

**TRANSHYDROGENASE:
UNDERSTANDING A MOLECULAR MACHINE**

BY PARVITAR SINGH MATHARU



A thesis submitted to the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Chemistry
College of Engineering and Physical Sciences
University of Birmingham
January 2014

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

Transhydrogenase couples hydride transfer between NADH and NADP⁺ to proton translocation across a membrane in animal mitochondria and bacteria. The product/reactant ratio can be driven to values greater than 400 by the energy of the proton gradient, demonstrating the importance of this gradient, and hence why transhydrogenase can be considered a molecular machine.

We isolated a number of possible transition states for the hydride transfer reaction in *R. rubrum* transhydrogenase, the lowest energy of which is symmetrical.

The nicotinamide ring orientation observed in the enzyme crystal structure most closely resembles an isolated asymmetrical transition state, not a symmetrical one. This observation is the basis for the asymmetric hypothesis. The asymmetric transition state hypothesis postulates that the asymmetry of the hydride transfer transition state for the reaction is responsible for the raised equilibrium constant.

We set out to probe the thermodynamics of the hydride transfer reaction using hybrid QM/MM (ONIOM) calculations on simple dIdIII model systems of *R. rubrum* transhydrogenase. Our results with these simple model systems allow us to infer that the whilst the asymmetry of the transition state is influential in altering the equilibrium constant of the hydride transfer reaction, the binding interactions of the surrounding protein environment also play a significant role.

Dedicated to my wonderful mother for never giving up on me, even when I was ready to give up on myself.

Acknowledgements

I would like to thank my Supervisor, Dr John Wilkie and the University of Birmingham for providing me with the fantastic opportunity to do this, and all the friends who provided hours of entertainment over many a rushed lunch time by distracting me when things weren't going so well and encouraging me when they were.

I would also like to thank Dr Graham Worth for his invaluable help with Fortran programming, without whom this project and by extension this thesis would not be possible.

Finally I would like to thank my family for all the love, support and sandwiches they provided over the last four years while I was doing this.

Contents

1. Introduction	1
1.1 Transhydrogenase as a Molecular Machine	2
1.2 The Biological Significance of the Hydride Transfer Reaction	3
1.3 <i>Rhodospirillum rubrum</i> Transhydrogenase Structure	5
1.4 Hydride Transfer and NADH/NADPH Analogues	12
1.5 Hydride Transfer and Nucleotide Binding Energies	13
1.6 Hydride Transfer and the Binding-Change Mechanism	15
1.7 The Asymmetric Hypothesis	19
1.8 The Research Challenge	24
2. Computational Resources and Methodology	25
<u>2.1 Computational Resources</u>	26
2.1-1 The Birmingham Environment for Academic Research	
(BlueBEAR) Cluster	26
2.1-2 Resources for Cluster Access and Computational	
Modelling	27

<u>2.2 Theoretical Principles</u>	28
2.2-1 Basic Molecular Orbital Theory	28
2.2-2 <i>Ab Initio</i> Hartree-Fock Molecular Orbital Theory	31
2.2-3 Density Functional Theory	35
<u>2.3 Computational Methodology</u>	50
2.3-1 The Hybrid DFT/HF Approach	50
2.3-2 The Transition State Search	51
2.3-3 Intrinsic Reaction Coordinate Calculations	53
2.4 Hybrid QM/MM (ONIOM) Methodology	56
2.5 Force Fields	57
2.6 Polarizable Continuum Model Calculations	59
3. Computational Modelling and Simulation	62
<u>3.1 Modelling Transition States For Hydride Transfer</u>	63
3.1-1 Generating N-methyl Pyridine Transition States	63
3.1-2 Generating Nicotinamide Hydride Transfer Transition States	65
<u>3.2 Transition State Nomenclature</u>	68
3.2-1 Common descriptors	68
3.2-2 The C ₂ Transition State Descriptors	69
3.2-3 The C _{2h} and C _{2v} Transition State Descriptors	70

<u>3.3 Modelling <i>Rhodospirillum rubrum</i> Transhydrogenase</u>	72
3.3-1 Preparing the Cofactors	72
3.3-2 Preparing the Protein Models	73
3.3-3 Modelling dIdIII Dimers	74
3.3-4 Modelling dI ₂ dIII ₂ Tetramers	75
3.3-5 The Amide Orientation of NAD(H)	77
 <u>3.4 Molecular Dynamics Simulations</u>	 77
3.4-1 Stages of an MD Simulation	78
3.5 Generating ONIOM Input Files: Development and Implementation of New Programs	83
3.6 Protein Modelling for ONIOM Study	87
 4. Preliminary Work	 90
4.1 The Sequential Electron/Proton/Electron Transfer Mechanism	91
4.2 Energetics of the N-methyl Pyridine Transition States	96
4.3 Disproving the Symmetrical Transition State Hypothesis	101

5. Nicotinamide Hydride Transfer Transition States **107**

5.1 The Transition State Search 109

5.2 New Transition States 122

5.21 New C₂ Structures 122

5.2-2 New C_{2v} Structures 125

5.3 Second Order C_{2h} Transition States 127

5.4 Transition State Energies 130

5.5 Mirror Images and Duplicates 136

5.6 Reactants, Products and True Transition States 148

5.7 Transition States In Solution 154

5.8 Implications for Hydride Transfer 159

6. Reaction Thermodynamics: Geometric and Electrostatic Effects

Upon Hydride Transfer **171**

6.1 ONIOM Simulation on Model Protein Systems 175

6.2 ONIOM Study of dIdIII Dimers 176

6.2-1 XE-OO Conformations 177

6.2-2 XE-ON Conformations 200

6.3 Preliminary ONIOM Study of dI₂dIII₂ Tetramers 221

6.3-1 XE-OO Tetramers 223

6.3-2 XE-ON Tetramers 227

6.4 Key Conclusions for the dIdIII Dimeric Systems	230
7. Conclusions and Future Work	232
References	240
Supplementary Information	250
Appendix	254

Glossary

Biology Terminology

NAD⁺ / (H) – Nicotinamide adenine dinucleotide (reduced)

NADP⁺ / (H) – Nicotinamide adenine dinucleotide phosphate (reduced)

H₂NAD(P)(H) – 1, 4, 5, 6 tetrahydro analogue of NAD(P)(H)

Nic(H) – oxidised/reduced nicotinamide moieties of NAD(P)(H)

H₂Nic(H) - oxidised/reduced nicotinamide moieties of H₂NAD(P)(H)

FMN - Flavin mononucleotide

(Computational) Chemistry Terminology

AMBER - Assisted Model Building with Energy Refinement (force field)

DFT - Density Functional Theory

HF - Hartree-Fock

IEF (PCM) - Integral Equation Formalism

IRC - Intrinsic Reaction Coordinate

LDA - Local Density Approximation

MM - Molecular Mechanical

MO - Molecular Orbital

ONIOM - "Our Own N-layered Integrated Molecular Orbital and Molecular Mechanics"

PCM - Polarizable Continuum Model

PES - Potential Energy Surface

QM – Quantum Mechanical

SCF - Self-Consistent Field

TS – Transition State

TV – Transition Vector

UFF - Universal Force Field

CHAPTER ONE

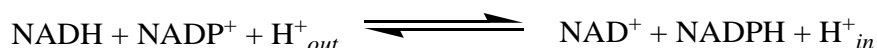
INTRODUCTION

This chapter provides a brief overview of the *R. rubrum* transhydrogenase structure and function, as well as a general review of previous studies of the hydride transfer reaction, and an outline of the motivation for this research.

