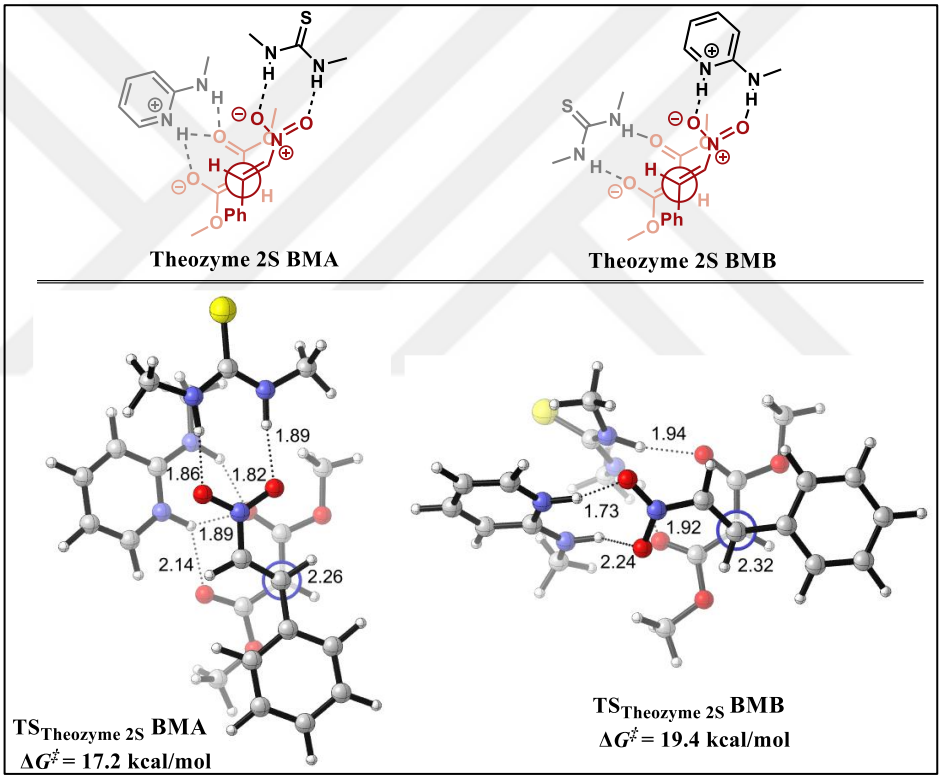


Figure 3.13. Newman projections (viewed along the forming C-C bond) of **Theozyme 2S** with different binding modes and their activation free energies relative to the starting materials

3.3.3. Matching and Evaluation of the Full Catalysts

The catalyst conformer library containing 1565 cinchona, 21 *β*-isocupredine, and 10 cyclohexyl derivatives (Figure 3.3) was screened against the CGM (Table 3.1) to identify

the catalysts that can stabilize the theozymes (Chapter 3.3.1 & 3.3.2) to discover more active and stereoselective catalysts for the target reaction. This procedure is explained in detail in Chapter 3.2.5. The most promising catalyst candidates with maximum catalytic group match avoiding the steric clash between substrates and catalysts are evaluated. Only the low energy structures are given and discussed in the text.

The most promising candidate based on QM evaluation results is the 2-aminoDMAP/urea (**2O**) derivative with a predicted activation energy 9.1 kcal/mol and enantioselectivity of 98.2% ee (Figure 3.14(a)). Intriguingly, the thiourea analogue (**2S**), which was predicted to be more stable by 2.1 kcal/mol in model calculations (Figure 3.12), is found to be disfavored by 2.5 kcal/mol in the presence of full catalyst (Figure 3.14(b)). It is unusual but not unique for catalysts based on urea to exhibit superior performance compared to their thiourea analogues [79, 80]. To better understand this unexpected result, a distortion/interaction analysis [81] was performed. The distortion/interaction analysis is a tool used to examine activation barriers that determine reaction rates. In the case of bimolecular reactions, the activation energies are the combined energies needed to distort the reactants into the geometries they have in TS, as well as the interaction energies between the two distorted molecules. The energy required for distorting the molecules is referred to as the distortion energy, which is the primary contributor to the activation barrier. The TS is reached when this distortion energy is overcome by the stabilizing interaction energy. By examining the changes in these energies along the reaction coordinate, valuable insights can be gained regarding the factors that control reactivity [81].

Figure 3.15 shows the distortion/interaction analysis of the full catalysts and model systems involving urea and thiourea. For both systems as the interaction energies for urea and thiourea analogues are essentially the same, the distortion energy determines the activity of catalyst.

In the model system, the catalyst distortion energies are similar because the catalytic units are treated independently, whereas lower distortion of the substrate in the thiourea favors **Theozyme 2S** over **Theozyme 2O**. The lower distortion of substrates in the thiourea model can be explained by forming C-C bond lengths in the located TSs. Upon formation of the C-C bond, sp2 hybridized planar carbons of the substrates transforms into tetrahedral sp3 carbons. Forming C-C bond in $\mathbf{TS}_{\mathbf{Theozyme\ 2S\ is}}$ 2.26Å compared to 2.19Å in $\mathbf{TS}_{\mathbf{Theozyme\ 2O}}$ indicating an earlier TS, which results in less distorted substrates.

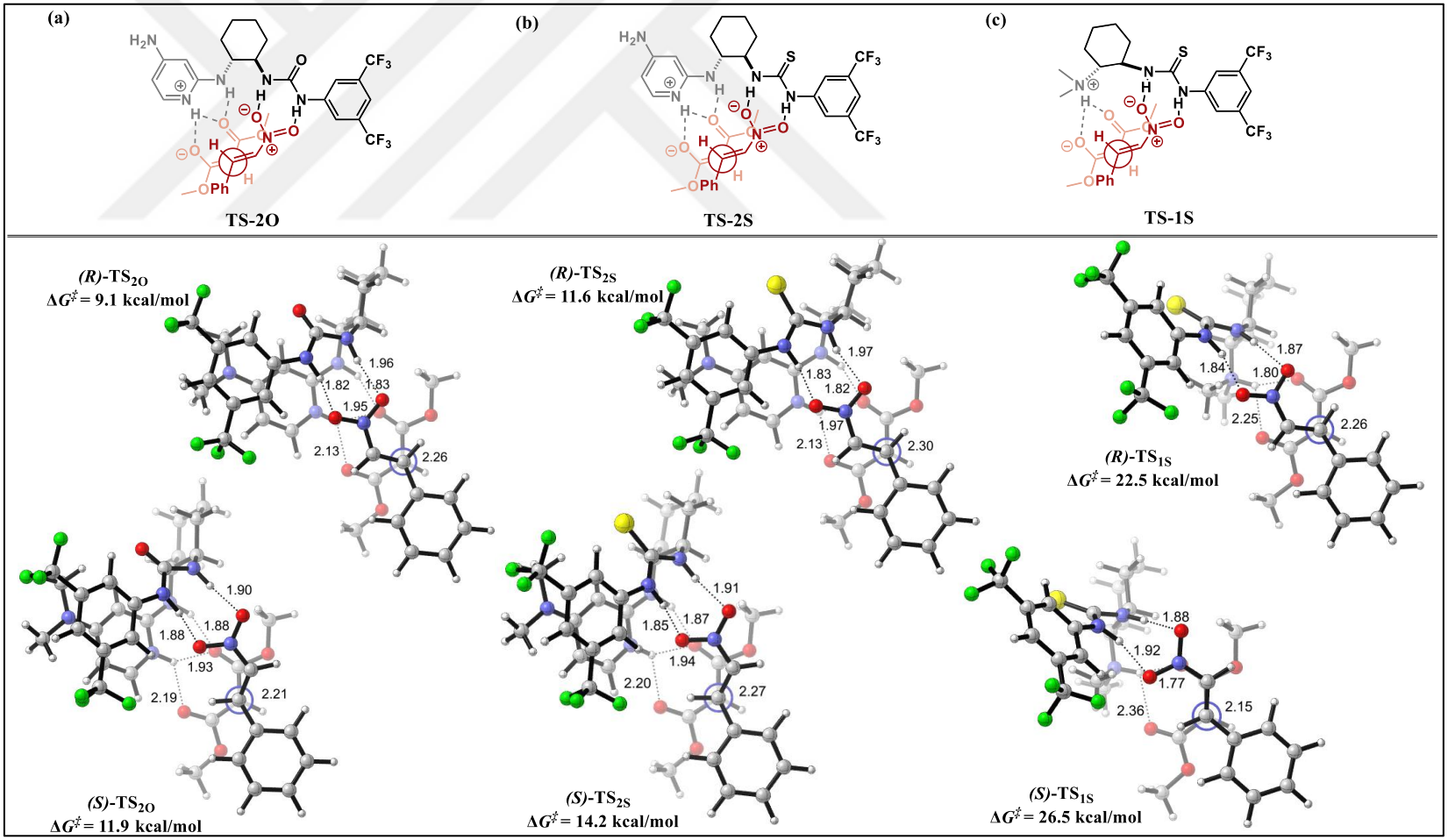


Figure 3.14. Newman projections (viewed along the forming C-C bond) of the transition states for *Re* and *Si* face attack in the presence of **2O, 2S**, and **1S** catalysts and their activation energies

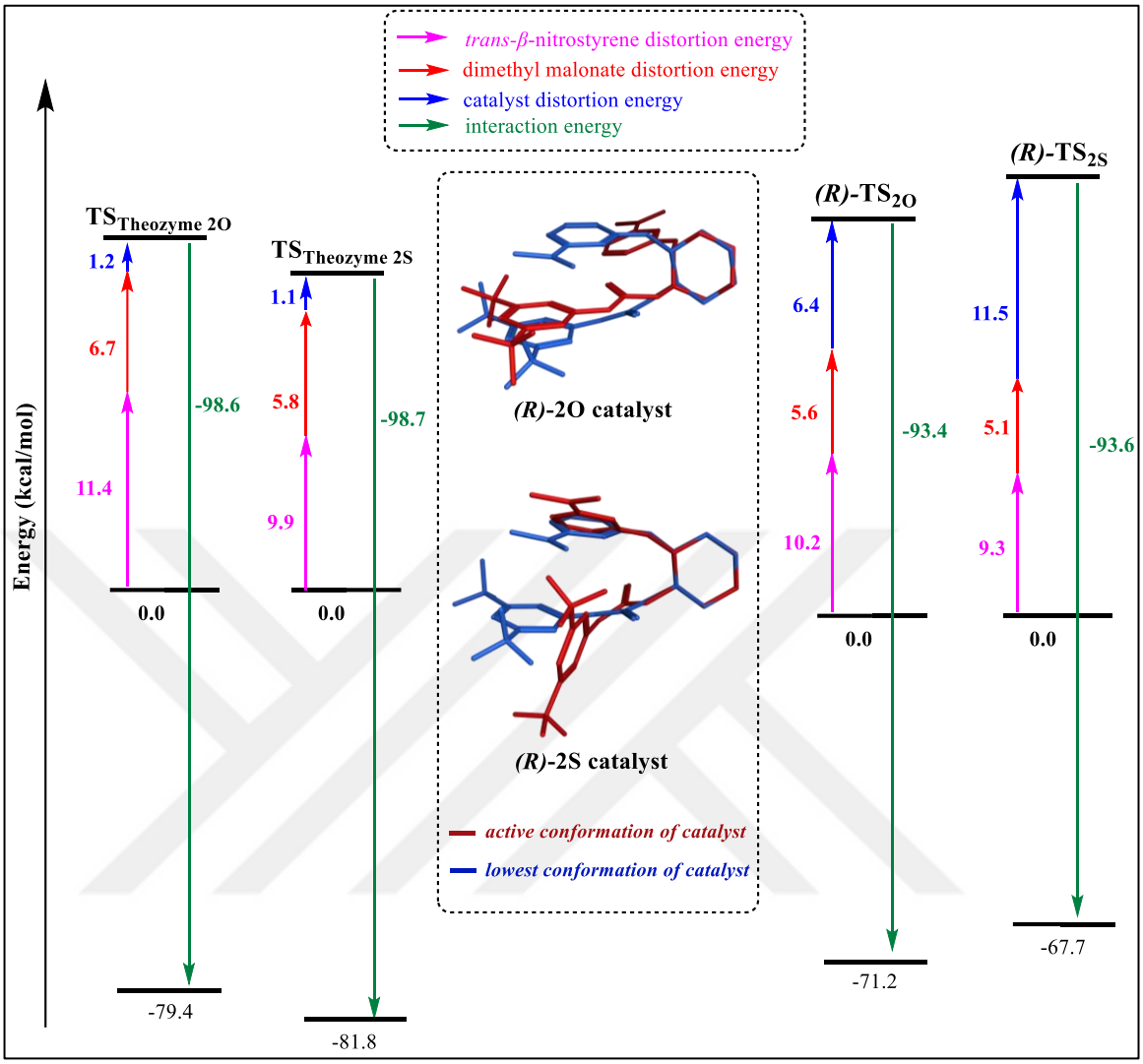


Figure 3.15. Distortion/Interaction Analysis of ***(R)*-TS2O** & ***(R)*-TS2S** and their theozyme transition states

For the full catalyst, when the distortion energy of substrates are compared, thiourea catalyzed transition state has 1.4 kcal/mol lower distortion energy of substrates (*trans-β*-nitrostyrene and dimethyl malonate). It is clear that ***(R)*-TS2S** is more favorable according to the distortion of substrates as in the model system. However, the reaction in the presence of the **2S** catalyst is less active (2.5 kcal/mol) than the **2O** catalyst (Figure 3.14). In the full catalytic system, however, the main difference is the distortion energy of the catalyst. Similar to the model system the distortion energy of substates is in favor of the thiourea catalyst (**2S**), yet, the distortion of the thiourea catalyst has 5.1 kcal/mol higher energetic cost than the urea analogue (**2O**), which results in a much lower activation barrier for the 2-

aminoDMAP/urea (**2O**) catalyst. These results suggest that the energetic cost of distorting the 2-aminoDMAP/thiourea (**2S**) catalyst to the active conformation plays a significant role in a destabilization of ***(R)*-TS$_{2S}$** with respect to ***(R)*-TS$_{2O}$**. This leads to an unexpected decrease in the reactivity of the **2S** catalyst.

The free energy profile for the dimethyl malonate addition to the *trans-β*-nitrostyrene reaction in the presence of a **2S** catalyst is given in Appendix, Figure C.1.

To examine the catalytic performance of the *organo-superbase* motif (2-aminoDMAP) the **2S** catalyst, and enantiomeric transition states in the presence of **1S** (Figure 3.14(c)), in which 2-aminoDMAP group is replaced with dimethyl amine **1S** is compared with **2S** (Figure 3.14(b)). It was found that a combined effect of bidentate H-bonding and a stabilizing π-π interaction strongly favors ***(R)*-TS$_{2S}$** ($\Delta G^{\neq}$=10.9 kcal/mol) relative to (***(R)*- TS$_{1S}$**. These interactions contribute differently to the stabilization of molecular structures, and the understanding individual contributions of these interactions to the stabilization of corresponding structures requires a detailed analysis. For this purpose, two different conformations of 2-amino-pyridine **(3S**) were studied. While ***(R)*-3S** takes advantage of both π-π stacking and H-bonding. ***(R)*-3S-conf** profits only from π-π stacking and double H-bonding is not allowed. It is clearly seen that conformation of ***(R)*- TS$_{3S}$** instead of ***(R)*- TS$_{3S\text{-}conf}$** decreased the activation free energy by 4.9 kcal/mol because of the bidentate H-bonding. So, approximately 4.9 kcal/mol is due to bidentate H-bonding and approximately 6.0 kcal/mol is due to π-π stacking. Moreover, the activation free energy of dimethyl amine (***(R)*-1S**) is 9.0 kcal/mol higher compared with 2-amino-pyridine (***(R)*-3S-conf**). The effect of stabilizing π-π interaction strongly favors ***(R)*-3S**. All computed results revealed that the *Re* face attack is more stable than the *Si* face attack in the presence of all studied catalysts. Thus, only *Re* face results were included in the comparison.

To understand the effect of dimethyl amine substituent, ***(R)*-2S** and 2-amino-pyridine ***(R)*-3S** were compared. Their structures and activation energies relative to the starting materials are given in Figure 3.16. The activation free energy of ***(R)*-TS$_{2S}$** is 2.0 kcal/mol higher than ***(R)*-TS$_{3S}$** (Figure 3.16(b)&(c)).