1. Introduction

1.1 Transhydrogenase as a Molecular Machine

Transhydrogenase couples hydride transfer between NADH and NADP^+ to proton translocation across a membrane in bacteria and animal mitochondria [1]. The enzyme functions at a crucial three-way interface between NAD(H) and NADP(H) and the proton electrochemical gradient (Δp) and is a consumer of this gradient under the majority of physiological conditions in bacteria and animal mitochondria [1]:



Equation 1.1 – Transhydrogenase functions at the three-way interface between the two major soluble redox cofactors of the cell [NAD(H) and NADP(H)] and the electrochemical gradient, Δp [1]

The $\frac{[\text{NADPH}][\text{NAD}^+]}{[\text{NADP}^+][\text{NADH}]}$ ratio can be driven to values greater than 400 by the energy of Δp [1], compared to the H^+/H^- ratio of the enzyme which is believed to be only in the range of 0.5-1.0 in the absence of Δp – transhydrogenase is not a very efficient proton pump [2-4], thus demonstrating the importance of the proton electrochemical gradient, and hence why transhydrogenase can be considered a molecular machine. This is in contrast to more typical enzymes which are known to catalyse reactions without having any effect on the equilibrium constant, and illustrates why the hydride transfer reaction in the enzyme is so interesting and worthy of research.

1.2 The Biological Significance of the Hydride Transfer Reaction

The resultant NADPH from the hydride transfer reaction described above is used in the reduction and biosynthesis of glutathione, which is required for the damage limitation of the free radicals produced in the respiratory chain, for example [1, 5]. The physiological role of transhydrogenase also has a more understated characteristic within the mitochondria of various animal tissues such as heart and muscle, where predominantly the NADP-linked isocitrate dehydrogenase has a greater capacity to produce NADPH than transhydrogenase [1]. It is theorised that together transhydrogenase and the two NAD- and NADP-linked isocitrate dehydrogenases catalyse a microcycle between isocitrate and α -ketoglutarate that adjusts the regulation of the Krebs tricarboxylic acid pathway. The metabolic expense of such improved control is that the rate of consumption of Δp also increases [3].

While transhydrogenase catalyses a redox reaction, it is probable that its energy-coupling processes have more in common with those of F_1F_0 -ATP synthase than with those of the respiratory-chain complexes in the sense that firstly, in transhydrogenase and the F_1F_0 -ATP synthase, nucleotide-binding changes at the catalytic site are essential to the mechanism, secondly, in both proteins the nucleotide-binding site is located at a considerable distance from the proton translocation pathway and connection is achieved through conformational change [1]. According to a 2003 review by Jackson recent thermodynamic and kinetic experiments and considerations from recent X-ray structures of transhydrogenase components suggest that crucially coupling depends on movement of the nicotinamide rings of the nucleotides to either allow or block hydride transfer [1]. The nicotinamide rings can either be held in a proximal or distal conformation. (Figure 1.1).

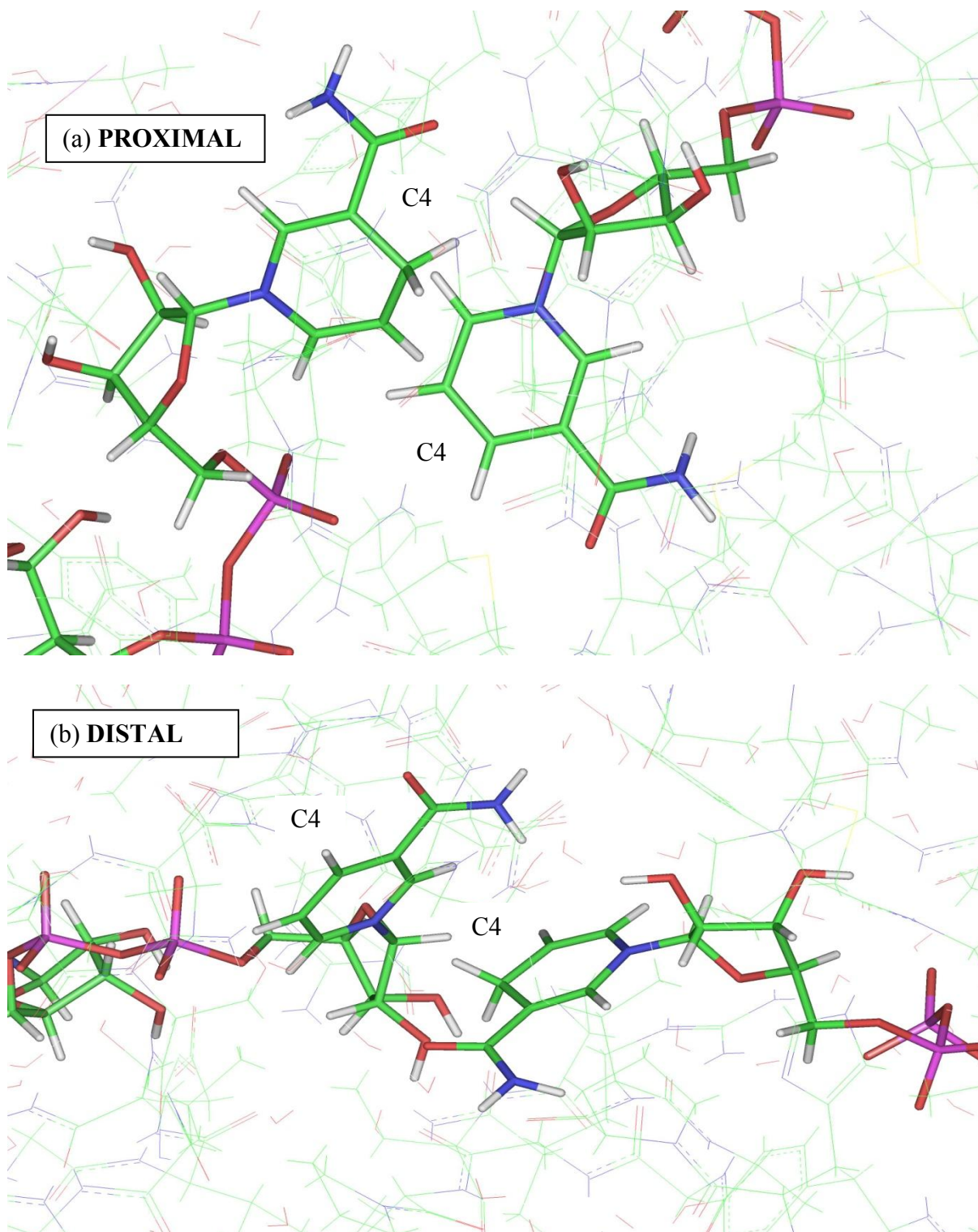


Fig 1.1 The two conformational states in the enzyme, the nicotinamide rings of NAD(H) and NADP(H) are either (a) held close together in a proximal conformation, allowing hydride transfer, or (b) held apart in a distal conformation, blocking hydride transfer.

In the proximal conformation (Figure 1.1 (a)), the nicotinamide rings of NAD(H) and NADP(H) are held close together (C4-C4 atom distance $< 4.00 \text{ \AA}$). In this state, hydride transfer can occur and is readily catalysed. Conversely, in the distal conformation (Figure 1.1 (b)), the nicotinamide rings of NAD(H) and NADP(H) are held far apart (C4-C4 atom distance $> 5.00 \text{ \AA}$). In this state, hydride transfer is impossible. This conformation corresponds to the enzymatic configuration either immediately after substrate binding or prior to product release.

Transhydrogenase is the only ion translocator to date where the standard free energies of the products are similar to those of the reactants [1], it is also important to note that one proton per hydride ion equivalent transferred from NADH to NADP^+ is translocated [1, 6]. The transhydrogenase reaction is also known to be stereospecific for the *C4A* (*pro-R*) hydrogen of the dihydronicotinamide ring of NADH and the *C4B* (*si*) face of the nicotinamide ring of NADP^+ [1, 6]. Furthermore, there is no inherent exchange of hydrogen between the water protons and the nucleotide during enzyme turnover [1, 6].

1.3 *Rhodospirillum rubrum* Transhydrogenase Structure

The global structure of transhydrogenase is known to be remarkably similar across a wide range of organisms. The *R. rubrum* structure, determined to a high resolution (2.60 \AA) [8] is discussed in detail here. Figure 1.2 below shows that both the dI component, which binds NAD^+/NADH , and the dIII component, which binds $\text{NADPH}/\text{NADP}^+$, protrude from the membrane – on the cytoplasmic side in bacteria, and on the matrix side of animal mitochondria [1, 6].

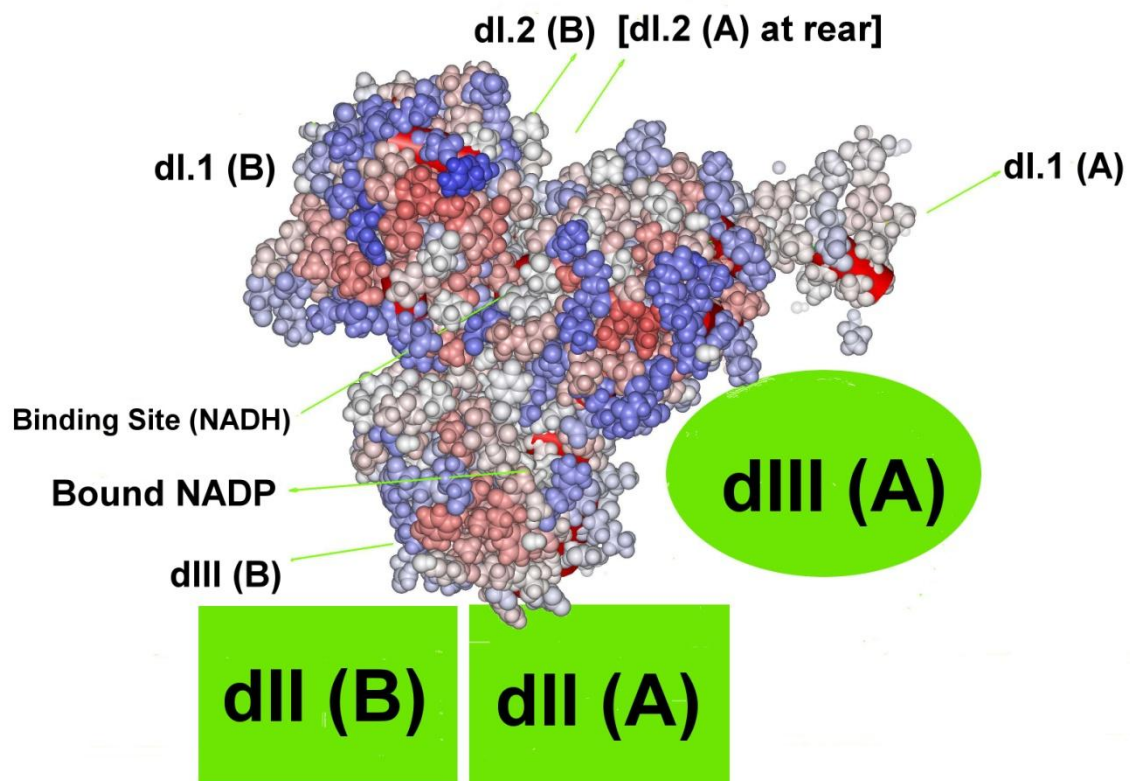


Fig. 1.2 The domain structure and some important secondary-structure features (helices in red are most prominent in this view) in transhydrogenase. The upper part shows a surface representation of the crystal structure of dI₂dIII₁ complex from *R. rubrum* transhydrogenase [8] coloured by hydrophobicity, generated by Insight II [10]. Bound NADP⁺ (stick format), is partially obscured by the ‘loop E’ lid in the dIII component but dI(B) lacks NAD(H) in the crystal structure. The lower part shows the likely position of the second dIII component and the two dII components of the intact enzyme in the membrane. Image adapted from [1].

The intact enzyme is a dimer of two dI+dII+dIII (monomers) (A and B in Figure 1.2), and differences to this basic structure amongst species arise from variations in the arrangement of polypeptides [1, 6].

The dI component is made up of two domains, dI.1 and dI.2 [7-9]. Both consist of parallel twisted β -sheets flanked by helices and take the form and connectivity of the Rossman fold (Figure 1.3) [8].

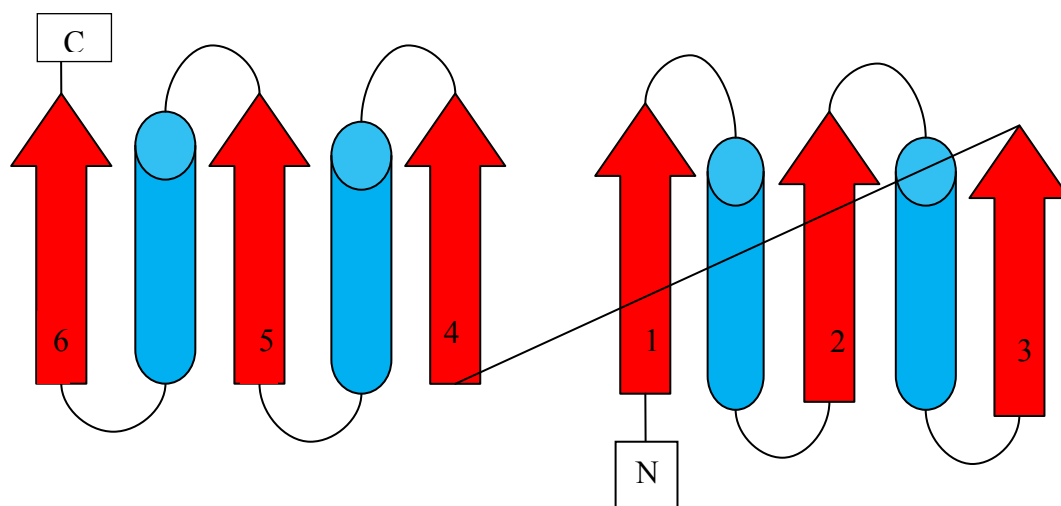


Fig. 1.3 A schematic representation of a Rossman fold. A common structural motif found in proteins that bind nucleotides, especially the NAD cofactor. The structure of two repeats consists of six β -strands (red arrows) which run in parallel, linked to two pairs of α -helices (blue cylinders). Each Rossman fold can bind one nucleotide, so in order to bind NAD (a dinucleotide), the binding domains are contain two paired Rossman folds that each bind one nucleotide moiety of the molecule. Single Rossman folds can bind mononucleotides such as FMN.

A deep cleft separates these two domains, which are connected by the long $\alpha 6$ and $\alpha 11$ helices and furthermore the NAD^+/NADH binding site exists within this cleft [8]. In the loop of the N-terminus of the ‘pyrophosphate-binding’ dI.2 helix, the characteristic GXGXXG ‘fingerprint’ of a class of nicotinamide nucleotide-binding proteins [11] exists. The specificity of NAD^+/NADH over $\text{NADP}^+/\text{NADPH}$ is thought to arise from the

hydrogen bonding between Asp-202 and the 2'-OH of the adenosine ribose [1] (Figure 1.4).