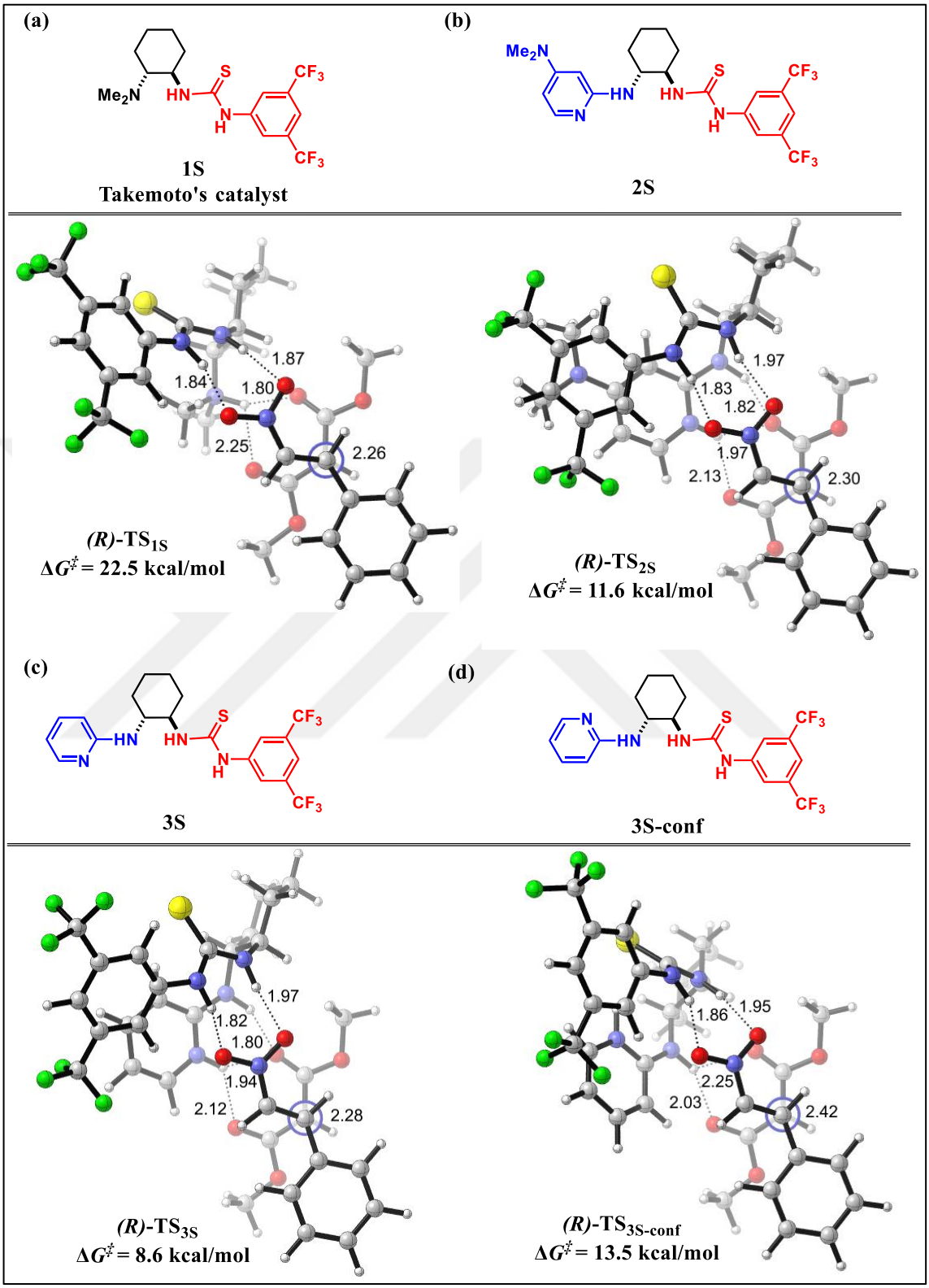


Figure 3.16. Newman projections (viewed along the forming C-C bond) of the transition states for *Re* face in the inclusion of **1S**, **2S,** and **3S** and their activation energies

Our full catalyst calculations were carried out via **BMA**, which is found more favorable for model theozymes (Figure 3.13). However, in the previously reported theoretical studies [58–61], results indicate that **BMB** is more favorable binding preference in the presence of Takemoto's catalyst (**1S**). To examine binding preferences of the 2-aminoDMAP containing full catalyst, CGM was generated based on the **Theozyme 2S BMB** and screened against catalyst library (Figure 3.3). Our calculations show that ***(R)*-TS$_{2S}$ BMB** is 4.5 kcal/mol higher in energy than ***(R)*-TS$_{2S}$ BMA**. Results suggest that **BMA** is more favorable pathway than **BMB** in the presence of catalyst which interacts with bidentate H-bonding. Because both computed results revealed that the *Re* face attack is more stable than the *Si* face attack, only *Re* face results were included in this comparison (Figure 3.17).

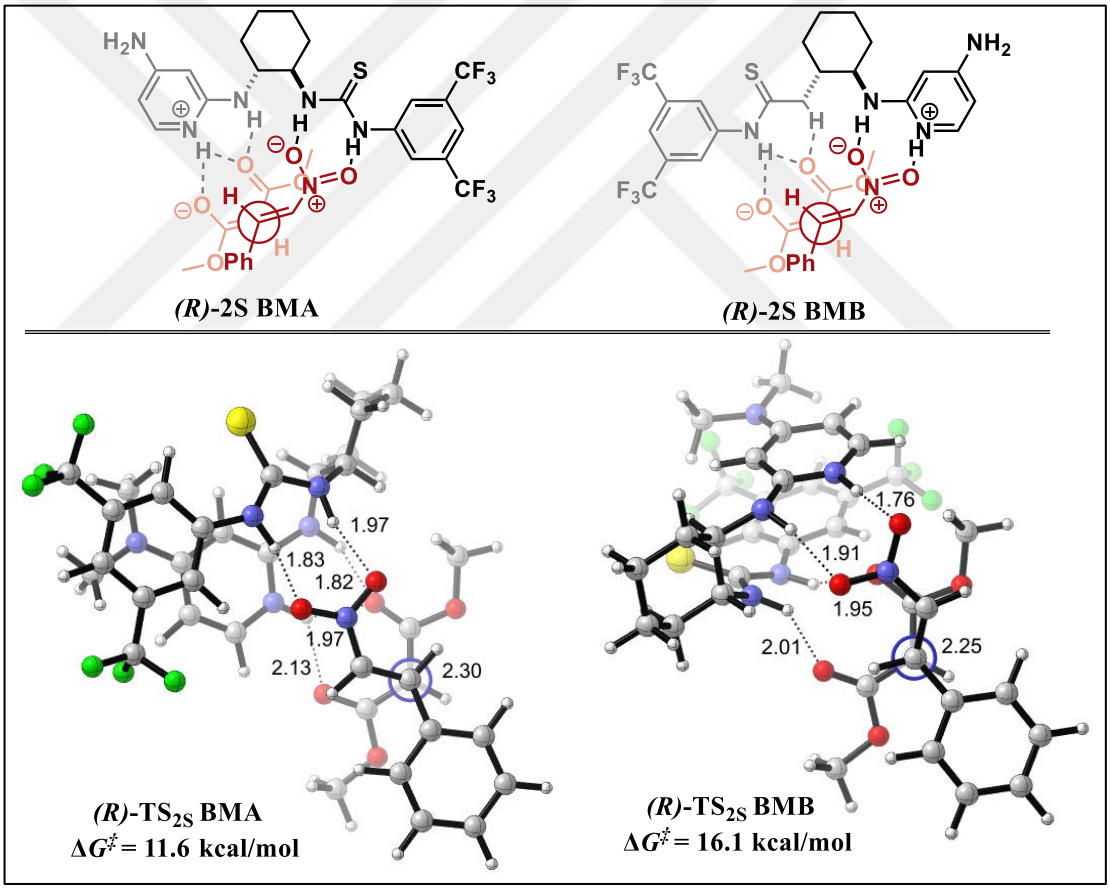


Figure 3.17. Newman projections (viewed along the forming C-C bond) of the transition states for *Re* face in the presence of **2S** with different binding modes and their activation free energies

3.3.4. Catalytic Performance of Proposed Catalysts

The **2S** and **2O** catalysts, which were identified through screening in the catalyst library include 1565 cinchona, 21 *β*-isocupredine derivatives, and 10 cyclohexyl derivatives (Figure 3.3) using the CIPOC protocol and determined to be more active than Takemoto's catalyst (**1S**) based on QM calculations, are proposed for experimental testing to our collaborating Prof. Cihangir Tanyeli's Research Group. Additionally, it is revealed theoretically that these catalysts (**2S/O**) can compete with Takemoto's catalyst (**1S**) in terms of selectivity. Thus, Tanyeli and co-workers concentrated on urea (**2O**) and thiourea (**2S**) of 2-aminoDMAP precatalyst to compare their (**2O/S**) activities with those of Takemoto (**1S**) in the same enantioselective transformation under identical reaction conditions [57]. When they ran conjugate addition of *trans*-*β*-nitrostyrene and diethyl malonate, both thiourea (**2S**) and urea (**2O**) acted two- to eight-fold faster (Table 3.3); and the urea catalyst (**2O**) gave complete conversion in a couple of hours affording distinctively higher enantioselectivity and yield, comparable to **1S**.

Table 3.2. Experimental comparison of the catalytic performances of 2-aminoDMAP/(thio)urea (**2S and 2O**) and those of Takemoto's (**1S**)

entry	catalyst	time[b] (h)	yield[c] (%)	ee (%)
1	**1S**	24	86	93
2	**2S**	12	80	75
3	**2O**	3	89	91

[a]Reactions were run in 0.2M concentration of *trans*-*β*-nitroalkene, (0.2 mmol) with 2 equivalent of diethyl malonate and 10% catalyst loading in toluene, rt. [b]Time for complete conversion. [c]Isolated yields.

3.4. CONCLUSION AND FUTURE WORK

CIPOC identified a novel chiral 2-aminoDMAP/(thio)urea catalysts (**2S** and **2O**) as being superior than Takemoto's catalyst (**1S**) for the conjugate addition of dimethyl malonate and *trans-β*-nitrostyrene. Calculations not only revealed the activity of this chiral 2-aminoDMAP/(thio)urea is attributed to the dual-function of 2-aminoDMAP unit (double H-bonding and π-stacking interactions) but also predict urea analog (**2O**) to be more active than the thiourea (**2S**). Results with the model catalyst have suggested that the thiourea-containing catalyst (**2S**) is highly favored than the urea-containing catalyst (**2O**). However, the inclusion of chiral bifunctional catalysts shows the importance of catalyst distortion in determining the activity. The proposed catalysts (**2O** and **2S**) was tested experimentally by Tanyeli's Research Groups for the target reaction. Experimental results are in excellent agreement with the computational predictions. This study not only demonstrates the applicability of CIPOC to determine the catalytic activity and selectivity of different multifunctional organocatalysts, but also shows its predictive power allowing easy and cheap identification of potential organocatalyst candidates for a selected target reaction.

The complete mechanism and free energy profile of Takemoto's catalyst (**1S**) have been previously examined in detail and the rate-determining step has been determined [61]. However, considering the different preferred binding preferences in Takemoto's (**1S**) and discovered catalysts (**2O** and **2S**), the full catalytic cycle can be further investigated in detail for both BMA and BMB binding preferences in the proposed catalysts. Additionally, in order to understand the effect of dimethyl amine substituent, ***(R)*-2S** and 2-amino-pyridine ***(R)*-3S** were compared in Chapter 3.3.4. Since the activity of 2-amino-pyridine containing catalyst ***(R)*-3S** is found to be more favorable theoretically, it can be experimentally tested for its catalytic activity for the target reaction.

4. EXPLORING THE MODE OF ACTION OF DIMERIC QUINIDINE CATALYZED ENANTIOSELECTIVE AMINOOXYLATION REACTION

In recent years, numerous highly effective synthetic protocols have been established utilizing cinchona alkaloids as chiral catalysts [82–86]. While progress has been made in understanding the mechanism of action of these compounds, the development of novel catalytic strategies still relies mostly on experimental approaches. The application of computational methods to gain insights into their mode of action has been limited to a small number of cases. This is mainly due to the complexity of cinchona alkaloids, which makes modeling their catalytic properties a challenging task, particularly in the context of explaining stereochemical outcomes.

Several theoretical studies demonstrate how the employment of straightforward models for the catalyst subunits can yield significant insights into the mechanism by which cinchona derivatives operate, particularly regarding the preferred binding modes between the catalyst and substrate(s) [87]. Although these simplified models have proven to be valuable in elucidating the mode of action of cinchona derivatives and understanding their interactions with substrates in a complementary manner, it is insufficient for determining the activation mechanism and mode of action of multifunctional organocatalysts. Additionally, in the previous chapter, we observed how the simplified models failed to determine the activity of the full catalyst. Especially in dimeric structures, the challenge lies not only in the complexity of the structure similar to that observed in monomeric cinchona structures, but also in the high degree of conformational flexibility. Although their capability for multisite binding and conformational flexibility lead to difficulties in investigations, these features are also the secret to their success in organic chemistry.

In this chapter of the thesis, the proposed protocol, CIPOC, was applied to investigate the ambiguous active conformations and mode of action of complex multifunctional dimeric quinidine organocatalysts with a high degree of conformational freedom for the aminooxylation reaction of oxindole with nitrosobenzene. In the reaction of *N*-benzyl-protected 3-substituted-2-oxindole with nitrosobenzene, which is often studied with cinchona alkaloid derivatives as catalysts [88–92]. Barbas and co-workers reported that this

reaction gives high yields of aminooxylation product in the presence of the dimeric derivative of quinidine [91] (Figure 4.1). Barbas' catalyst is the first example of an organocatalyst that can create a tetrasubstituted chiral center at the C3 position of oxindoles through asymmetric synthesis and the formation of a C-O bond in the reaction of oxindoles with nitrosobenzene. Despite its success, the mechanism of the enantioselective aminooxygenation catalyzed by the dimeric quinidine-based catalyst remains undetermined due to the complex structure of the catalyst. In this study, the impact of catalysts on the mechanism and selectivity has been thoroughly investigated and explained. The essential interactions and binding preferences of catalysts in critical transition states have been examined, and the factors that influence the stereoselectivity of the reaction are clarified. This work presents a new perspective on the challenges of exploring the mode of action and active conformations of complex multifunctional organocatalysts.

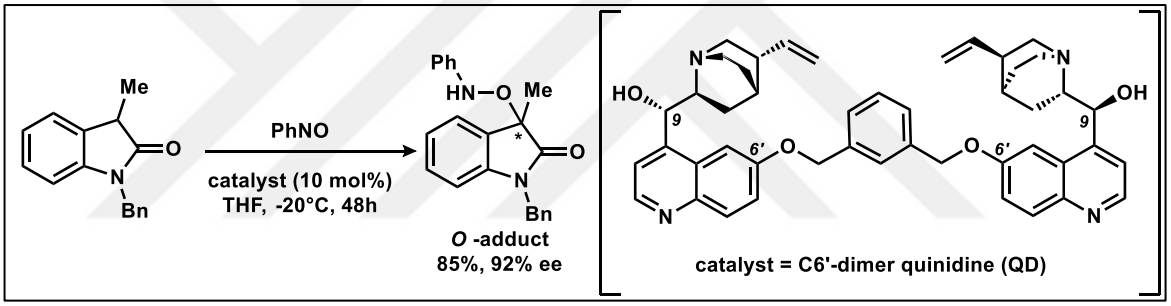


Figure 4.1. Reaction of *N*-protected 3-methyl-2-oxindole with nitrosobenzene [91]

4.1. INTRODUCTION

The development of chiral catalysts for the asymmetric synthesis of oxindole derivatives with a tetrasubstituted carbon at the 3-position has gained significant attention in recent years. This is due to the privileged nature of the oxindole skeleton, which serves as a core motif in numerous bioactive natural products and pharmaceutically active compounds. [93, 94]. For applying CIPOC in this chapter of the thesis, the reactions of 3-substituted 2-oxindoles with nitrosobenzene were selected as targets because of their attractive product distribution. It is possible to obtain two different products with different biological activities in this reaction (Figure 4.2). 3-substituted 3-hydroxy indole units achieved through aminooxylation are found in many natural products with a broad biological activity such as maremycin, convalutimidine, donaxaridine, dioxybrasinine, and celogentine K, and also in

pharmaceuticals such as SM-130686 and TMC-95A [95]. On the other hand, hydroxyamination products are found in the structure of potential drug candidates (NITD609, SSR-149415, and AG-041R) for the treatment of malaria and stress-related illnesses [95]. Even though obtaining the required product in high yield and purity by steering the product distribution in favor of one of the adducts is of critical importance, the reaction leads to different major products in the presence of different catalysts and the catalytic mode of action of cinchona alkaloids and how they affect the product distribution are not yet known.

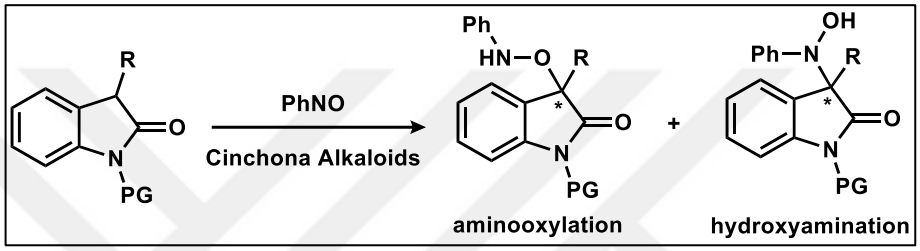


Figure 4.2. Reactions of 3-prochiral 2-oxindoles with nitrosobenzene

In the reaction of 2-oxindoles with nitrosobenzene, which is frequently studied in the presence of cinchona alkaloid derivatives as catalysts, interestingly, the main product obtained varies depending on the catalyst used. For instance, Chen and co-workers obtained hydroxyamination products with 90% yield by using quinine C6'-hydroxy derivative as the catalyst in the reaction of *N*-benzyl-protected 3-substituted-2-oxindole with nitrosobenzene [96]. In contrast, Barbas and co-workers reported 3-hydroxy oxindoles as aminooxylation products from the same reaction with 85% yield and 92% ee in the presence of a dimerized quinidine derivative at the C6' position, a.k.a. Barbas' catalyst [91] (Figure 4.1). On the other hand, Jia et al. achieved the highest yield of 91% in obtaining 3-amino-oxindoles by using chiral bifunctional tertiary amine thiourea derivatives as catalysts in the asymmetric hydroxyamination reaction [97]. Similarly, Zhang's group developed cinchona alkaloid derivatives that show *N*-selectivity in the reactions of *N*-unprotected 2-oxindoles [96]. Lastly, Companyó et al. [92] obtained *N*-adduct from *N*-unprotected 3-substituted-2-oxindoles and nitrosobenzene reaction catalyzed by bifunctional amino-thioureas. After five days aminooxylation product was observed in the inclusion of protecting group of the starting oxindole. They reported that protecting group of 3-substituted oxindole determines the regioselectivity of the reaction.