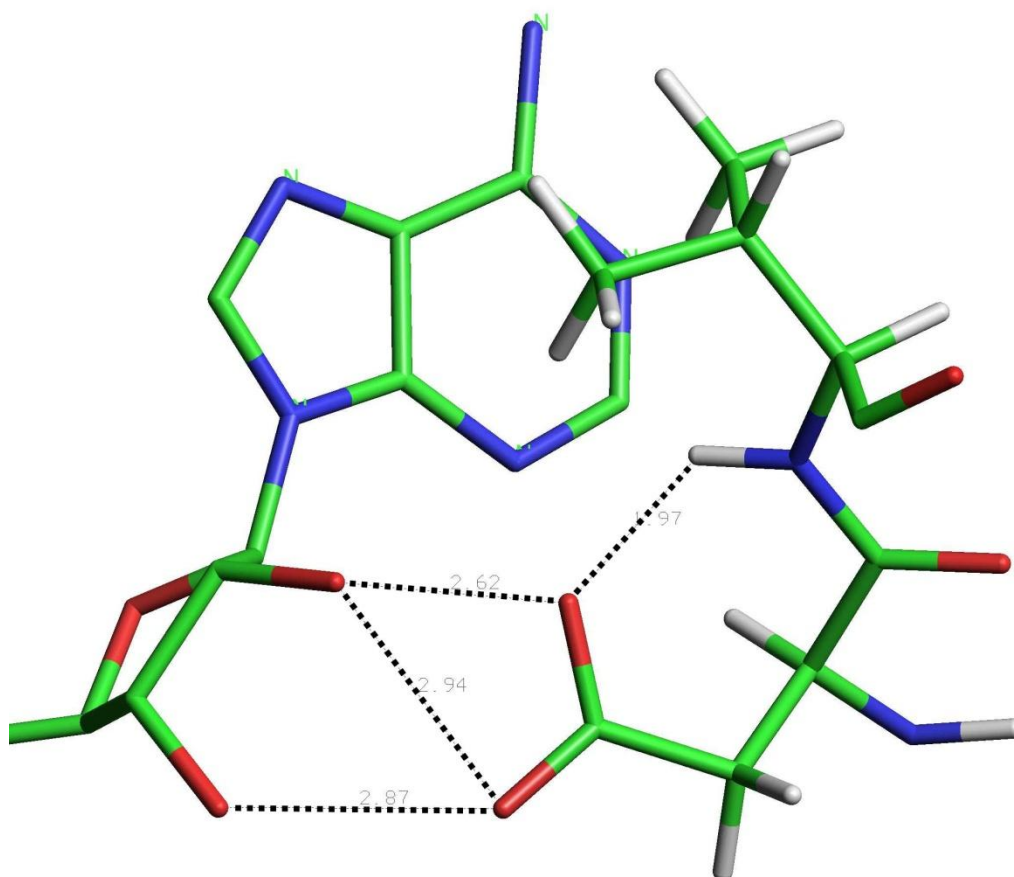


Fig. 1.4 Hydrogen bonding interactions between the 2'-OH of the adenosine ribose and Asp-202 is believed to give rise to the specificity of NAD^+/NADH over $\text{NADP}^+/\text{NADPH}$ [1]. Image generated by Insight II [10].

The dI polypeptides that comprise the dimer are related by a two-fold pseudo-symmetry axis which runs through the 'core' formed by the dI.2 domains. A β -hairpin of about thirteen residues spreads across from each dI.2 domain along the back of this symmetry-related subunit to contribute to its nucleotide-binding domain wall. [8].

Transhydrogenase dIII has one domain also with the structure and connectivity of the Rossman fold [12-14]. An atypical feature of bound NADP^+ is that it has a flipped orientation compared to the common nucleotides found in the Rossman fold, consequently the adenosine moiety is located over the second β - α - β - α - β motif of the fold and the nicotinamide mononucleotide of the first [8]. A GXGXXA sequence coinciding with the ‘fingerprint’ of NADP-dependent enzymes, but surprisingly the functional relationship with the bound nucleotide is different than in other proteins [1].

The $\text{NADP}^+/\text{NADPH}$ specificity over NAD^+/NADH arises from a network of hydrogen bonds between three residues (Lys-164, Arg-165 and Ser-166) and the adenosine ribose 2'-phosphate group [1] (Figure 1.5).

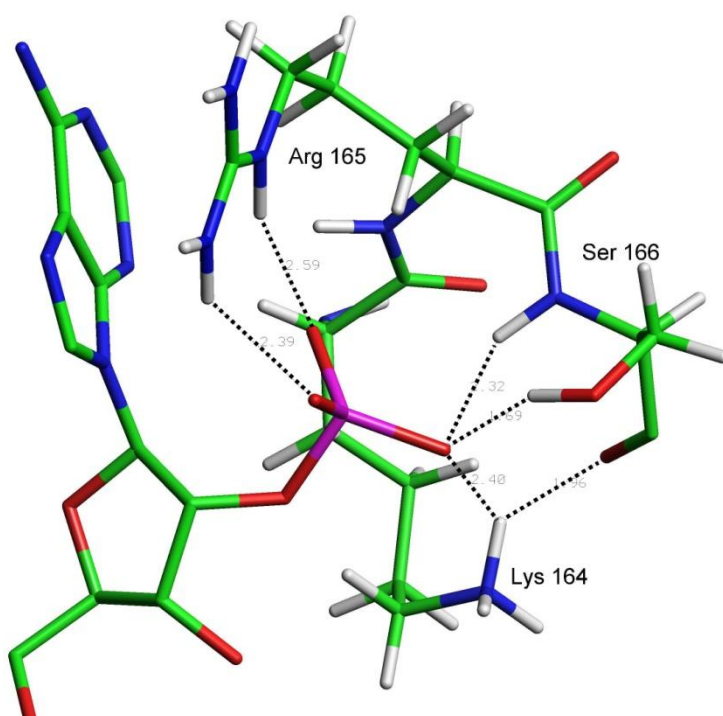


Fig. 1.5 The NAD^+/NADH specificity over $\text{NADP}^+/\text{NADPH}$ arises from a network of hydrogen bonds between three residues (Lys-164, Arg-165 and Ser-166) and the adenosine ribose 2'-phosphate group [1]. Image generated by Insight II [10]

Mixtures of separately prepared recombinant dI and dIII undergo spontaneous complexation to form dI_2dIII_1 [14, 15], and the crystal structure of this complex shows that the dIII polypeptide is linked with the cleft of only one of the two dI polypeptides (B in Figure 1.2). The dIII polypeptide extensively interacts with the dI.2 core of the dI dimer. The few interactions between dIII and dI.1(B) involve highly conserved regions of the enzyme and are thought to be essential in the spread of changes in conformation during the coupling of proton translocation to the redox reaction. The dI/dIII interface in the complex contains the hydride transfer site [8].

It is easy to explain the reason for the absence of the dII component from the X-ray crystal structure shown in Figure 1.2 as simply that because this component spans the membrane it is unobservable in crystal structures, but the reason why a dI_2dIII_1 heterotrimer is formed readily in solution, whilst a dI_2dIII_2 tetramer is not is a little more complicated.