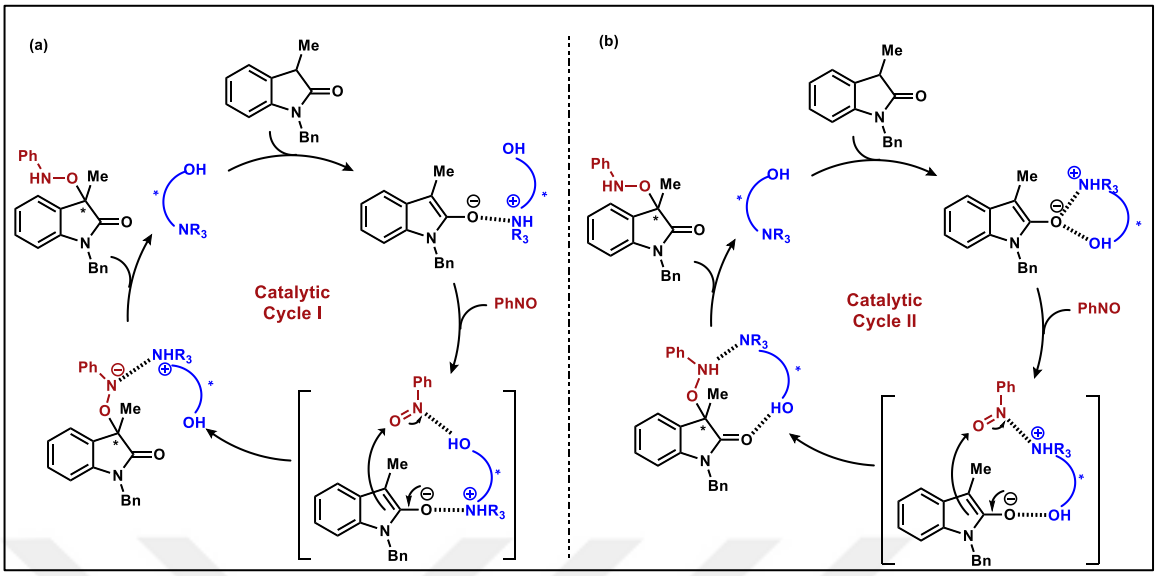


Figure 4.3. The proposed catalytic mechanism of the reaction between *N*-benzyl-protected 3-methyl-2-oxindole and nitrosobenzene in the presence of cinchona alkaloid-derived catalyst by (a) Barbas and co-workers [91] and (b) Çamlısoy's Thesis [98]

The proposed catalytic cycles of the reaction between 3-substituted 2-oxindoles and nitrosobenzene with a cinchona alkaloid-derived catalyst are given in Figure 4.3. The **catalytic cycle I** (Figure 4.3 (a)) proposed by Barbas and co-workers is assumed to begin with the deprotonation of the oxindole C3 methine proton by the catalyst's amino group. The protonated catalyst is assumed to form an ion pair with the enolate and direct the nitrosobenzene with its hydroxyl group. In the next step, proton transfer from the protonated amine catalyst to the intermediate results in the formation of aminooxylation product while releasing the catalyst. Although they could not explain the mechanism of the dimeric quinidine-catalyzed enantioselective aminooxylation, they propose **catalytic cycle I** (Figure 4.3 (a)). Barbas and co-workers mentioned that *N*-adducts were observed as minor product (<15%) occasionally, and the reaction was *O*-selective. **Catalytic cycle II** (Figure 4.3 (b)) also begins with the deprotonation of the oxindole C3 methine proton by the catalyst's amino group. In contrast to **catalytic cycle I**, the protonated amine of the catalyst activates the nitrosobenzene, paving the way to a low energy concerted addition proton transfer step, while the hydroxyl group of the catalyst stabilizes the enolate. This enolate then reacts with nitrosobenzene in an *N*- or *O*- selective fashion. The C-N or C-O bond forms simultaneously with the transfer of proton from the catalyst in a single step. Çamlısoy investigated cinchona

alkaloid catalyzed enantioselective reaction of 2-oxindoles with nitrosobenzene in her Ph.D. thesis [98] and reported this reaction in the presence of cinchona derivatives including alcohol functional group gives 15% ee. The product distribution disfavor the formation of the *O*-adduct, and the probability of the formation of *N*-adduct was calculated as 83%. In addition, the cinchona derivative including the thiourea functional group gives also hydroxyamination products (97%) in high enantioselectivities (100% ee). The results revealed that hydroxyamination is the major reaction pathway, which proceeds through a concerted addition-proton transfer mechanism. None of the catalysts in the database were identified to be capable of directing the product distribution towards the aminooxylation products.

Even though access to *O*-adducts from the reaction of oxindoles with nitrosobenzene in the presence of Barbas' catalyst is a great success, the mechanism of the reaction and how the catalyst provides a catalytic arrangement of the multiple functional groups in its structure to stabilize the rate-determining step remains unclear. Moreover, there is no computational strategy so far to explore the mode of action of such flexible complex multifunctional chiral catalytic structures. No computational study is performed with an organocatalyst of this size in the literature.

4.2. COMPUTATIONAL APPROACH

CIPOC summarized in Figure 1.1 was employed to investigate the ambiguous active conformations and mode of action of dimeric quinidine organocatalysts for the reaction of oxindole with nitrosobenzene. Initially, theozymes including representative functional groups involved in the dimeric quinidine catalyst (Figure 4.4) were generated at the B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31G(d) level to evaluate their catalytic performances, optimal 3D arrangements for maximum TS stabilization taking into account their binding mode preferences. Based on the lowest energy theozymes catalytic group maps (CGMs) were generated and diversified followed by their screening in the conformer library of the catalyst. Final step involves the evaluation of hit conformers of catalyst using QM calculations. The details of each step is described in the following subsections.

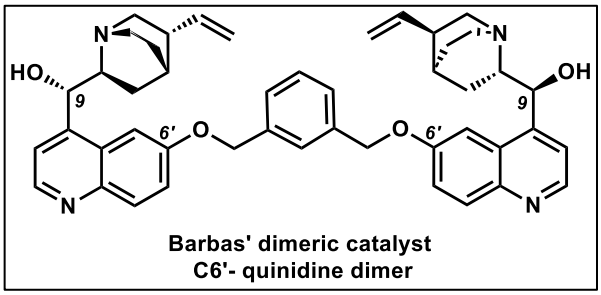


Figure 4.4. Barbas' dimerized quinidine derivative catalyst [91]

4.2.1. Theozyme Generation

Two main catalytic cycles for cinchona alkaloid-derived catalysts have been proposed (Figure 4.3). The **catalytic cycle I** (Figure 4.3 (a)) is the proposal of Barbas, in which enolate interacts with nitrosobenzene in an *O*-selective fashion, as a result of the directing effect of a hydroxyl group of the catalyst. On the other hand, in the **catalytic cycle II** (Figure 4.3 (b)) enolate interacts via hydrogen bonding with the hydroxyl group of the catalyst, and nitrosobenzene interacts with the protonated amine group of the catalyst. In this cycle bond formation and proton transfer occurs simultaneously, as previously mentioned in the introduction section (Chapter 4.1). Possible new binding preferences due to the dimeric structure of the catalyst were also taken into consideration, in addition to the binding preferences in the proposed catalytic cycles (Figure 4.3). The binding modes studied are summarized in Figure 4.6 [99].

In this chapter, while aminooxylation was named **A**, hydroxyamination was labeled as **H** in the nomenclature. In addition, the binding preferences of the **catalytic cycle I** (Figure 4.3 (a)) are named **1**, and the binding preferences of **catalytic cycle II** (Figure 4.3 (b)) as **2**. For example, **Model A1** (shown in Figure 4.5) signifies aminooxylation model TS of **catalytic cycle I**. The nomenclature was further modified to include a suffix: **1** for no additional hydroxyl group of catalyst interaction, **2** for additional hydroxyl group of catalyst interacts with nitrosobenzene, and **3** for additional hydroxyl group of catalyst interacts with oxindole after the point depending on the presence of the second hydroxyl group in the dimeric structure of the catalyst. For example, **A2.3** (shown in Figure 4.6) signifies aminooxylation TS of **catalytic cycle II** in which additional hydroxyl group of catalyst interacts with oxindole.

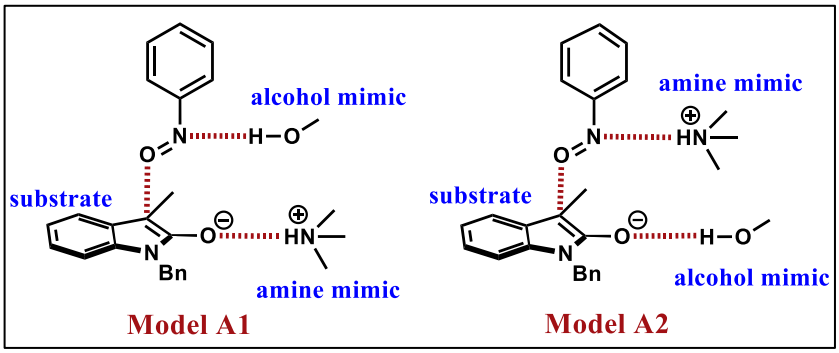


Figure 4.5. Theozymes of aminooxylation reaction for each binding mode

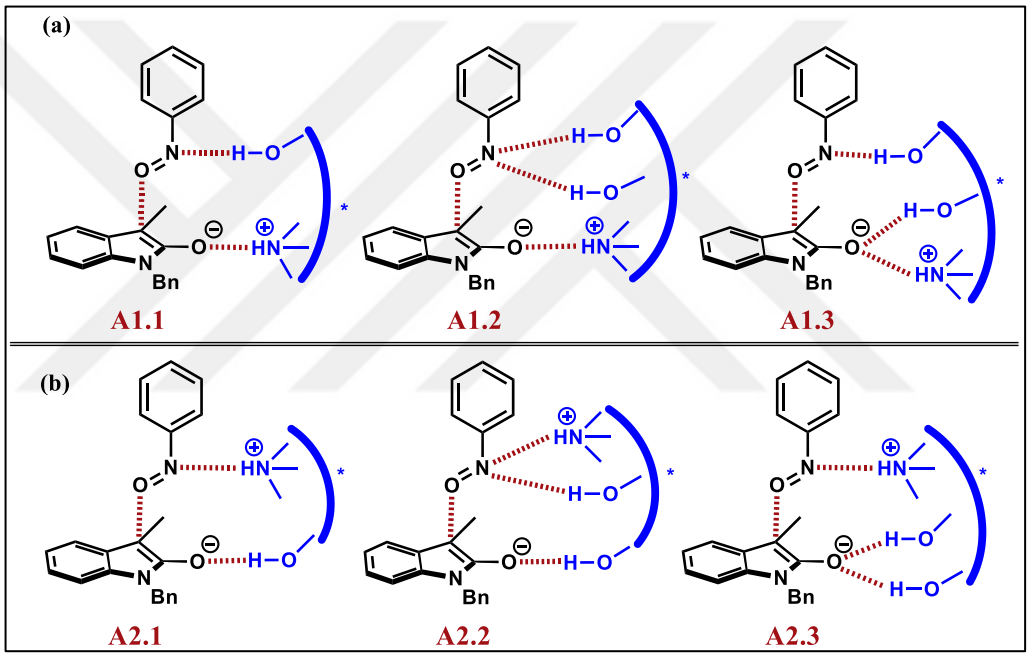


Figure 4.6. Proposed binding modes for cinchona alkaloid derived catalysts for (a) **catalytic cycle I** (b) **catalytic cycle II**

The theozymes were constructed using the catalytic functional groups (hydroxyl and amino groups) which are involved in the structures of the studied multifunctional dimeric organocatalyst (Figure 4.4). The theozymes for main binding modes, which are shown in Figure 4.6, were created by representing the hydroxyl group of the catalyst with one methanol and the amine group with one trimethylamine, as shown in Figure 4.5. These created models were diversified by rotating around the forming bond (C-N or C-O). Enantiotopic faces were also taken into account to examine enantioselectivity. The lowest

energy models were determined for each binding mode by optimizing them with quantum mechanical calculations. QM calculation details are explained in Chapter 4.2.5.

4.2.2. Creating Catalytic Group Map & Pharmacophore Search

The pharmacophore analysis was carried out using MOE [64]. Catalytic group maps (CGM) were constructed with pharmacophore-like features such as H-bond acceptor and donor, aromatic ring centers, etc. based on the quantum mechanically optimized theozyme.

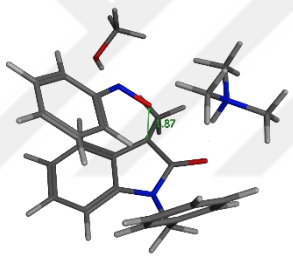
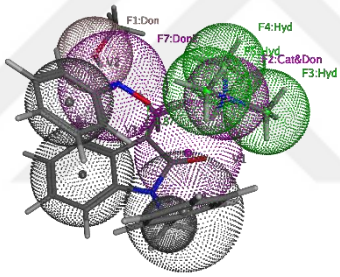
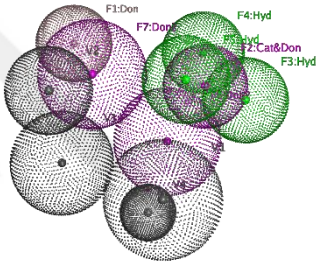
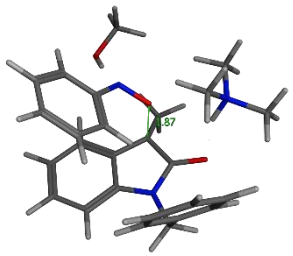
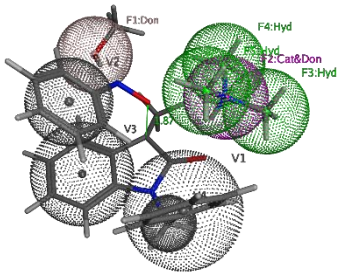
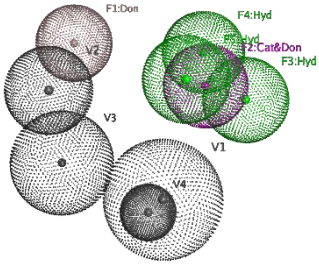
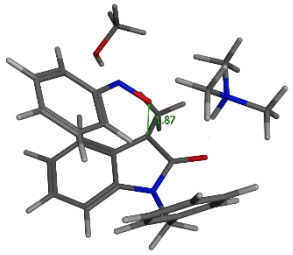
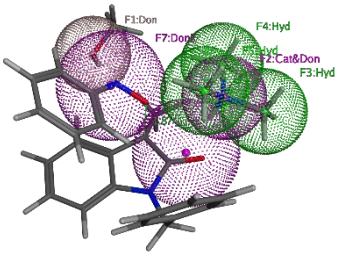
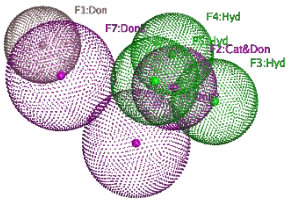
Table 4.1. Selected pharmacophore groups for CGM (3D layout is shown in Table 4.2)

CGM	**Hydroxyl Group**			**Protonated Amine Group**			**Excluded Volume**		
		Radius	**Feature**		**Radius**	**Feature**		**Radius**	**Feature**
CGM-1	F1	1.0	Don	F2	1.0	Cat&Don	V1	2.0	Exc1
	F7	1.4	Don2 vector	F3	1.0	AtomQ	V2	1.5	Exc1
				F4	1.0	AtomQ	V3	2.0	Exc1
				F5	1.0	AtomQ	V4	1.0	Exc1
				F6	1.4	Don2 vector			
CGM-2	F1	1.0	Don	F2	1.0	Cat&Don	V1	2.0	Exc1
				F3	1.0	AtomQ	V2	1.5	Exc1
				F4	1.0	AtomQ	V3	2.0	Exc1
				F5	1.0	AtomQ	V4	1.0	Exc1
CGM-3	F1	1.0	Don	F2	1.0	Cat&Don			
	F7	1.4	Don2 vector	F3	1.0	AtomQ			
				F4	1.0	AtomQ			
				F5	1.0	AtomQ			
				F6	1.4	Don2 vector			

Different CGMs, consist of different features were generated using different interaction types (Figure 3.7) or variant radius determining the tolerance on the deviation from the exact position in the pharmacophore groups contained in the lowest energy theozymes for each binding mode. These differences in pharmacophore queries (Table 4.1) are shown in Table 4.2 as an example on the generated theozyme (**Model A1**). Additionally, excluded volume property was taken into account to avoid steric clash between substrates and catalyst in

CGM-1 and **CGM-2** which are shown in Table 4.1. **CGM-1** which has seven features and four excluded volume is the most restrictive pharmacophore query for this target reaction. **CGM-2** and **CGM-3** were constructed as the subsets of **CGM-1**. All the essential features, which are representations of trimethyl amine and methanol, must be present in all CGMs. Increasing the number of pharmacophore features involves implementing more stringent criteria. This can significantly reduce the probability of finding a suitable match during the virtual screening process. So, if hit numbers were not sufficient using **CGM-1**, subsets of **CGM-1** (**CGM-2** and **CGM-3**) were taken into account.

Table 4.2. Representation of 3D layout of CGMs for aminooxylation modelTS

	ModelTS	**Selected Pharmacopore Groups**	**CGM**
CGM-1			
CGM-2			
CGM-3			

4.2.3. Generating Conformer Library of Organocatalyst

A conformational search was performed for the explored dimeric catalyst using MOE [64] software with MMFF94x. A LowModeMD search was utilized based on its capability to explore conformations that are close to relevant and realistic low-energy states, as well as its efficiency in identifying most of the local minima of complex multicomponent systems [100]. An iteration limit of 500000 with a rejection limit of 500000 was used to obtain a sufficient sample of conformational minima because of the highly complex multifunctional organocatalyst with a large number of conformational degree of freedom. When the RMSD values of the two conformations are less than 0.25, these were considered duplicates. 11103 conformers were generated for the dimeric catalysts. All conformations were ranked according to the MM force field calculations, and the lowest 40 of them were optimized with QM to calculate the lowest energy conformer of each studied catalyst.

4.2.4. Matching the Theozymes

All CGMs based on the low energy theozymes were screened against the catalyst conformer library containing 11103 conformers (Figure 4.7). The resulting hit conformers were primarily sorted by their matching scores (RMSD) and their deviations from their optimum catalytic geometries. As an example, **CGM-1** (Table 4.1 & 4.2) based on the **Model A1** was screened against the catalyst conformer library and 32 hit conformers were determined. After these 32 hit conformers were sorted with respect to ascending matching scores (RMSD), and they were examined based on predefined criteria, which are,

- Results showing an overlap between substrates (oxindole and nitrosobenzene) and matched catalysts were ignored.
- The cutoff criterion was set to 0.8 RMSD, and matches with RMSD values higher than 0.8 were discarded.

A total of seven hits were determined as possible active conformers for further evaluation using QM calculations. Screening of **CGM-1** (Table 4.1 & 4.2) based on the enantiomeric model of **Model A1** gave only three hit conformers.

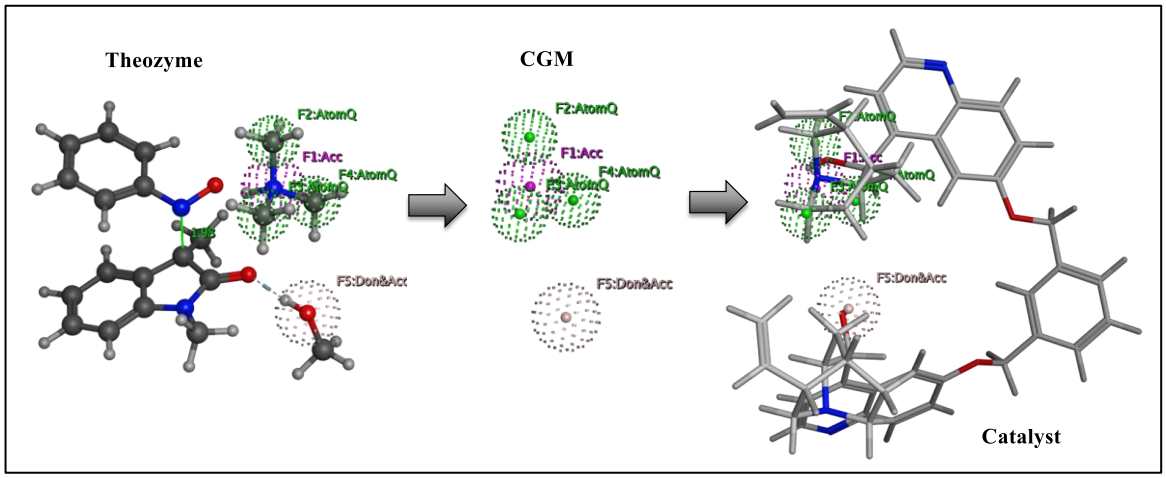


Figure 4.7. Construction of CGM using modelTS and matched conformer of dimeric catalyst

4.2.5. Quantum Mechanical Calculations

On the quantum mechanically optimized theozyme, replacing the model catalytic groups (methanol and trimethyl amine) with the matching conformer of the catalyst gives the transition structure in the presence of the chiral Barbas' catalyst (Figure 4.8). The product distribution and enantioselectivity were explored by applying the same protocol on the model transition state constructed for each binding preference which is given in Figure 4.5. Finally, the rates and selectivity of the leads were evaluated using QM calculations.

Detailed explanations about QM calculations were mentioned previously in Chapter 3.2.5. The same methodology was applied in this chapter. The only difference is the use of THF as a solvent instead of toluene to represent the experimental conditions [90–92].

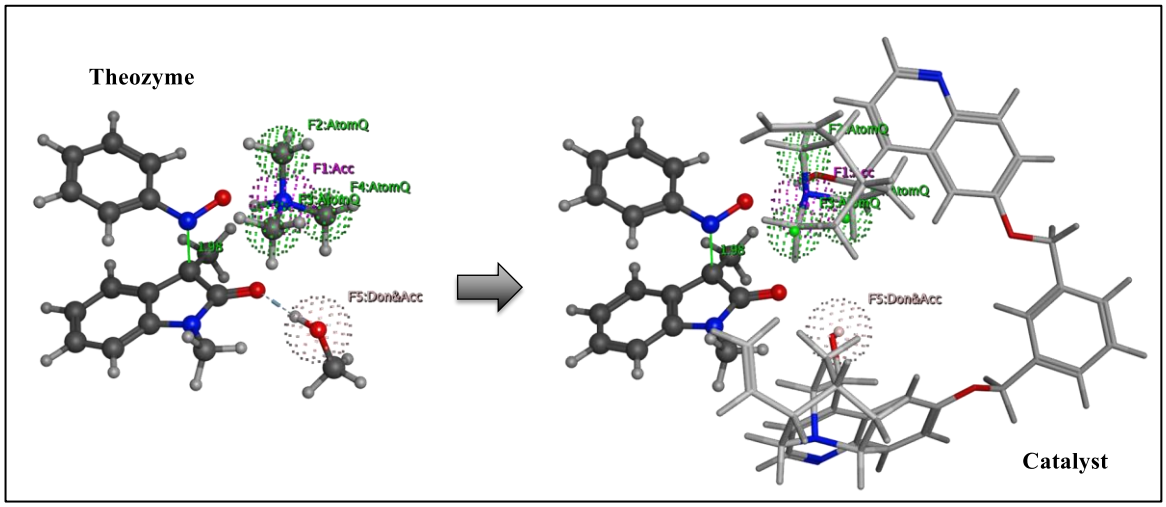


Figure 4.8. Matching the catalytic active site model for hydroxyamination

4.3. RESULTS AND DISCUSSION

4.3.1. Theozymes Based on Different Binding Modes

We generated theozymes (Figure 4.5) using different binding modes to understand their effect on the activation barrier. In this chapter, since the Barbas' dimeric catalyst was studied, only the alcohol and tertiary amine functional groups present in the structure of this catalyst were used to construct the theozymes. Firstly, considering the proposed catalytic cycles (Figure 4.3) theozymes leading to aminooxylation were generated using described protocol (Chapter 4.2.1). Quantum mechanically optimized theozymes and their activation barriers relative to the starting materials are given in Figure 4.9.

Model A2, which is the theozyme for **catalytic cycle II**, has a lower activation barrier than **Model A1**, which is the theozyme for **catalytic cycle I**. $\mathbf{TS_{Model\ A1}}$ describing a stepwise pathway proposed by Barbas, is calculated to be 1.2 kcal/mol higher in energy than the lowest theozyme. These calculations suggest that aminooxylation reactions proceed through a concerted addition-proton transfer pathway. This result agrees with the reported theozyme calculations by Çamlısoy [98].

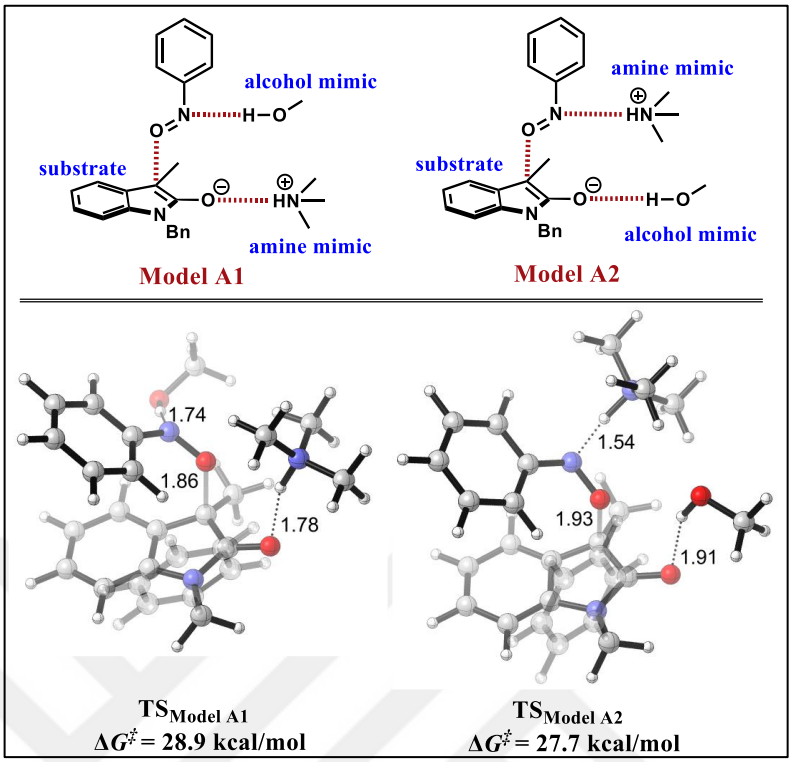


Figure 4.9. Theozymes and their relative Gibbs free energies relative to the starting materials for aminooxylation reaction

The effect of the protecting group (benzyl) of oxindole on the activation barrier was also investigated. Figure 4.10 shows the optimized two different arrangements of the benzyl group (highlighted in green) and their activation barries relative to the starting materials. **Model A1(a)** is 5.8 kcal/mol lower in energy than **Model A1(b)**. **Model A1(a)** protfits form a stabilizing edge to face π-π interaction between benzene ring of oxindole with nitrosobenzene [101, 102]. Although **Model A1(b)** is higher energy, both of these theozymes were included in screening against catalyst library to explore all the binding modes.

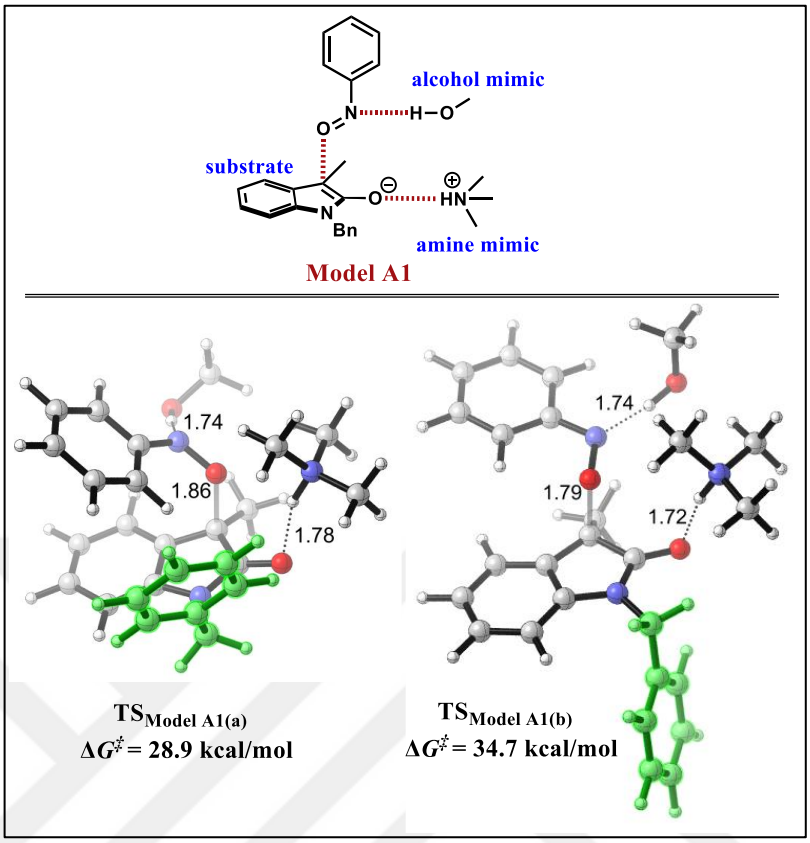


Figure 4.10. Different arrangement of benzyl group of substrate for aminooxylation model and their Gibbs free energies relative to starting materials

4.3.2. Matching and Evaluation of the Full Catalysts

On the quantum mechanically optimized theozymes, replacing the model catalytic groups with the matching conformer of the catalyst gives the TS in the presence of the full catalyst. Finally, the rates and selectivity of the hit conformers of the catalyst were evaluated using QM calculations. Only the lowest energy enantiomeric transition structures in the presence of matching catalyst are discussed in the following sections.

4.3.2.1. Stepwise Pathway

Figure 4.11 represents the enantiomeric transition state structures for the rate-determining addition step of aminooxylation for **catalytic cycle I** in the presence of Barbas' dimeric catalyst with different binding modes and their activation barriers with respect to reactants. The important interactions on these structures that need to be compared are highlighted in green.

***(R)*-TS$_{A1.1}$** is the lowest energy transition state showing only two polar interactions between the catalyst and the substrates. Protonated amine form an ion-pair with the enolate and alcohol of the same quinidine unit orients nitrosobenzene.

Even though matching resulted in conformers allowing an additional polar contact from the C9-alcohol of the second quinidine unit (Model A1.2 and Model A1.3), the corresponding TS structures are found to be higher in energy.

***(R)*-TS$_{A1.2}$** is 2.3 kcal/mol higher in energy than ***(R)*-TS$_{A1.1}$** with a forming C-O bond length of 1.92 Å compared to ***(R)*-TS$_{A1.1}$** with a forming C-O bond length of 1.80 Å. ***(R)*-TS$_{A1.1}$** shows a strong ion pair interaction ($d_{O..HN} = 1.68$Å). A weakening in this ion pair interaction has in ***(R)*-TS$_{A1.2}$** ($d_{O..HN} = 1.86$Å) This strong ion pair interaction has been disrupted in ***(R)*-TS$_{A1.2}$** resulted in a 2.3 kcal/mol energetic cost.

***(R)*-TS$_{A1.3}$** has the highest activation barrier (20.8 kcal/mol) among all the binding modes of aminooxylation TS through stepwise path. This can be explained by the negative charge on enolate stabilized by an additional hydroxyl group of dimeric catalyst. So, a decrease in electron density leads to the deactivation of the nucleophile that significantly increases the activation barrier.

The calculations of the rate determining addition step of aminooxylation for **catalytic cycle I** (Figure 4.11) showed that the *Re* face attack is more stable than the *Si* face for binding mode 1 and 2 (**A1.1 & A1.2**). Contrary, *Si* face attack is more favorable for binding mode 3 (**A1.3**).

Different arrangements of the protecting benzyl group of substrate for ***(R)*-TS$_{A1.1}$** is also explored and results are given in Figure 4.12. The upward orientation of the protecting benzyl group of oxindole did not provide any additional π-π interactions. Although both of them have the same polar interactions between substrates and catalyst, ***(R)*-TS$_{A1.1(a)}$** is 10.1 kcal/mol lower in energy than ***(R)*-TS$_{A1.1(b)}$** in energy because of stronger H-bonds as can be depicted from shorter H-bond distances and close to linear H-bond angles ($d_{O..HN} = 1.68$Å, 174.3° and $d_{N..HO} = 1.59$Å, 177.1° versus $d_{O..HN} = 1.95$Å, 140.1° and $d_{N..HO} = 1.75$Å, 172.7°). On the side of the dimer that does not interact with the substrates, a stabilizing intramolecular interaction is also observed between the hydroxyl group and the amine, which is 2.01 Å.

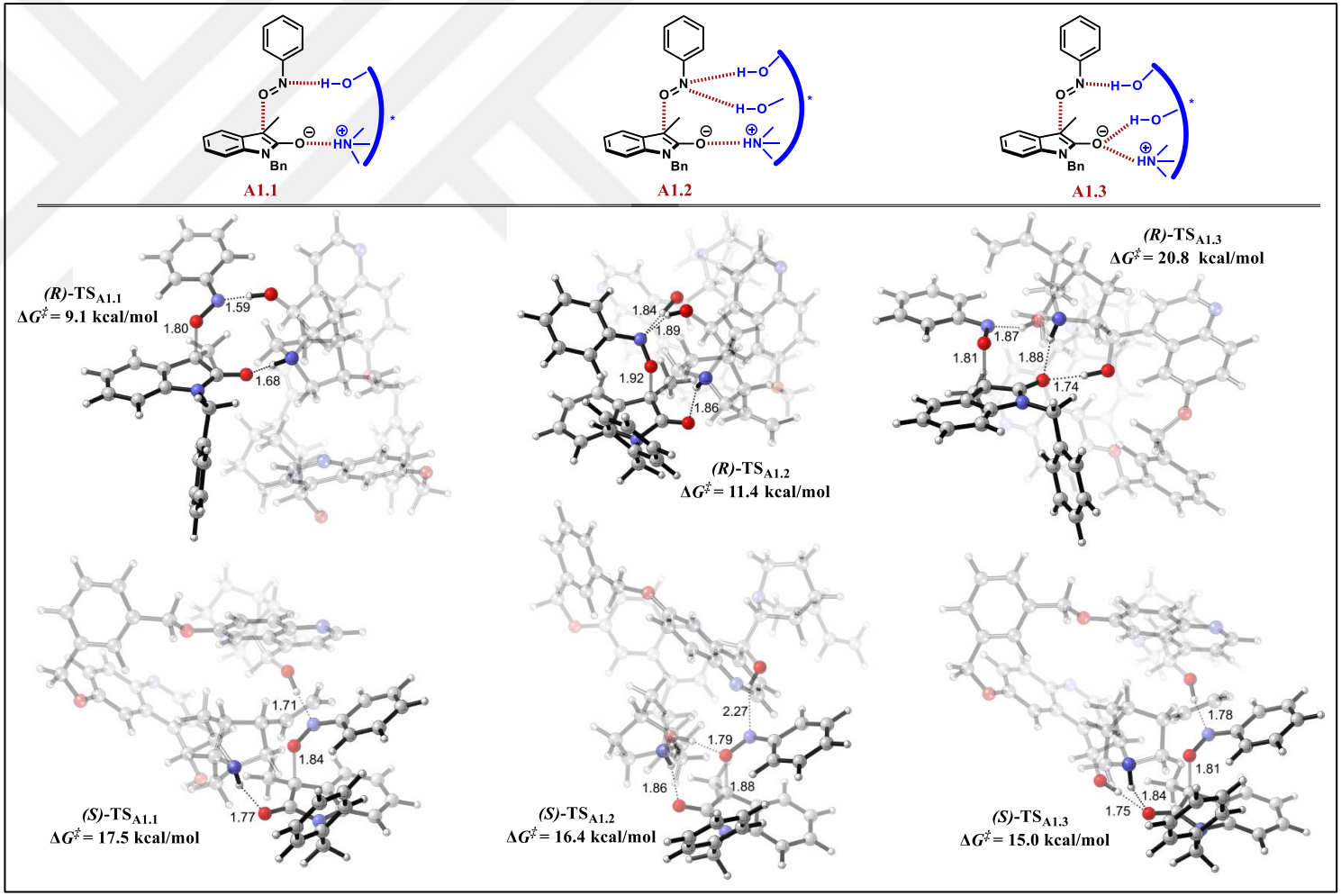


Figure 4.11. Transition states of aminooxygenation for the **catalytic cycle I** in the presence of dimeric catalyst with different binding modes and their Gibbs free energies relative to starting materials