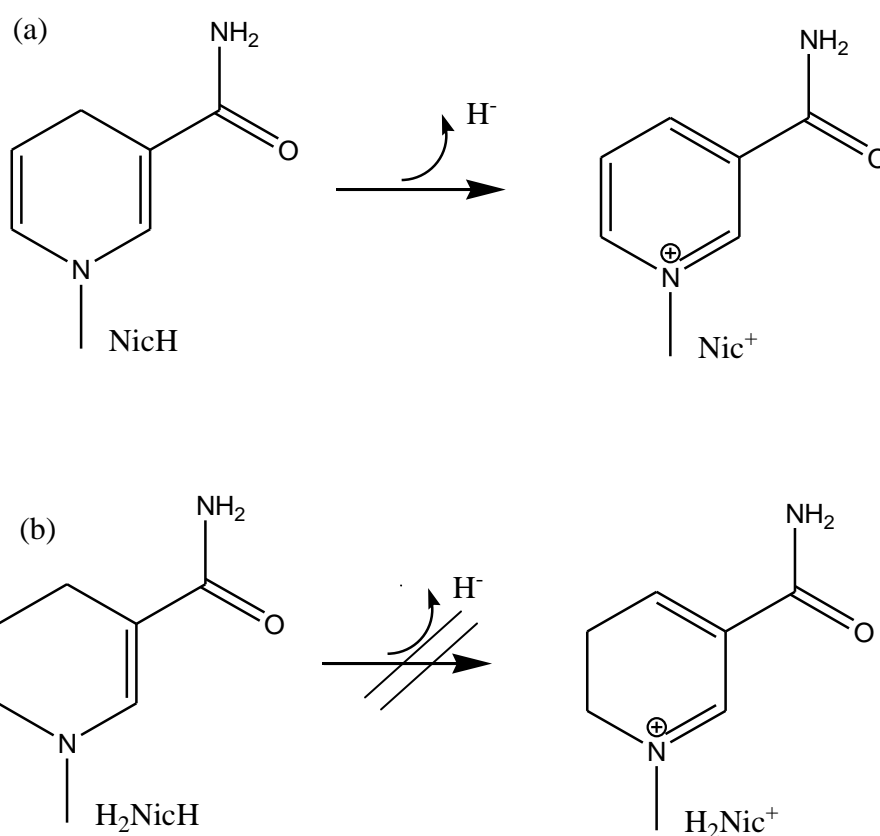
The position of the single dIII component is at the side of the cleft of a single dI and the binding loop of the $NADP^+$ nicotinamide on dIII is inserted into the cleft of the dI, clearly a second dIII must be able to interact with the second dI component in the intact enzyme during turnover. Crystallographic modelling studies have shown that the docking of a second dIII component to the dI_2dIII_1 heterotrimer complex in a conformation identical to the first is blocked by the clashing of amino acid side-chains at the dIII/dIII interface, explaining at least in part, why the tetramer is not formed during crystallisation [15]. The major consequence of this is that the conformation of the second binding site, that is to say whether it is proximal or distal in nature, is unknown.

Hydropathy analysis of amino-acid sequences [16-18] and the engineering of unique Cys residues in the protein [19, 20] suggest that the mammalian single-subunit

transhydrogenases have fourteen transmembrane helices (28 per dimer), two-subunit bacterial transhydrogenases, from *E. coli* for example, have 13 (26 per dimer) and three-subunit bacterial transhydrogenases, from *R. rubrum* for example, have 12 (24 per dimer). Numbering the helices from mammalian transhydrogenases, numbers 3, 4, 9, 10, 13 and 14 are seen to be the most conserved, however little evidence exists to suggest the nature of these helical clusters, though predictions based on the location of these conserved amino-acid residues have been made in [20, 21]. The loops between the transmembrane helices are thought to be short, less than 14 residues on the cytoplasmic side and less than 25 residues on the periplasmic side in bacteria. The cytoplasmic loops are more conserved and some of these interact with dIII [1].

The dI_2dIII_1 complex (Figure 1.2 [8]) can be used to infer the overall arrangement of components in the intact enzyme. A second structurally identical dIII, likely to be in a conformationally different state (distal) from that observed in the trimeric complex where the dIII(B) is proximal to its dI(B), must be located at or in close proximity to the vacant dI(A) cleft. The two dII components are likely to be coupled with the two dIII components. It is also likely that there is no interaction between dI and dII in the three-subunit transhydrogenases, consequently dI can be readily displaced from the intact enzyme and subsequent dI reconstitution allows reformation of the intact enzyme to occur [22]. A single strand of unconserved polypeptide chain provides a direct interaction between dI and dII in both the one- and two-subunit transhydrogenases [1,8].

1.4 Hydride Transfer and NADH/NADPH Analogues



Scheme 1.1 NicH and Nic $^+$ represent the reduced and oxidized nicotinamide moieties of the physiological nucleotides, NAD(P)(H). H_2NicH and H_2Nic^+ represent the reduced and biologically oxidized moieties of the analogue nucleotides, $\text{H}_2\text{NAD(P)(H)}$. The results of the study in [6] show that despite having similar binding affinities, the analogues of the natural nucleotides cannot donate hydride for transhydrogenation (b). Adapted from [6].

Studies of dI_2dIII_1 complexes have shown that they catalyse a swift single-turnover burst of transhydrogenation. [23-26]. When dII is not present, further turnover is restricted due to the slow rate of product release (dependent on direction, NADP^+ or NADPH) from dIII.

Later studies in [6] with the complex isolated from *R. rubrum* suggest that the 1,4,5,6-tetrahydro analogue (H_2NADH) binds to dI with similar affinity to NADH and that when bound, both lead to the closure of the dI mobile loop. It was also observed that though H_2NADPH also bound tightly to dIII, the dissociation rate constant increased compared to NADPH. Substitution of dIII NADP^+ with either NADPH or its 1,4,5,6-tetrahydro analogue resulted in similar chemical shift changes, indicating equivalent conformational alterations. Ultimately this study [6] concludes that irrespective of having similar binding affinities, the analogues of the natural nucleotides could not donate hydride for transhydrogenation (Scheme 1.1 (b)).

Similar studies on other NADH analogues such as Acetylpyridine adenine dinucleotide (AcPdADH) are shown to be better in transhydrogenation but that the optical absorbance is shifted relative to the natural nucleotide. Consequently, it was observed that for the reaction of AcPdAD^+ and NADPH on dI_2dIII_1 complexes, the kinetics of NADPH oxidation and AcPdAD^+ reduction are identical [23]. The AcPdADH and NADP^+ reaction kinetics of AcPdADH oxidation and NADP^+ reduction tell a similar story, in that they too coincide exactly [25]. This demonstrates hydride transfer is direct and that there are no redox intermediates in transhydrogenase.

It can be concluded that the *pro*-R C4H of the NADH dihydronicotinamide ring must be juxtaposed to the *si* face of C4 of the NADP^+ nicotinamide at the redox step during the physiological reaction of the intact enzyme.

1.5 Hydride Transfer and Nucleotide Binding Energies

Furthermore, (dihydro)nicotinamide and nicotinamide rings are known to be quite reactive. In aqueous solution without enzymes the rate of reaction between oxidised and reduced nucleotides and their analogues is significant - 10^{-2} - $10^{-3} \text{ M}^{-1} \text{ s}^{-1}$ are the typical orders of

second-order rate constants [27, 28]. The non-enzymatic rate is not significant at the low nicotinamide nucleotide concentrations present in most transhydrogenase experiments and within biological cells.

The consequence of the coupling mechanism reactivity of transhydrogenase suggests that during nucleotide binding to the enzyme, both the dihydro- and nicotinamide rings of NADP^+ and NADH respectively have to be kept away from each other. An unfavourable redox reaction would occur uncoupled from proton translocation and driven by the solution ΔG of the nucleotides if NADH and NADP^+ were allowed to bind with the C4 atoms of their juxtaposed nicotinamides [1]. The operating environment of the transhydrogenase active site (bacterial cytoplasm or mitochondrial matrix) is naturally a solution in which NAD^+/NADH is oxidised whilst $\text{NADP}^+/\text{NADPH}$ is usually significantly reduced (i.e. the product state of transhydrogenase). In its open state, transhydrogenase is expected to bind both NADH and NADP^+ with high affinity relative to NAD^+ and NADPH , to promote the binding of substrate and the release of product. The caveat here however is that experimental K_d values to support this supposition are unavailable [1]. If during conversion to the occluded state these relative affinities were to remain constant, then the equilibrium constant for hydride transfer would be unfavourable, essentially trapping the enzyme in a ‘thermodynamic pit’, whereby the rate at which transhydrogenase operated as a fully coupled proton translocator would be significantly lowered [29]. It has been theorized that this problem is avoided by a change in the nucleotide binding energies of the two nucleotides during inter-conversion of the open and occluded states. Nucleotide occlusion is necessary not only to prevent the redox reaction being uncoupled from proton transfer, but also for the change in binding affinities. These changes can then allow destabilisation of the newly-bound substrates – NADP^+ and NADH without either dissociating from the protein prior to proton

translocation [1]. This is the key to the binding-change mechanism as described in the next section, however since the evidence for this being the case is unconvincing, the counter proposal of the asymmetric hypothesis described in section 1.7 is offered as an alternative rationale in the course of this research.