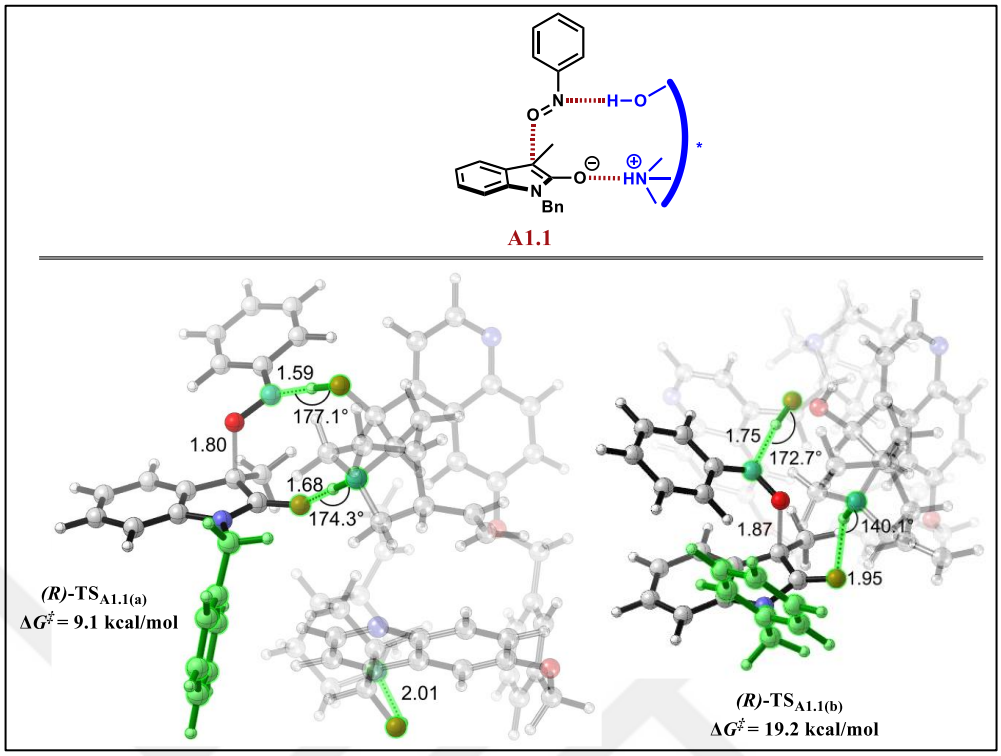


Figure 4.12. Different arrangement of benzyl group of substrate for ***(R)*-TS$_{A1.1}$** and their Gibbs free energies relative to starting materials

Figure 4.13 shows the different interactions of the hydroxyl group of catalyst with nitrosobenzene for ***(R)*-TS$_{A1.2}$**. Although both of the transition states have the same interaction according to Figure 4.6, the additional hydroxyl group of dimeric catalyst interacts with nitrosobenzene in two different arrangements. In **catalytic cycle I** (Figure 4.3 (a)), addition of the enolate to nitrosobenzene develops a negative charge on its nitrogen. In ***(R)*-TS$_{A1.2(a)}$** additional hydroxyl group interacts with the nitrogen of nitrosobenzene, while in ***(R)*-TS$_{A1.2(b)}$** it interacts the oxygen of nitrosobenzene which is involved in the bond formation. A combined effect of the destabilization of ***(R)*-TS$_{A1.2(b)}$** together with better stabilization of the negative charge on the nitrogen atom of nitrosobenzene by bidentate H-bonding in ***(R)*-TS$_{A1.2(a)}$** disfavors ***(R)*-TS$_{A1.2(b)}$** by 7.3.

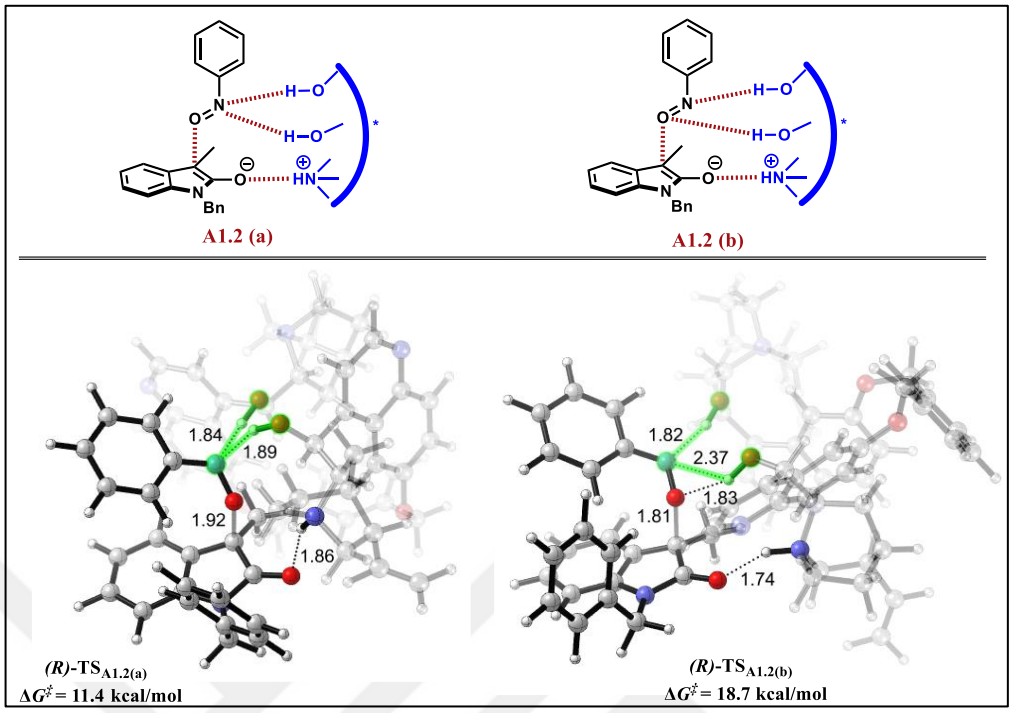


Figure 4.13. Transition state alternatives for ***(R)*-TS$_{A1.2}$** and their Gibbs free energies relative to starting materials

4.3.2.2. Concerted Pathway

Figure 4.14 represents the enantiomeric transition state structures for the rate-determining addition step of aminooxylation for **catalytic cycle II** in the presence of Barbas' dimeric catalyst with different binding modes and their activation barriers with respect to reactants. ***(R)*-TS$_{A2.2}$** is the lowest energy transition state.

The structure of ***(R)*-TS$_{A2.2}$** from two different views is shown in Figure 4.15. The important interactions are highlighted in green. In ***(R)*-TS$_{A2.2}$**, an additional hydroxyl group interacts with the nitrogen of nitrosobenzene, so the negative charge on the nitrogen atom of nitrosobenzene can be better stabilized by the presence of two strong H-bonds, as explained before with the example of ***(R)*-TS$_{A1.2(a)}$**. In addition to H-bonds, there are three different π-π interactions. First one (pi-pi int 1) is an intramolecular interaction between quinoline and the alkene group of catalysts. The second (pi-pi int 2) is an intermolecular interaction between the alkene group of the catalyst and the benzene ring of oxindole. The last one (pi-pi int 3) is also an intermolecular interaction between the quinolone part of the catalyst and the benzene ring of oxindole.

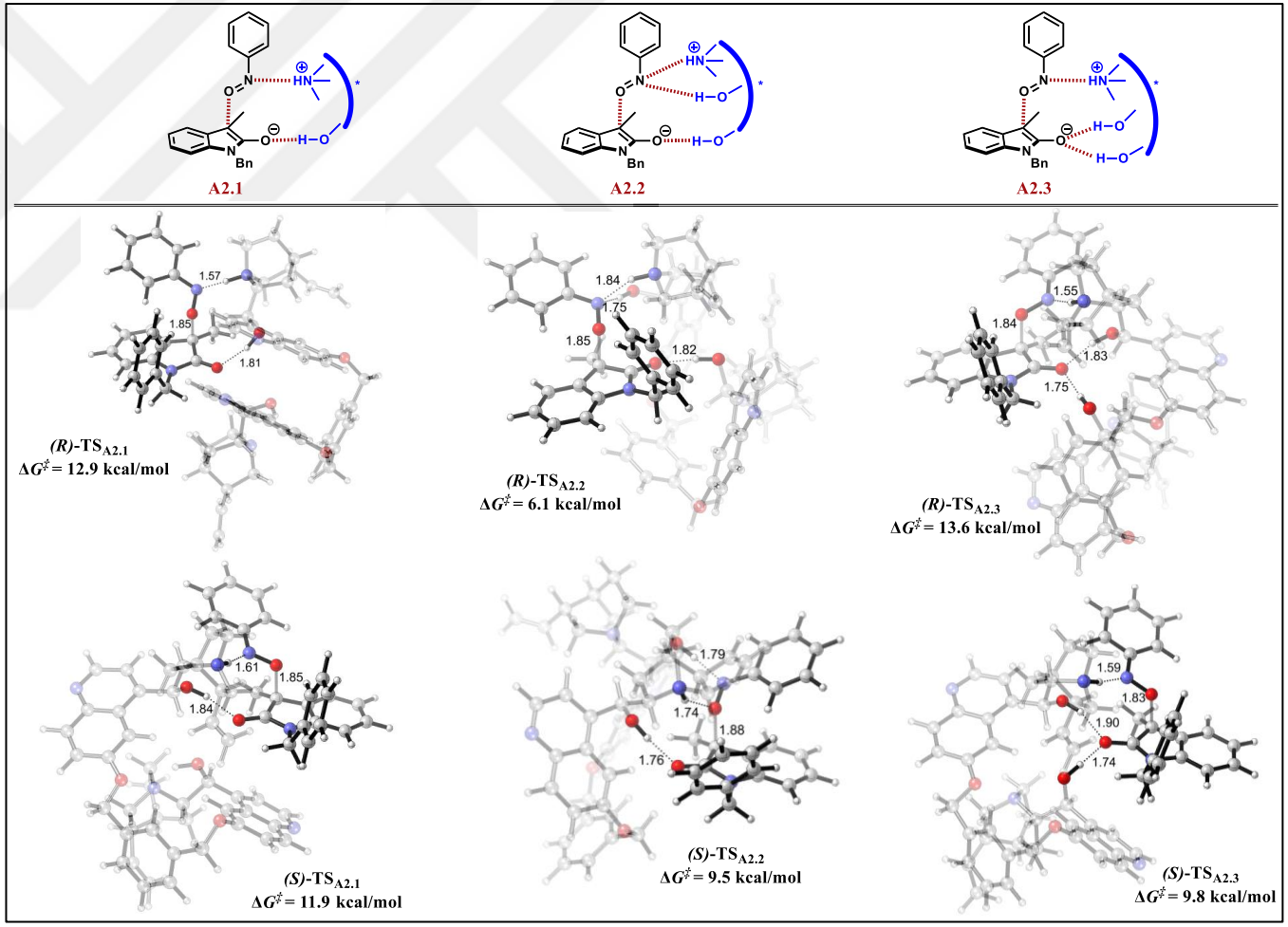


Figure 4.14. Transition states of aminooxygenation for **catalytic cycle II** in the presence of a dimeric catalyst with different binding modes and their Gibbs free energies relative to starting materials

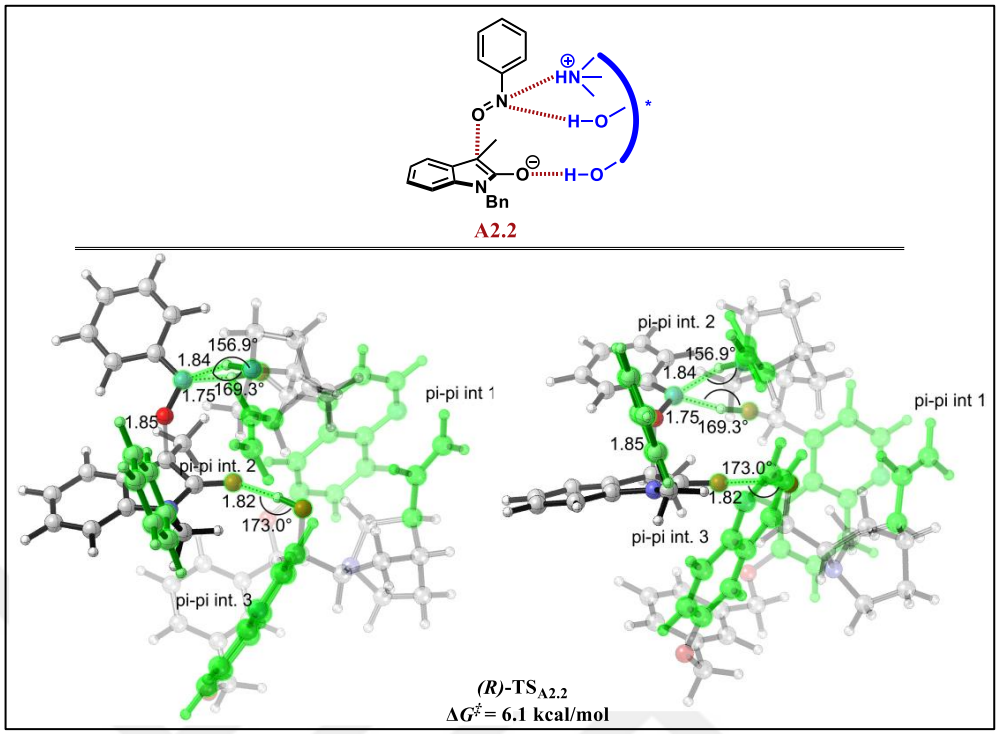


Figure 4.15. Lowest energy transition state geometry, ***(R)*-TS$_{A2.2}$** from two different views and its Gibbs free energies relative to starting materials

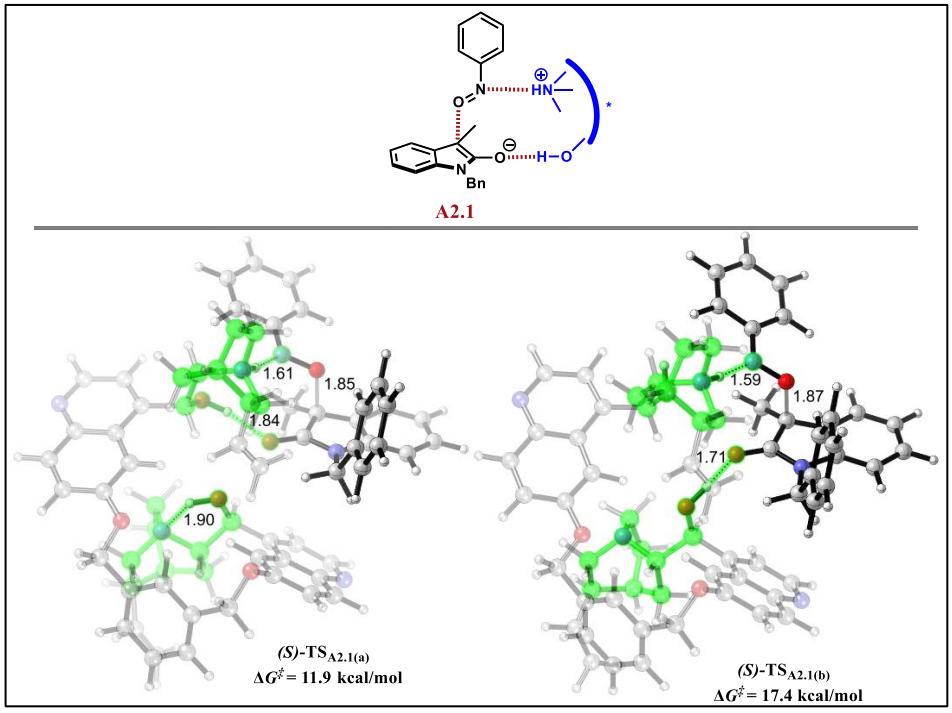


Figure 4.16. Transition state alternatives for ***(S)*-TS$_{A2.1}$** and their Gibbs free energies relative to starting materials

***(S)*-TS$_{A2.2}$** is has the same binding mode that leads to the *S* enantiomer. In ***(S)*-TS$_{A2.2}$** 3.4 kcal/mol higher in energy than ***(R)*-TS$_{A2.2}$**, protonated amine group interacts with the oxygen of nitrosobenzene, which is involved in the bond formation, instead of the nitrogen of nitrosobenzene. This interaction supposably increased its activation barrier.

***(R)*-TS$_{A2.1}$** is 6.8 kcal/mol higher in energy ***(R)*-TS$_{A2.2}$** (Figure 4.14) leading preferably to *S* enantiomer. Figure 4.16 shows the different arrangements of the hydroxyl group of catalyst for ***(S)*-TS$_{A2.1}$**. ***(S)*-TS$_{A2.1(a)}$** is 5.5 kcal/mol lower in energy than ***(S)*-TS$_{A2.1(b)}$**. Both TSs have the same binding mode involving a total of two polar interactions. The major difference between the two structures is the alcohol group in the catalyst interacting with the enolate. In ***(S)*-TS$_{A2.1(b)}$**, protonated amine and alcohol units belong to different quinidine units, whereas in ***(S)*-TS$_{A2.1(a)}$** interacting alcohol and amine groups are on the same quinidine moiety, and an intramolecular H-bond is observed between the hydroxyl group and the amine in the second quinidine moiety ($d_{N..HO} = 1.90$Å highlighted in green). This intramolecular H-bonding interaction presumably stabilizes the catalyst decreasing its distortion energy.

***(R)*-TS$_{A2.3}$** has the highest activation barrier (13.6 kcal/mol) among all the binding modes of aminooxylation TS through concerted path, as aforementioned ***(R)*-TS$_{A1.3}$** in Figure 4.11. This can be also explained by the negative charge on enolate stabilized by an additional hydroxyl group of dimeric catalyst. So, a decrease in electron density leads to the deactivation of the nucleophile that significantly increases the activation barrier.

The calculations of the rate determining addition step of aminooxylation for **catalytic cycle II** (Figure 4.14) showed that the *Si* face attack is more stable than the *Re* face for binding mode 1 and 3 (**A2.1 & A2.3**). Contrary, *Re* face attack is more favorable for binding mode 2 (**A2.2**).

***(R)*-TS$_{A2.2}$** (in Figure 4.15) is the lowest energy transition state among all aminooxylation TSs. Its energy is 3 kcal/mol higher than ***(R)*-TS$_{A1.1}$** (in Figure 4.12) describing a stepwise pathway proposed by Barbas. These calculations suggest that aminooxylation reaction proceeds via a concerted addition-proton transfer pathway.

4.3.2.3. Evaluation of N- Selectivity

To explain the conflicting experimental results based on product distribution in the literature, considering the proposed catalytic cycles (Figure 4.3) theozymes leading to

hydroxyamination were generated. Quantum mechanically optimized theozymes and their activation barriers relative to the starting materials are given in Figure 4.17.

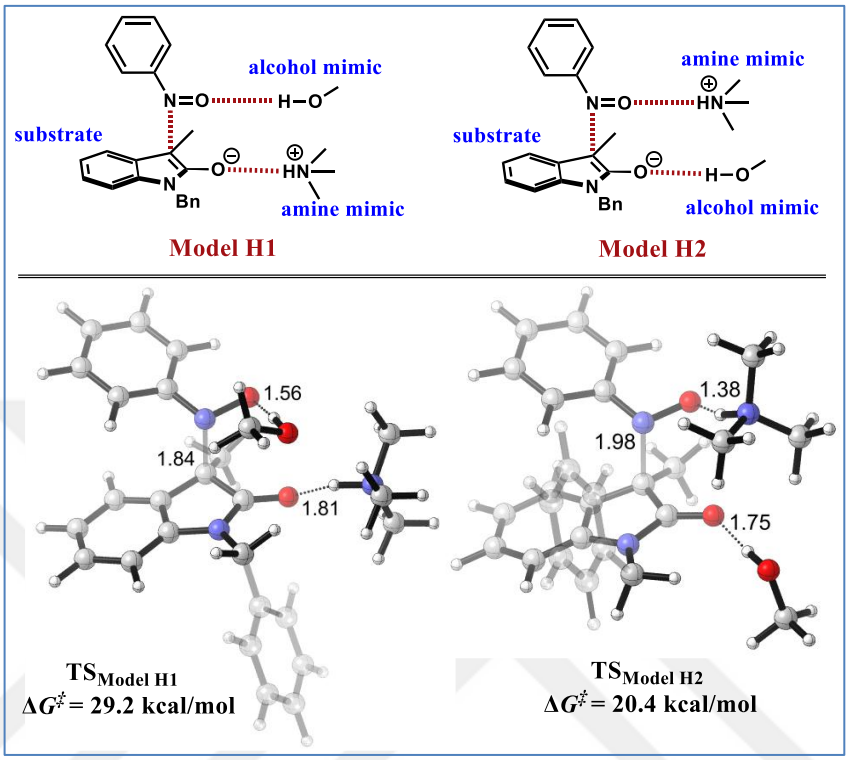


Figure 4.17. Theozymes and their relative Gibbs free energies relative to the starting materials for hydroxyamination reaction

Model A2 and **Model H2**, which are theozymes for **catalytic cycle II**, have lower activation barrier than **Model A1** and **Model H1,** which are theozymes for **catalytic cycle I**. These calculations suggest that both aminooxylation and hydroxyamination reactions proceed through a concerted addition-proton transfer pathway. This result agrees with the reported theozyme calculations by Çamlısoy [98]. **TS**$_{\textbf{Model A1}}$ and **TS**$_{\textbf{Model H1}}$, describing a stepwise pathway proposed by Barbas, is calculated to be 28.9 kcal/mol and 29.2 kcal/mol respectively.

Figure 4.18 compares the polar interactions in the optimized theozymes of aminooxylation and hydroxyamination reactions for **catalytic cycle II**. Although there is an extra face to edge π-π interaction between nitrosobenzene and the benzyl group of oxindole in **TS**$_{\textbf{Model A2}}$, **TS**$_{\textbf{Model H2}}$ is significantly stabilized by stronger H-bonds as can be depicted from shorter H-

bond distances and close to linear H-bond angles ($d_{N..HN}$ = 1.54Å, 168.3° and $d_{O..HO}$ = 1.91Å, 152.6° versus $d_{O..HN}$ = 1.38Å, 163.9° and $d_{O..HO}$ = 1.75Å, 176.3°).

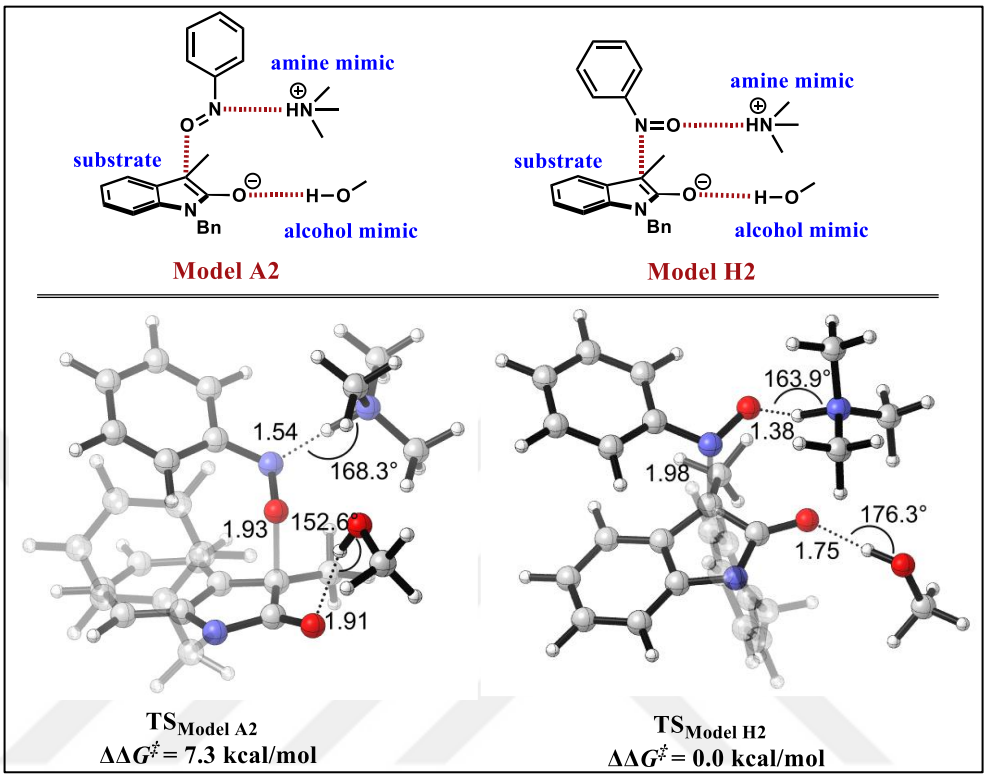


Figure 4.18. Theozyme comparison **Model A2** vs **Model H2** and their relative Gibbs free energies relative to the lowest energy **Model H2**

Model calculations suggest that the hydroxyamination reaction is favoured energetically for both catalytic cycles. While this energetic difference is 5.5 kcal/mol for **catalytic cycle I**, it is 7.3 for **cycle II**. These theoretically obtained results are in line with the experimentally observed hydroxyamination product in the presence of certain catalyst [88, 90, 92].

In order to evaluate the unexpected *O*-selectivity of the Barbas' quinidine dimer catalyst, we investigated the active catalyst conformers based on **Model H2.2**, especially focusing on the hydroxyamination analogue of ***(R)*-TS$_{A2.2}$,** which is the lowest energy binding mode and TS for the aminooxylation reaction.

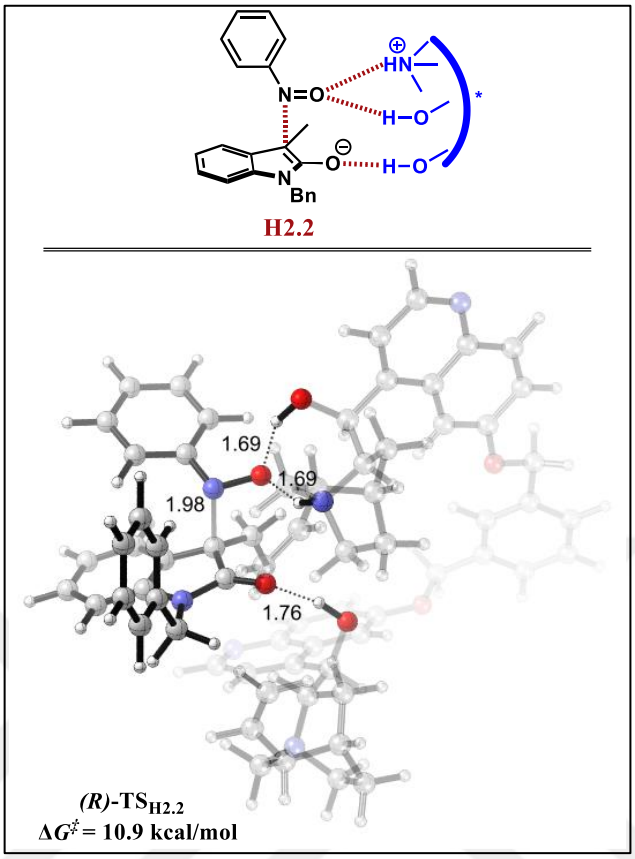


Figure 4.19. ***(R)*-TS$_{H2.2}$** and its Gibbs free energies relative to starting materials

Although calculations with the model catalyst have suggested that the hydroxyamination reaction is favored over the aminooxylation reaction, ***(R)*-TS$_{H2.2}$** (Figure 4.19) is 4.8 kcal/mol higher in energy than ***(R)*-TS$_{A2.2}$**. To better understand the reason of energy difference between ***(R)*-TS$_{A2.2}$** and ***(R)*-TS$_{H2.2}$**, the distortion/interaction analysis (Figure 4.20) was performed for the *Re* face attack TSs as they are more favorable than the *Si* face (Figure 4.14). The distortion/interaction analysis [81] of the transition states can be used to provide insights into the source of enantioselectivity and the factors that influence the activation energy barrier in the step responsible for determining regioselectivity and enantioselectivity.

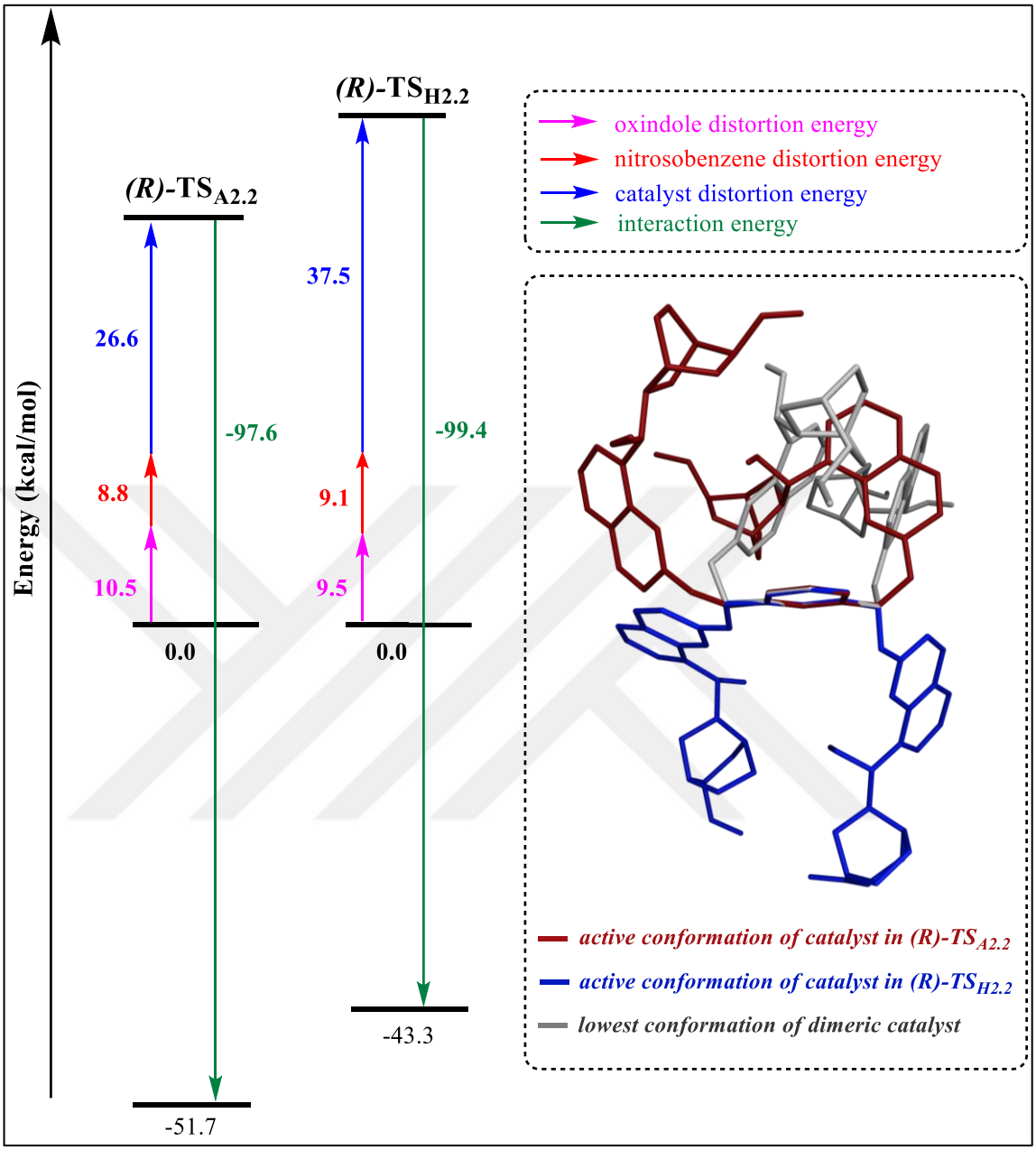


Figure 4.20. Distortion/Interaction Analysis of ***(R)*-TS$_{A2.2}$** & ***(R)*-TS$_{H2.2}$**

For both transition state as the interaction energies are close to each other, the distortion energy determines the activity of catalyst. Lower catalyst distortion favorably contribute to the stabilization of aminooxylation with respect to hydroxyamination reaction. When the distortion energy of substrates was compared, hydroxyamination transition state has 0.7 kcal/mol lower distortion energy of substrates (oxindole and nitrosobenzene). The lower distortion of substrates in the ***(R)*-TS$_{H2.2}$** can be also explained by bond lengths in the located TSs. The located transition structure ***(R)*-TS$_{H2.2}$** which has a forming C-N bond 1.98Å and

its substrates are less distorted than ***(R)*-TS$_{A2.2}$** which has a forming C-O bond of 1.84Å. It is clear that ***(R)*-TS$_{H2.2}$** is more favorable according to the distortion of substrates. However, the main difference is the distortion energies of catalysts. Dimeric catalyst pays 10.9 kcal/mol higher energetic penalty in order to attain its active conformation in ***(R)*-TS$_{H2.2}$** relative to its conformer in ***(R)*-TS$_{A2.2}$**, which results in a much lower activation barrier for aminooxylation reaction. These results suggest that the energetic cost of distorting the dimeric catalyst to the active conformation may play a significant role in the destabilization of ***(R)*-TS$_{H2.2}$** with respect to ***(R)*-TS$_{A2.2}$**. This leads to a decrease in the reactivity of the dimeric catalyst in hydroxyamination reaction.

Overall, calculations favor a concerted addition-proton transfer pathway leading to both hydroxyamination and aminooxylation products. The free energy profile for the lowest energy pathways toward aminooxylation and hydroxyamination reaction is given in Figure 4.21. In aminooxylation pathway, the catalyst in its active conformation binds nitrosobenzene through the protonated amine and the hydroxyl group on the same quinidine moiety and orients the oxindole to a near attack conformation by forming a H-bond with the alcohol functionality on the other quinidine moiety. This complex which is 2.6 kcal/mol higher in energy than the separated reactants, leads to ***(R)*-TS$_{A2.2}$**, which is 3.5 kcal/mol uphill. The product is 4.0 kcal/mol more stable than the separated reactants leading to a free energy span of 10.1 kcal/mol for the aminooxylation catalytic cycle. The binding mode of the catalyst in the hydroxyamination pathway is essentially the same, yet the complex is much higher in energy (ΔG=9.1 kcal/mol), presumably due to catalyst distortion. The corresponding TS is only 1.8 kcal/mol higher in energy (indicative of the facile formation of the *N*-adduct how the complex) leading to *N*-adduct, 2.5 kcal/mol more stable than the reactants. The calculated free energy span for the hydroxyamination reaction is 13.4 kcal/mol over the catalytic cycle. The calculations therefore suggest that the aminooxylation pathway is 3.3 kcal/mol favoured over the hydroxyamination pathway as a result of the catalyst distortion.