1.6 Hydride Transfer and the Binding-Change Mechanism

Quantum-mechanical calculations (AM1 [30]) reveal that the energetic cost of puckering the planar 1,4-dihydronicotinamide ring into a quasi-boat conformer ($\alpha_N=\alpha_C=10^\circ$) is 17 kcal/mol, this is relatively small compared to the energy decrease of the TS for hydride transfer of dogfish lactate dehydrogenase experienced through deformation [31]. Ring puckering forces the *pro*-R C4 atom to be pseudoaxial, thereby probably increasing the driving force for hydride transfer through greater orbital overlap during the reaction. The kinetic significance of this effect was illustrated by MD calculations on a number of dehydrogenases, the net result of which indicate the partiality of the 1,4-dihydronicotinamide C4H atom to be transferred so as to pucker towards the substrate-binding pocket [30]. The rate of transhydrogenase hydride transfer between NADH and NADP⁺ is very swift - k_{app} is approximately 2100 s⁻¹ in dI₂dIII₁ complexes [24]. This is considerably quicker than either soluble dehydrogenase substrate oxidation or reduction by NAD(P)(H). Even though rather modest rates of transhydrogenase hydride transfer are expected due to nucleotide binding and orientation, nicotinamide ring activation through puckering, is not impossible [1, 31].

Crystal structures in [7-9, 14-16] allowed the conserved residues in the hydride transfer site to be identified. The almost invariant run of 10 residues between Ile-84 and Gly-93 in

the NADP⁺ nicotinamide binding loop in dIII, the conserved tyrosine residues (Tyr-235 in dI and Tyr-171 and Tyr-54 in dIII) and the invariant charged polar residues, Arg-127, Gln-132, Asp-135 and Ser-138 (in the RQD loop) [8] of dI, which can all have interactions with NAD(H) nicotinamide, are of particular interest.

These residues will partake in the positioning the dihydro- and nicotinamide rings of NADH and NADP⁺ respectively, for hydride transfer during occluded state formation, as described by the binding-change mechanism (Figure 1.6).

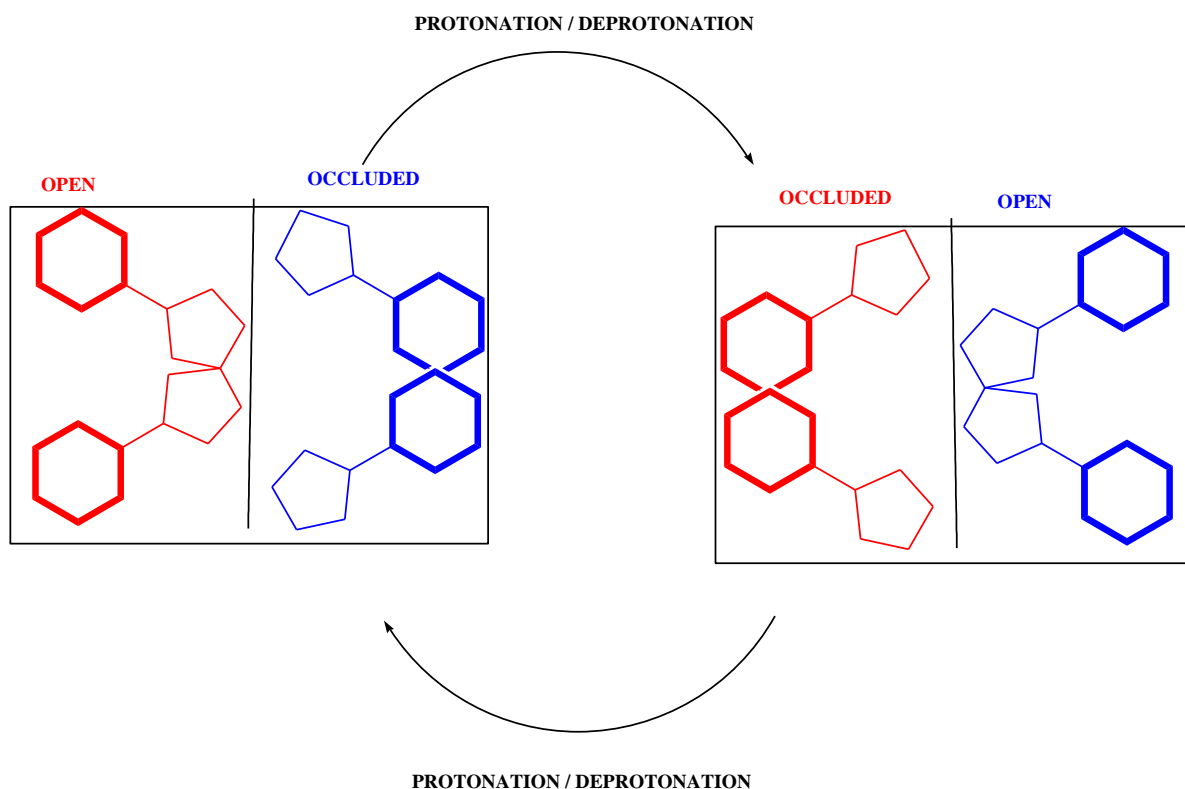


Fig. 1.6 A simplified illustration of the binding-change mechanism of transhydrogenase. In the 'open' state, exchange between bound nucleotides and those in the solvent is possible and hydride transfer between NADP^+ and NADH is blocked - the bold dihydronicotinamide and nicotinamide rings are held apart (distal). Conversely, in the 'occluded' state, exchange between bound nucleotides and those in the solvent is no longer possible and hydride transfer is allowed due to both bold rings being brought close together in a proximal fashion. Protonation and deprotonation reactions related to proton translocation through dII interconvert between these two 'open' and 'occluded' states as shown. The conformational states of the two monomers of transhydrogenase (one red, one blue) alternate during turnover - adapted from [1].

The binding-change mechanism as outlined above essentially describes two states of transhydrogenase. In the 'open' state, bound nucleotides can exchange with those in the solvent, blocking hydride transfer between NADH and NADP^+ as the dihydronicotinamide and nicotinamide rings are held apart in a distal configuration.

Conversely, in the 'occluded' state, bound nucleotides cannot exchange with those in the solvent allowing hydride transfer as both rings are held close together in a proximal configuration. Protonation and deprotonation reactions related to proton translocation through dII interconvert between these two 'open' and 'occluded' states as shown in Figure 1.6.

It has been proposed that following the closure of the site for the expulsion of water, they will also be responsible for the polarisation shown by the raised redox reaction equilibrium constant [1].

The driving force for hydride transfer in the binding-change mechanism can either be when the nicotinamide rings are brought together in the occluded conformation (see Figure 1.6) or the proton gradient itself. If it is the former, this is known to favour the product, and so an extension to the binding-change mechanism to include a theoretical third state of transhydrogenase which is a reactant-like open state can be hypothesised. This third state would ensure a continuous hydride transfer cycle by shifting the equilibrium back to favouring reactant (and the pulling apart of the rings) after the rings have been brought together and the equilibrium favours the product.

As the intact enzyme is a dI_2dIII_2 tetramer, both hydride transfer sites are represented in the binding-change mechanism, leading to an overall two-electron transfer for hydride transfer to proton translocation, but when each site is considered individually, it can be seen that the 1:1 ratio for a single electron transfer process is maintained.

As yet, a true transhydrogenase hydride transfer mechanism – different from a sequential electron/proton/electron mechanism – is still not proven. However, investigation of the reaction through experiments with nucleotide analogues in solution seem to suggest single-

electron reduction products (of nicotinamide) and single-electron oxidation products (of dihydronicotinamide), which would be ‘high energy intermediates’ are not formed [33]. The energy required in generating $\text{NADP}^{+\bullet}$ and NADH^\bullet by a single-electron transfer process would likely be prohibitive [34], therefore hydride transfer is mechanistically more likely. Quantum-mechanical tunnelling of the hydride ion through the activation barrier is indicated by the temperature dependencies of the primary kinetic isotope effect for AcPdAD^+ reduction by $\text{NADP}^2\text{H}/\text{NADP}^1\text{H}$ in dI_2dIII_1 complexes [24]. Hydrogen tunnelling has been detailed for some soluble enzymes . It has also been suggested that thermal oscillations in the ground state of the enzyme-substrate complex may decrease the distance over which the hydrogen tunnels, consequently accelerating the reaction rate [35, 36].

1.7 The Asymmetric Hypothesis

In the course of this work I intend to offer a counter proposal to the binding-change mechanism discussed in Section 1.6, whereby it is theorized that the transition state itself is responsible for the raised equilibrium constant. This is the basis for the asymmetric hypothesis. So far it has been shown that transhydrogenase structure and function has been extensively studied and that most of the experimental evidence comes from the dI_2dIII_1 complexes isolated from the *R. rubrum* bacterium. Although the hydride transfer mechanism has been theorised for this enzyme, it has yet to be fully proven, and it is hoped that this work will go some way in addressing this issue, essentially by building upon some of the work conducted in [6] where computational modelling techniques and quantum-mechanical and molecular dynamics calculations were used to investigate the transition state for the hydride transfer between $\text{NAD}(\text{H})$ and $\text{NADP}(\text{H})$. The conclusions of this

2007 study are that the transition state resembles the arrangement of nucleotides in the transhydrogenase active site in the crystal structure, and that the molecular dynamics simulation of the enzyme and that the dihydro- and nicotinamide rings remain close to the ground state for hydride transfer throughout a 1.4 ns trajectory. Also the study in [6] shows that the transition states which most closely corresponds to the crystal structure are not the lowest energy transition states for a simple hydride transfer process. They also have asymmetric C-H bond distances, leading to the hypothesis that the asymmetrical nature of these transition states itself may be sufficient enough a driving force to change the equilibrium constant.

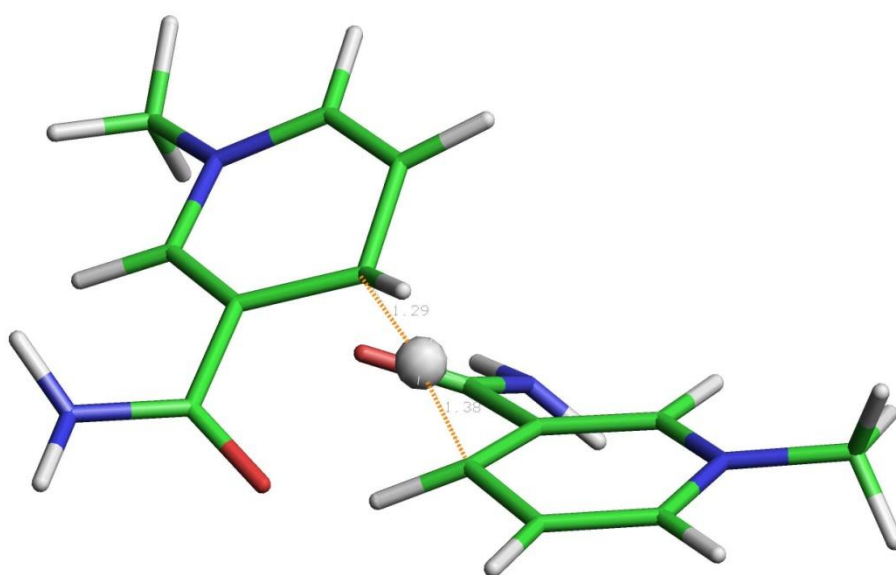


Fig. 1.7 - An asymmetric transition state for hydride transfer, C-H bond distances of 1.29 Å and 1.38 Å as shown.

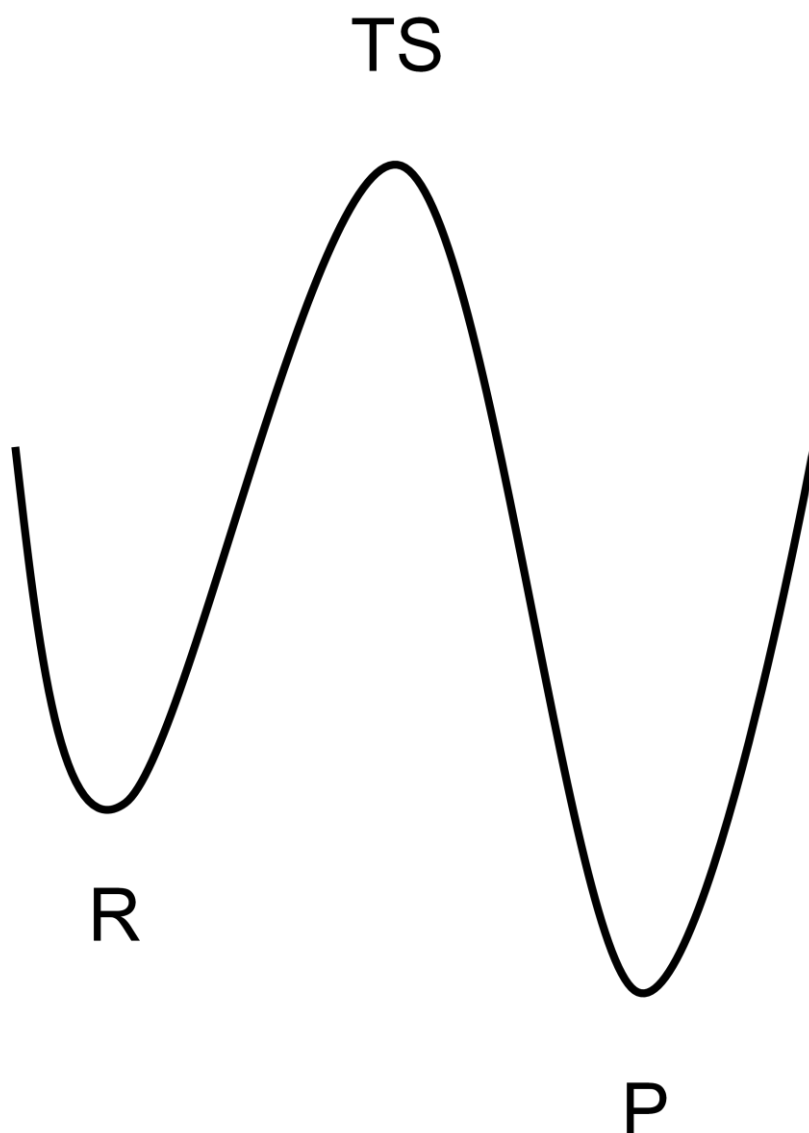
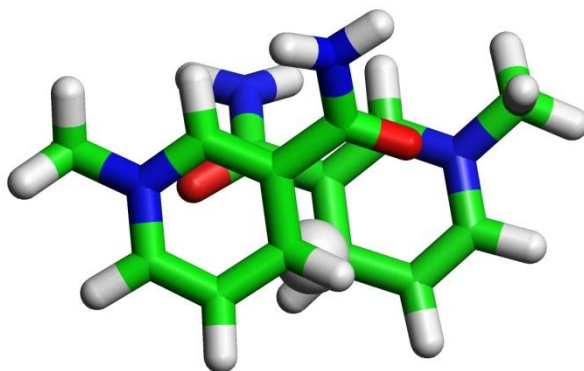


Fig. 1.8 - A typical reaction profile

Figures 1.7 and 1.8 respectively show the asymmetric nature of a transition state which closely corresponds to the crystal structure, and a typical reaction profile. The asymmetry of this transition state, which is not the lowest in energy, is readily apparent as the two C-H bond distances between the C4 of each ring and the transferring hydride are 1.29 Å and 1.38 Å respectively. It is this observation that has lead to the proposal of the asymmetric hypothesis.