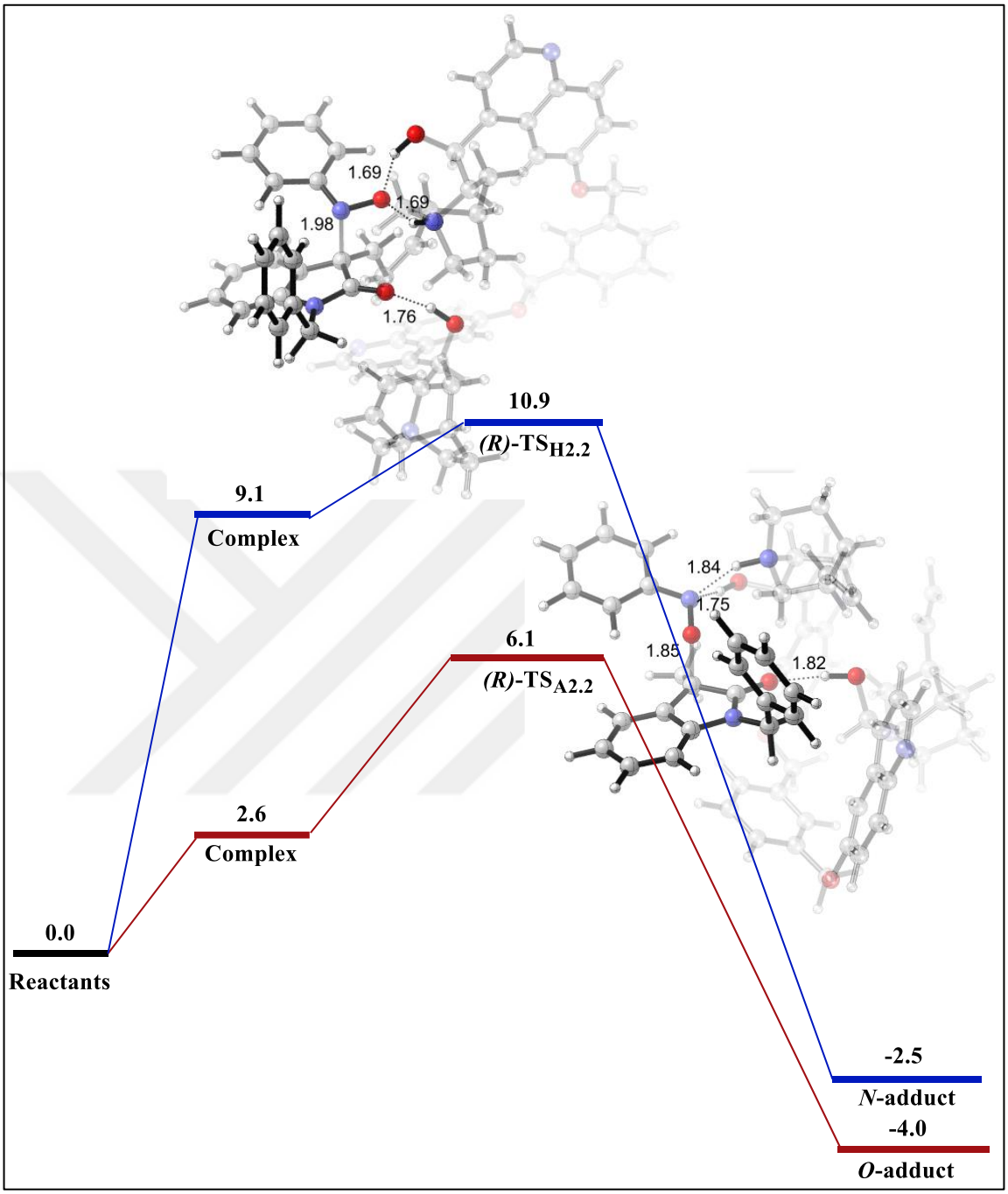


Figure 4.21. Free energy profile ($\Delta G^{\neq}$) in kcal/mol for the reaction of *N*-protected 3-methyl-2-oxindole with nitrosobenzene

4.4. CONCLUSION AND FUTURE WORK

CIPOC was applied to explore the ambiguous active conformations and mode of action of complex multifunctional dimeric quinidine organocatalysts with a high degree of conformational freedom for the aminooxylation reaction of *N*-benzyl-protected 3-substituted-2-oxindole with nitrosobenzene.

Calculations with the model catalyst have suggested that the hydroxyamination reaction is favored over the aminooxylation reaction. The inclusion of a dimeric multifunctional catalyst has additionally exposed the importance of the 3D arrangement of functional groups of dimeric catalysts and shown the significance of catalyst distortion in determining the product distribution. Calculations with the dimeric catalyst showed that aminooxylation reaction is more highly favored than hydroxyamination reaction. Our results revealed that C6' quinidine dimer binds nitrosobenzene through its protonated amine, while both hydroxyl groups in the catalyst structure interact and orient substrates. The reaction proceeds via a concerted addition-proton transfer pathway to give the *O*-adduct as major product. Although these results are in line with Barbas' experimental observations, calculations revealed that reaction proceeds via **catalytic cycle II** proposed by Çamlısoy.

This study not only demonstrates the applicability of CIPOC to the investigation of the catalytic activity and selectivity of highly complex multifunctional organocatalysts with a large number of conformational degree of freedom but also shed light on the ambiguous active conformations and mode of action of the dimeric catalyst.

Quantum Theory of Atoms in Molecules (QTAIM) and Independent Gradient Model (IGM) analyses can be performed to facilitate the visualization and better understanding of intermolecular and intramolecular interactions. QTAIM [103, 104] can be employed to analyze the intramolecular and intermolecular interactions formed in TSs. This approach examines the topological properties at hydrogen bonds, van der Waals interactions, and π-π interactions [105–110]. Additionally, IGM [111] analysis is a method that can be used to visualize intermolecular interactions in the structures. Energy decomposition analysis (EDA) [112] also can be utilized for dissecting the factors influencing selectivity.

REFERENCES

1. Cheong PHY, Legault CY, Um JM, et al. Quantum mechanical investigations of organocatalysis: Mechanisms, reactivities, and selectivities. *Chem Rev* 2011; 111: 5042–5137.

2. Çelebi Nihan. *TUBITAK 114Z791 - Determination of Active Organocatalysts Using Computational Methods*. İstanbul, 2013.

3. Richter F, Blomberg R, Khare SD, et al. Computational design of catalytic dyads and oxyanion holes for ester hydrolysis. *J Am Chem Soc* 2012; 134: 16197–16206.

4. Savile CK, Janey JM, Mundorff EC, et al. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science (1979)* 2010; 329: 305–309.

5. Jiang L, Althoff EA, Clemente FR, et al. De novo computational design of retro-aldol enzymes. *Science (1979)* 2008; 319: 1387–1391.

6. Kheirabadi M, Çelebi-Ölçüm N, Parker MFL, et al. Spiroligozymes for transesterifications: Design and relationship of structure to activity. *J Am Chem Soc* 2012; 134: 18345–18353.

7. Jensen F. *Introduction to Computational Chemistry*. Second Edition. Odense, Denmark: John Wiley & Sons, Ltd., 2007.

8. Foresman James B., Frisch A. *Exploring Chemistry with Electronic Structure Methods*. Third Edition. Wallingford, USA: Gaussian, Inc, 2015.

9. Basavaiah D, Rao AJ, Satyanarayana T. Recent advances in the Baylis - Hillman reaction and applications. *Chem Rev* 2003; 103: 811–891.

10. Lowe John P., Peterson Kırk A. *Quantum Chemistry*. Third Edition. MA,USA: Elsevier Academic Press, 2006.

11. Cramer CJ. *Essentials of Computational Chemistry; Theories and Models*. Second Edition. Minnesote, USA: John Wiley & Sons, 2004.

12. Leach Andrew R. *Molecular Modelling Principles and Applications*. Second Edition. Harlow, England: Pearson Prentice Hall, 2001.

13. Lee C, Yang eitao, Parr RG. Development of the Colic-Salvetti correlation-energy formula into a functional of the electron density. 37.

14. Becke AD. Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 1993; 98: 5648–5652.

15. Cramer CJ, Truhlar DG. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem Rev* 1999; 99: 2161–2200.

16. Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. *Chem Rev* 2005; 105: 2999–3093.

17. Norjmaa G, Ujaque G, Lledós A. Beyond Continuum Solvent Models in Computational Homogeneous Catalysis. *Top Catal* 2022; 65: 118–140.

18. McDonagh JL. *Computing The Aqueous Solubility of Organic Drug-like Molecules and Understanding Hyrdophobicity*. University of St Andrews, http://research-repository.st-andrews.ac.uk/ (2015).

19. Luque FJ, Curutchet C, Muñoz-Muriedas J, et al. Continuum solvation models: Dissecting the free energy of solvation. *Physical Chemistry Chemical Physics* 2003; 5: 3827–3836.

20. Herbert JM. Dielectric continuum methods for quantum chemistry. Epub ahead of print 14 March 2022. DOI: 10.1002/wcms.1519.

21. Rollinger JM, Stuppner H, Langer T. Virtual screening for the discovery of bioactive natural products. *Progress in Drug Research*; 65.

22. Vyas V, Jain A, Jain A, et al. Virtual screening: A fast tool for drug design. *Sci Pharm* 2008; 76: 333–360.

23. Qing X, Lee XY, De Raeymaeker J, et al. Pharmacophore modeling: Advances, Limitations, And current utility in drug discovery. *J Receptor Ligand Channel Res* 2014; 7: 81–92.

24. Tilahun Muhammed M, Aki-Yalcin E. Pharmacophore Modeling in Drug Discovery: Methodology and Current Status. 2021; 8: 749–762.

25. Govender T, Arvidsson PI, Maguire GEM, et al. Enantioselective Organocatalyzed Transformations of β-Ketoesters. *Chem Rev* 2016; 116: 9375–9437.

26. de Castro PP, Carpanez AG, Amarante GW. Azlactone Reaction Developments. *Chemistry - A European Journal* 2016; 22: 10294–10318.

27. Gaunt MJ, Johansson CCC, McNally A, et al. Enantioselective organocatalysis. *Drug Discov Today* 2007; 12: 8–27.

28. Odagi M, Nagasawa K. Recent Advances in Natural Products Synthesis Using Bifunctional Organocatalysts Bearing a Hydrogen-Bonding Donor Moiety. *Asian J Org Chem* 2019; 8: 1766–1774.

29. Giacalone F, Gruttadauria M, Agrigento P, et al. Low-loading asymmetric organocatalysis. *Chem Soc Rev* 2012; 41: 2406–2447.

30. Dondoni A, Massi A. Asymmetric organocatalysis: From infancy to adolescence. *Angewandte Chemie - International Edition* 2008; 47: 4638–4660.

31. List B, Yang JW. The organic approach to asymmetric catalysis. *Science (1979)* 2006; 313: 1584–1586.

32. MacMillan DWC. The advent and development of organocatalysis. *Nature* 2008; 455: 304–308.

33. Narayanaperumal S, Rivera DG, Silva RC, et al. Terpene-derived bifunctional thioureas in asymmetric Organocatalysis. *ChemCatChem* 2013; 5: 2756–2773.

34. Doyle AG, Jacobsen EN. Small-molecule H-bond donors in asymmetric catalysis. *Chem Rev* 2007; 107: 5713–5743.

35. Takemoto Y. Recognition and activation by ureas and thioureas: Stereoselective reactions using ureas and thioureas as hydrogen-bonding donors. *Org Biomol Chem* 2005; 3: 4299–4306.

36. Miyabe H, Takemoto Y. Discovery and application of asymmetric reaction by multi-functional thioureas. *Bull Chem Soc Jpn* 2008; 81: 785–795.

37. Alemán J, Parra A, Jiang H, et al. Squaramides: Bridging from molecular recognition to bifunctional organocatalysis. *Chemistry - A European Journal* 2011; 17: 6890–6899.

38. Palomo C, Oiarbide M, López R. Asymmetric organocatalysis by chiral Brønsted bases: Implications and applications. *Chem Soc Rev* 2009; 38: 632–653.

39. Sohtome Y, Shin B, Horitsugi N, et al. Entropy-controlled catalytic asymmetric 1,4-type Friedel-crafts reaction of phenols using conformationally flexible guanidine/bisthiourea organocatalyst. *Angewandte Chemie - International Edition* 2010; 49: 7299–7303.

40. Sohtome Y, Tanaka S, Takada K, et al. Solvent-dependent enantiodivergent mannich-type reaction: Utilizing a conformationally flexible guanidine/bisthiourea organocatalyst. *Angewandte Chemie - International Edition* 2010; 49: 9254–9257.

41. Chen X, Dong S, Qiao Z, et al. Guanidine organocatalyst for the asymmetric mannich-type reaction between α-isothiocyanato imide and sulfonyl imines. *Chemistry - A European Journal* 2011; 17: 2583–2586.

42. Yu Z, Liu X, Zhou L, et al. Bifunctional guanidine via an amino amide skeleton for asymmetric michael reactions of ß-ketoesters with nitroolefins: A concise synthesis of bicyclic β- Amino acids. *Angewandte Chemie - International Edition* 2009; 48: 5195–5198.

43. Ishikawa T, Araki Y, Kumamoto T, et al. Modified guanidines as chiral superbases: Application to asymmetric Michael reaction of glycine imine with acrylate or its related compounds. *Chemical Communications* 2001; 245–246.

44. Núñez MG, Farley AJM, Dixon DJ. Bifunctional iminophosphorane organocatalysts for enantioselective synthesis: Application to the ketimine nitro-Mannich reaction. *J Am Chem Soc* 2013; 135: 16348–16351.

45. Farley AJM, Sandford C, Dixon DJ. Bifunctional Iminophosphorane Catalyzed Enantioselective Sulfa-Michael Addition to Unactivated α-Substituted Acrylate Esters. *J Am Chem Soc* 2015; 137: 15992–15995.

46. Bandar JS, Lambert TH. Enantioselective Brønsted base catalysis with chiral cyclopropenimines. *J Am Chem Soc* 2012; 134: 5552–5555.

47. Bandar JS, Lambert TH. Cyclopropenimine-catalyzed enantioselective mannich reactions of tert -butyl glycinates with N -Boc-imines. *J Am Chem Soc* 2013; 135: 11799–11802.

48. Davis TA, Vilgelm AE, Richmond A, et al. Preparation of (-)-nutlin-3 using enantioselective organocatalysis at decagram scale. *Journal of Organic Chemistry* 2013; 78: 10605–10616.

49. Schwieter KE, Johnston JN. Enantioselective Addition of Bromonitromethane to Aliphatic N-Boc Aldimines Using a Homogeneous Bifunctional Chiral Organocatalyst. *ACS Catal* 2015; 5: 6559–6562.

50. Shen B, Makley DM, Johnston JN. Umpolung reactivity in amide and peptide synthesis. *Nature* 2010; 465: 1027–1032.

51. Dobish MC, Johnston JN. Achiral counterion control of enantioselectivity in a Brønsted acid-catalyzed lactonization. *J Am Chem Soc* 2012; 134: 6068–6071.

52. Davis TA, Johnston JN. Catalytic, enantioselective synthesis of stilbene cis-diamines: A concise preparation of (-)-Nutlin-3, a potent p53/MDM2 inhibitor. *Chem Sci* 2011; 2: 1076–1079.

53. Rabalakos C, Wulff WD. Enantioselective organocatalytic direct Michael addition of nitroalkanes to nitroalkenes promoted by a unique bifunctional DMAP-thiourea. *J Am Chem Soc* 2008; 130: 13524–13525.

54. Işık M, Tanyeli C. Cu-catalyzed selective mono-N-pyridylation: Direct access to 2-aminoDMAP/sulfonamides as bifunctional organocatalysts. *Journal of Organic Chemistry* 2013; 78: 1604–1611.

55. Işık M, Unver MY, Tanyeli C. Modularly evolved 2-aminodmap/squaramides as highly active bifunctional organocatalysts in Michael addition. *Journal of Organic Chemistry* 2015; 80: 828–835.

56. Okino T, Hoashi Y, Takemoto Y. Enantioselective Michael Reaction of Malonates to Nitroolefins Catalyzed by Bifunctional Organocatalysts. *J Am Chem Soc* 2003; 125: 12672–12673.

57. Okino T, Hoashi Y, Furukawa T, et al. Enantio- and diastereoselective michael reaction of 1,3-dicarbonyl compounds to nitroolefins catalyzed by a bifunctional thiourea. *J Am Chem Soc* 2005; 127: 119–125.

58. Hamza A, Schubert G, Soós T, et al. Theoretical studies on the bifunctionality of chiral thiourea-based organocatalysts: Competing routes to C-C bond formation. *J Am Chem Soc* 2006; 128: 13151–13160.

59. Zhu R, Zhang D, Wu J, et al. Theoretical study of the bifunctional-urea catalyzed Michael reaction of 1,3-dicarbonyl compounds and nitroolefins: reaction mechanism and enantioselectivity. *Tetrahedron Asymmetry* 2006; 17: 1611–1616.

60. Izzo JA, Myshchuk Y, Hirschi JS, et al. Transition state analysis of an enantioselective Michael addition by a bifunctional thiourea organocatalyst. *Org Biomol Chem* 2019; 17: 3934–3939.

61. Pliego JR. Theoretical free energy profile and benchmarking of functionals for amino-thiourea organocatalyzed nitro-Michael addition reaction. *Physical Chemistry Chemical Physics* 2020; 22: 11529–11536.

62. Tantillo DJ, Chen J, Houk KN. Theozymes and compuzymes: theoretical models for biological catalysis. *Curr Opin Chem Biol* 1998; 2: 743–750.

63. Kiss G, Çelebi-Ölçüm N, Moretti R, et al. Computational enzyme design. *Angewandte Chemie - International Edition* 2013; 52: 5700–5725.

64. Chemical Computing Group Inc. Molecular Operating Environment (MOE), https://www.chemcomp.com/ (2015, accessed 12 May 2023).

65. Chen IJ, Foloppe N. Conformational sampling of druglike molecules with MOE and catalyst: Implications for pharmacophore modeling and virtual screening. *J Chem Inf Model* 2008; 48: 1773–1791.

66. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision C.01. *Gaussian, Inc.*

67. DeChancie J, Clemente FR, Smith AJT, et al. How similar are enzyme active site geometries derived from quantum mechanical theozymes to crystal structures of enzyme-inhibitor complexes? Implications for enzyme design. *Protein Science* 2007; 16: 1851–1866.

68. Zhang X, DeChancie J, Gunaydin H, et al. Quantum mechanical design of enzyme active sites. *Journal of Organic Chemistry* 2008; 73: 889–899.

69. Stephens PJ, Devlin FJ, Chabalowski CF, et al. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *Physical Chemistry The Journal*; 0.

70. Wilk A N L, Nusair DM. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis1. *J Phys* 1980; 58: 1200.

71. Grimme S, Antony J, Ehrlich S, et al. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94

elements H-Pu. *Journal of Chemical Physics*; 132. Epub ahead of print 21 April 2010. DOI: 10.1063/1.3382344.

72. Gonzalez C, Bernhard Schlegel H. An improved algorithm for reaction path following. *J Chem Phys* 1989; 90: 2154–2161.

73. Gonzalez Carlos, Schegel H. Bernhard. Reaction Path Following in Mass-Weighted Internal Coordinates. *Journal of Physical Chemistry* 1990; 94: 5523–5527.

74. Marenich A V., Cramer CJ, Truhlar DG. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *Journal of Physical Chemistry B* 2009; 113: 6378–6396.

75. Ribeiro RF, Marenich A V., Cramer CJ, et al. Use of solution-phase vibrational frequencies in continuum models for the free energy of solvation. *Journal of Physical Chemistry B* 2011; 115: 14556–14562.

76. Balcells D, Maseras F. Computational approaches to asymmetric synthesis. *New Journal of Chemistry* 2007; 31: 333–343.

77. Gómez DE, Fabbrizzi L, Licchelli M, et al. Urea vs. thiourea in anion recognition. *Org Biomol Chem* 2005; 3: 1495–1500.

78. Bordwell FG. Equilibrium Acidities in Dimethyl Sulfoxide Solution. *Acc Chem Res* 1988; 21: 456–463.

79. Vachal P, Jacobsen EN. Chiral Auxiliaries and Ligands in Asymmetric Synthesis. *Angew Chem, Int Ed Engl* 1992; 31: 867–870.

80. Maher DJ, Connon SJ. Acceleration of the DABCO-promoted Baylis-Hillman reaction using a recoverable H-bonding organocatalyst. *Tetrahedron Lett* 2004; 45: 1301–1305.

81. Bickelhaupt FM, Houk KN. Analyzing Reaction Rates with the Distortion / Interaction - Activation Strain Model. *Angewandte Chemie* 2017; 129: 10204–10221.

82. Yoon TP, Jacobsen EN. Privileged chiral catalysts. *Science (1979)* 2003; 299: 1691–1693.

83. Boratyński PJ, Zielińska-Błajet M, Skarżewski J. Cinchona Alkaloids—Derivatives and Applications. In: *Alkaloids: Chemistry and Biology*. Academic Press Inc., 2019, pp. 29–145.

84. Hughes DL. Highlights of the Recent Patent Literature: Focus on Asymmetric Organocatalysis. *Org Process Res Dev* 2022; 26: 2224–2239.

85. Dalko PI, Moisan L. In the golden age of organocatalysis. *Angewandte Chemie - International Edition* 2004; 43: 5138–5175.

86. Singh G, Yeboah E. Recent applications of Cinchona alkaloid-based catalysts in asymmetric addition reactions. *Reports in Organic Chemistry* 2016; Volume 6: 47–75.

87. Marcelli T. Organocatalysis: Cinchona catalysts. *Wiley Interdiscip Rev Comput Mol Sci* 2011; 1: 142–152.

88. Jia LN, Huang J, Peng L, et al. Asymmetric hydroxyamination of oxindoles catalyzed by chiral bifunctional tertiary amine thiourea: Construction of 3-amino-2-oxindoles with quaternary stereocenters. *Org Biomol Chem* 2012; 10: 236–239.

89. Yang Z, Wang Z, Bai S, et al. Highly enantioselective synthesis of 3-amino-2-oxindole derivatives: Catalytic asymmetric α-amlnation of 3-substituted 2-oxindoles with a chiral scandium complex. *Chemistry - A European Journal* 2010; 16: 6632–6637.

90. Zhang T, Cheng L, Liu L, et al. Asymmetric organocatalytic N-nitroso-aldol reaction of oxindoles. *Tetrahedron Asymmetry* 2010; 21: 2800–2806.

91. Bui T, Candeias NR, Barbas CF. Dimeric quinidine-catalyzed enantioselective aminooxygenation of oxindoles: An organocatalytic approach to 3-hydroxyoxindole derivatives. *J Am Chem Soc* 2010; 132: 5574–5575.

92. Companyó X, Valero G, Pineda O, et al. Enantioselective organocatalytic oxyamination of unprotected 3-substituted oxindoles. *Org Biomol Chem* 2012; 10: 431–439.

93. Badillo JJ. Enantioselective synthesis of substituted oxindoles and spirooxindoles with applications in drug discovery. *Curr Opin Drug Discov Devel* 2010; 13: 758–776.

94. Zhou F, Liu YL, Zhou J. Catalytic asymmetric synthesis of oxindoles bearing a tetrasubstituted stereocenter at the C-3 position. *Adv Synth Catal* 2010; 352: 1381–1407.

95. Ishikura M, Abe T, Choshi T, et al. Simple indole alkaloids and those with a nonrearranged monoterpenoid unit. *Nat Prod Rep* 2015; 32: 1389–1471.

96. Zhang T, Cheng L, Liu L, et al. Asymmetric organocatalytic N-nitroso-aldol reaction of oxindoles. *Tetrahedron Asymmetry* 2010; 21: 2800–2806.

97. Jia LN, Huang J, Peng L, et al. Asymmetric hydroxyamination of oxindoles catalyzed by chiral bifunctional tertiary amine thiourea: Construction of 3-amino-2-oxindoles with quaternary stereocenters. *Org Biomol Chem* 2012; 10: 236–239.

98. Çamlısoy Y. *Development of Cinchona Based Catalysts Guided By Computational Insights For Asymmetric Synthesis*. Yeditepe University, 2023.

99. Unpredicted binding preferences were also detected in the matching step (Chapter 4.2.5) of the developed protocol and added to Figure 4.6.

100. Labute P. LowModeMD - Implicit low-mode velocity filtering applied to conformational search of macrocycles and protein loops. *J Chem Inf Model* 2010; 50: 792–800.

101. Tsuzuki S, Honda K, Uchimaru T, et al. Origin of attraction and directionality of the π/π interaction: Model chemistry calculations of benzene dimer interaction. *J Am Chem Soc* 2002; 124: 104–112.

102. Sinnokrot MO, Valeev EF, Sherrill CD. Estimates of the ab initio limit for π-π interactions: The benzene dimer. *J Am Chem Soc* 2002; 124: 10887–10893.

103. Bader RFW. A Quantum Theory of Molecular Structure and Its Applications.

104. Arnold WD, Oldfield E. The chemical nature of hydrogen bonding in proteins via NMR: J-couplings, chemical shifts, and AIM theory. *J Am Chem Soc* 2000; 122: 12835–12841.

105. Zhou PP, Qiu WY. Red-shifted hydrogen bonds and blue-shifted van der waals contact in the standard watson-crick adenine-thymine base pair. *Journal of Physical Chemistry A* 2009; 113: 10306–10320.

106. Wang Y, Qiao Y, Lan Y, et al. Predicting the origin of selectivity in NHC-catalyzed ring opening of formylcyclopropane: A theoretical investigation. *Catal Sci Technol* 2021; 11: 332–337.

107. Czyznikowska Z, Göra RW, Zaleśny R, et al. Structural variability and the nature of intermolecular interactions in watson-crick B-dna base pairs. *Journal of Physical Chemistry B* 2010; 114: 9629–9644.

108. Wang Y, Tang M, Wang Y, et al. Insights into stereoselective aminomethylation reaction of α,β-unsaturated aldehyde with N,O-acetal via N-heterocyclic carbene and brønsted acid/base cooperative organocatalysis. *Journal of Organic Chemistry* 2016; 81: 5370–5380.

109. Zhang QY, Li SJ, Wang Y, et al. Insights into the chiral sulfide/selenide-catalyzed electrophilic carbothiolation of alkynes: Mechanism and origin of axial chirality. *Organic Chemistry Frontiers* 2021; 8: 1983–1990.

110. Qiao Y, Yang W, Wei D, et al. Theoretical investigations toward TMEDA -catalyzed [2 + 4] annulation of allenoate with 1-aza-1,3-diene: Mechanism, regioselectivity, and role of the catalyst. *RSC Adv* 2016; 6: 70723–70731.

111. Lefebvre C, Rubez G, Khartabil H, et al. Accurately extracting the signature of intermolecular interactions present in the NCI plot of the reduced density gradient: Versus electron density. *Physical Chemistry Chemical Physics* 2017; 19: 17928–17936.

112. Emamian S, Lu T, Kruse H, et al. Exploring Nature and Predicting Strength of Hydrogen Bonds: A Correlation Analysis Between Atoms-in-Molecules Descriptors,

Binding Energies, and Energy Components of Symmetry-Adapted Perturbation Theory. *J Comput Chem* 2019; 40: 2868–2881.

APPENDIX A: COMPLETE ENERGETICS FOR ALL REACTANTS, TRANSITION STATES, AND PRODUCTS

Table A.1. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of bifunctional acid-base catalyzed model transition states of *trans-β*-nitrostyrene and dimethyl malonate

Structure	ΔE_0	ΔH	ΔG	$\Delta G_{Quasi\ corr.}$
Reactants	0.0	0.0	0.0	0.0
$TS_{Theozyme\ 2S}$ / $TS_{Theozyme\ 2S}$ BMA	-20.4	-20.2	13.4	17.2
$TS_{Theozyme\ 2S}$ BMB	-18.5	-18.5	15.6	19.4
$TS_{Theozyme\ 2O}$	-17.4	-16.5	14.4	19.3
$TS_{Theozyme\ 1S}$	-19.0	-19.2	15.0	17.6

Table A.2. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of bifunctional acid-base catalyzed transition states of *trans-β*-nitrostyrene and dimethyl malonate

Structure	ΔE_0	ΔH	ΔG	$\Delta G_{Quasi\ corr.}$
Reactants	0.0	0.0	0.0	0.0
***(R)*-TS$_{2S}$ / *(R)*-TS$_{2S}$ BMA**	-16.6	-16.9	9.2	11.6
***(S)* -TS$_{2S}$**	-14.2	-14.8	12.7	14.2
***(R)*-TS$_{2S}$ BMB**	-12.2	-12.4	13.3	16.1
***(R)*-TS$_{1S}$**	-5.7	-6.0	19.0	22.5
***(S)*-TS$_{1S}$**	-1.7	-2.1	23.2	26.5
***(R)*-TS$_{3S}$**	-19.5	-19.9	7.2	8.6
***(R)*-TS$_{3S\text{-}conf}$**	-14.4	-14.8	12.3	13.5

Table A.3. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of energy profile for bifunctional thiourea catalyzed transition states of *trans-β*-nitrostyrene and dimethyl malonate

Structure	ΔE_0	ΔH	ΔG	$\Delta G_{Quasi\ corr.}$
Reactants	0.0	0.0	0.0	0.0
***(R)*-2S-complex**	-16.9	-17.2	9.6	11.0
***(S)-2*S-complex**	-17.3	-17.4	8.5	10.7
***(R)*-TS_{2S}**	-16.6	-16.9	9.2	11.6
***(S)*-TS_{2S}**	-14.2	-14.8	12.7	14.2
***(R)-2*S-product**	-26.3	-26.8	-0.3	2.1
***(S)-2*S-product**	-23.0	-23.4	2.6	5.5

Table A.4. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of model transition states of *N*-benzyl-protected 3-substituted-2-oxindole with nitrosobenzene

Structure	ΔE_0	ΔH	ΔG	$\Delta G_{Quasi\ corr.}$
Reactants	0.0	0.0	0.0	0.0
Model A1 / Model A1(a)	-4.8	-4.9	27.2	28.9
Model A1(b)	0.9	0.8	32.2	34.7
Model A2	-6.1	-6.0	25.8	27.7
Model H1	-5.1	-5.5	27.3	29.2
Model H2	-13.4	-13.3	18.0	20.4

Table A.5. The B3LYP-D3/6-311+G(d,p)//B3LYP-D3/6-31+G(d) energetics (kcal/mol) of dimeric quinidine catalyzed transition states of *N*-benzyl-protected 3-substituted-2-oxindole with nitrosobenzene

Structure	ΔE_0	ΔH	ΔG	$\Delta G_{Quasi\ corr.}$
Reactants	0.0	0.0	0.0	0.0
***(R)*-TS$_{A1.1}$ / *(R)*-TS$_{A1.1(a)}$**	-19.8	-20.4	8.2	9.1
***(R)*-TS$_{A1.1(b)}$**	-9.5	-9.9	17.6	19.2
***(S)*-TS$_{A1.1}$**	-11.1	-11.2	15.7	17.5
***(R)*-TS$_{A1.2}$ / *(R)*-TS$_{A1.2(a)}$**	-17.4	-17.9	10.2	11.4
***(S)*-TS$_{A1.2}$**	-12.4	-12.9	15.3	16.4
***(R)*-TS$_{A1.3}$**	-8.1	-8.6	19.3	20.8
***(S)*-TS$_{A1.3}$ / *(S)*-TS$_{A1.3(a)}$**	-13.9	-14.3	13.1	15.0
***(S)*-TS$_{A1.3(b)}$**	-12.9	-13.3	14.3	15.9
***(R)*-TS$_{A2.1}$**	-16.0	-16.4	10.2	12.9
***(S)*-TS$_{A2.1}$ / *(S)*-TS$_{A2.1(a)}$**	-16.8	-17.1	9.7	11.9
***(S)*-TS$_{A2.1(b)}$**	-11.1	-11.2	15.6	17.4
***(R)*-TS$_{A2.2}$**	-22.5	-23.1	6.8	6.1
***(S)*-TS$_{A2.2}$**	-19.1	-19.4	8.3	9.5
***(R)*-TS$_{A2.3}$**	-15.2	-15.7	12.2	13.6
***(S)*-TS$_{A2.3}$**	-18.9	-19.3	9.3	9.8
***(R)*-TS$_{H2.2}$**	-17.6	-18.0	9.3	10.9

APPENDIX B: DISTORTION/INTERACTION ANALYSIS

Table B.1. Distortion/Interaction Analysis of ***(R)*-TS$_{2O}$** & ***(R)*-TS$_{2S}$** *

	ΔE_{dist} (dimethyl malonate)	ΔE_{dist} (*trans-(β)*-nitrostyrene)	ΔE_{dist} (catalyst)	ΔE_{int}	TS C-C lentgh (Å)
***(R)*-TS$_{2O}$**	5.6	10.2	6.4	-93.6	2.26
***(R)*-TS$_{2S}$**	5.1	9.3	11.5	-93.4	2.30

* The activation energies, ΔE, the distortion energies, $\Delta E_{dist,}$ and the interaction energies, ΔE_{int}, are given in kcal/ mol.

Table B.2. Distortion/Interaction Analysis of **TS$_{Theozyme\ 2O}$** & **TS$_{Theozyme\ 2S}$** *

	ΔE_{dist} (dimethyl malonate)	ΔE_{dist} (*trans-(β)*-nitrostyrene)	ΔE_{dist} (catalyst-DMAP)	ΔE_{dist} (catalyst-urea/thiourea)	ΔE_{int}	TS C-C lentgh (Å)
TS$_{Theozyme\ 2O}$	6.7	11.4	0.7	0.5	-98.6	2.19
TS$_{Theozyme\ 2S}$	5.8	9.9	0.7	0.4	-98.7	2.26

* The activation energies, ΔE, the distortion energies, $\Delta E_{dist,}$ and the interaction energies, ΔE_{int}, are given in kcal/mol.

Table B.3. Distortion/Interaction Analysis of ***(R)*-TS$_{A2.2}$** & ***(R)*-TS$_{H2.2}$** *

	ΔE_{dist} (oxindole)	ΔE_{dist} (nitrosobenzene)	ΔE_{dist} (catalyst)	ΔE_{int}	TS C-C lentgh (Å)
***(R)*-TS$_{A2.2}$**	10.5	8.8	26.6	-97.6	1.84
***(R)*-TS$_{H2.2}$**	9.5	9.1	37.5	-99.4	1.98

* The activation energies, ΔE, the distortion energies, $\Delta E_{dist,}$ and the interaction energies, ΔE_{int}, are given in kcal/ mol.

APPENDIX C: ENERGY PROFILES

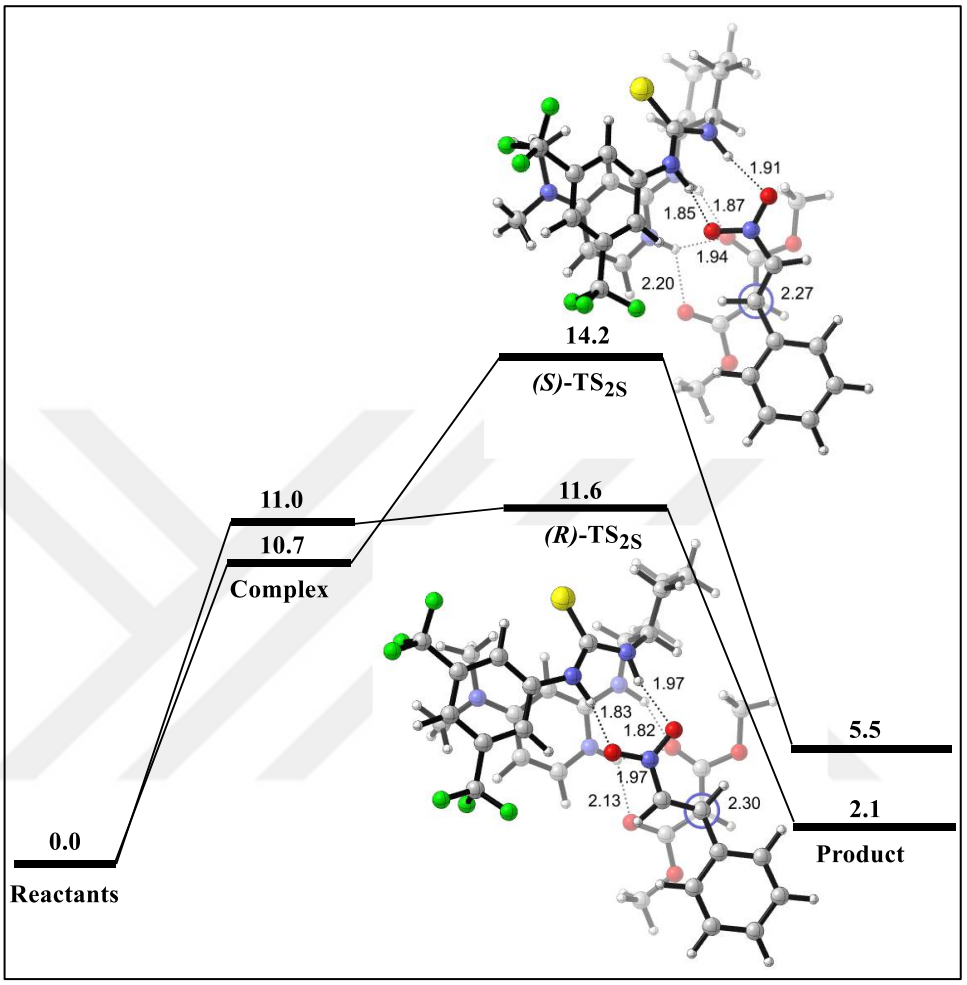


Figure C.1. Free energy profile ($\Delta G^{\neq}$) in kcal/mol for the addition reaction of dimethylmalonate to *trans*-*β*-nitrostyrene [B3LYP-D3/6-311+G(d,p)-SMD(Toluene)//B3LYP-D3/6-31G(d)-SMD(Toluene)]