An asymmetric transition state is indicative of an asymmetric reaction - a more reactant-like transition state corresponds to an exothermic reaction. The question to consider here then is whether an asymmetric transition state gives rise to different reactant and product energies when compared to fully symmetric transition state, and whether this sufficiently explains the change in reactant/product ratio brought about by the proton gradient.

(a) **Fully Symmetric Transition State** - *NOT* observed in the enzyme.



(b) **Asymmetric Transition State**

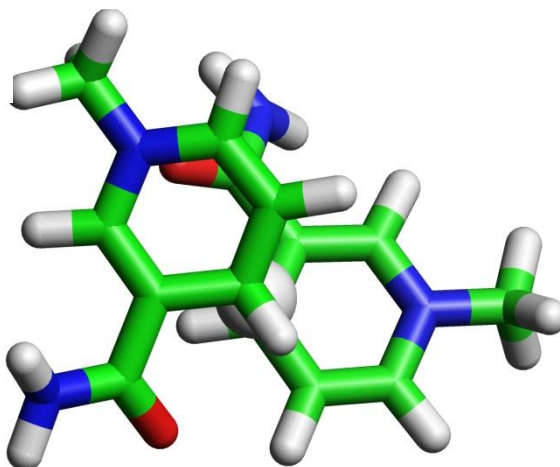


Fig. 1.9 Hydride Transfer Transition States: (a) fully symmetric transition state, both C-H bond distances, 1.34 Å, and (b) asymmetric transition state, C-H bond distances, 1.29 Å and 1.34 Å respectively.

Transition states are essentially the highest potential energy configurations along the reaction coordinate (see Figure 1.8), and as such resemble the reactants and products of a reaction, and as noted above, a more reactant-like transition state is indicative of an exothermic reaction, whilst a more product-like transition state is indicative of an endothermic reaction. As well as the reaction thermodynamics from the transition state for a given reaction, the geometry of the reaction - whether it is symmetric or asymmetric in nature - can also be inferred. As a transition state resembles the geometry of the starting and end points of a reaction, it follows that a symmetric reaction must have a symmetric transition state, conversely an asymmetric transition state must be indicative of an asymmetric reaction.

In the case of the hydride transfer reaction of transhydrogenase, the fully symmetric transition state conformation (both C-H bond distances of 1.34 Å, Figure 1.9 (a)) is not observed in the enzyme despite being the lowest energy conformation. The conformation that is observed (Figure 1.9 (b)) is the second lowest in energy and is an asymmetric transition state (C-H bond distances of 1.29 Å and 1.38 Å respectively, see Figure 1.7), thus indicating an asymmetric reaction and forming the basis for the asymmetric hypothesis. The striking difference between the fully symmetric and asymmetric transition states is that the rather neat overlap between the amide groups of both adjacent rings is lost in the asymmetric transition states leading to fewer electrostatic interactions within the structure when compared to the fully symmetric, however as described in section 1.8 below, the geometry of the transition state may not be the only factor which influences the observed change in the equilibrium constant of the hydride transfer reaction, there is also the ambiguity of the amide orientation of NAD(H) as arising from the crystal structure to consider.

1.8 The Research Challenge

As well as probing the validity of the asymmetric hypothesis described above, there is also the ambiguous nature of one of the bonds in the amide orientation observed in the crystal structure of the enzyme [6] – arising from the fact that in X-ray crystal structures, protons do not appear and both nitrogen and oxygen are indistinguishable – to consider.

Therefore the aim of this work is to use ONIOM calculations to determine the proper orientations of the substrate amide groups, and also to determine energies of the reactant, product and transition states for hydride transfer within the active site of the protein. Hybrid QM/MM (ONIOM) calculations of the reactant and product bound states in both distal and proximal configurations of the enzyme active site of *R. rubrum* transhydrogenase will probe the effects of ring geometry on hydride transfer in the enzyme, the effects of the electrostatics of the surrounding protein environment will also be considered. The results of these calculations will test the aforementioned hypothesis that the asymmetrical nature of the transition state itself is sufficient enough a driving force to change the equilibrium constant.

CHAPTER TWO

COMPUTATIONAL RESOURCES

AND METHODOLOGY

This chapter provides a brief overview of the computational resources used in this research, as well as the theoretical principles behind the implemented computational methodologies.

2. Methodology

In this chapter, the key computational methodologies and resources that were employed during the course of this work are discussed. The basic theoretical principles underlying the computation are also presented, so as to demonstrate that the study undertaken was not just the result of many ‘black box’ calculations over many months. Understanding the principles behind the chosen methods also allows for the advantages and disadvantages of them to be readily identified and thus the accuracy of the final results obtained and discussed in the subsequent chapters to be fairly evaluated.

2.1 Computational Resources

2.1-1 The Birmingham Environment for Academic Research (BlueBEAR) Cluster

The BlueBEAR cluster [37] was used in order to run the necessary calculations to investigate the reaction mechanism of proton translocation by *R. rubrum* transhydrogenase. The operating system is Scientific Linux 5.2. This cluster consists of:

- 3 logon nodes in a round-robin configuration for resiliency
- 384 dual-processor dual-core (4 cores/node) 64-bit 2.6 GHz AMD Opteron 2218 worker nodes giving a total of 1536 cores. Most of these nodes have 8 GB of memory whilst 16 of them have 16 GB

- 4 quad-processor dual-core (8 cores per node) 64-bit 2.6 GHz AMD Opteron 8218 nodes with 32 GB of memory
- Over 150 TB of raw disk space primarily allocated to BlueBEAR users using IBM's GPFS cluster file system

2.1-2 Resources for Cluster Access and Computational Modelling

The BlueBEAR cluster was accessed primarily through an SGI Workstation [38] running a UNIX operating system. The Insight II [10] molecular modelling software package was also run on this system and used to model the enzyme and build the initial transition state structures used in this study as detailed below. The PuTTY 0.60 (beta) telnet and SSH client [39] was also used for convenience as this provided access to BlueBEAR on machines running either Windows XP or Windows 7 through the Linux based toc filesystems at the University of Birmingham.

2.2 Theoretical Principles

The fundamental theoretical principles behind computational methods used in this work essentially fall into three areas, each of which will be discussed in turn below.

These are:

- Molecular Orbital (MO) Theory
- *Ab Initio* Hartree-Fock MO Theory
- Density Functional Theory (DFT)

2.2-1 Basic Molecular Orbital Theory [40]

The foundation of quantum mechanics is that a wave function, designated ψ , exists for any chemical system, and that the appropriate functions which act upon this wave function return the observable properties of the system, this relationship can be expressed mathematically in the form:

(2.1)

$$\hat{\theta}\psi = e\psi$$

In this equation $\hat{\theta}$ represents an operator and e represents a scalar value for some system property. When this relationship holds, ψ is defined as an eigenfunction, and e as an eigenvalue. The product of and its complex conjugate, $|\psi^*\psi|$ is the probability density. Since only real and not complex wave functions are considered here, the * symbol can be dropped, therefore the probability that a chemical system will be located in some region of multidimensional space equates to $\int |\psi|^2 d\tau$ over that region of space.

This means that an acceptable ψ has particular constraints. For a bound particle, the normalised $\int |\psi|^2 d\tau$ must be unity, which is to say that the probability of its location is one, which in turn requires that in addition to being continuous and single valued, ψ must also be quadratically integrable.

The Hamiltonian operator, H , of a system returns the system energy, E as an eigenvalue, giving the Schrödinger equation:

(2.2)

$$H\psi = E\psi$$

For simple systems, H takes into account the kinetic energies of the electrons and nuclei, the attraction between the electrons and nuclei as well as the inter-electronic and inter-nuclear repulsions. As the complexity of the system increases, for example when the spin-orbit coupling interactions in heavy elements are significant, additional terms in the Hamiltonian are necessary, and such is the case when MO is extended to include Hartree-Fock.

For many-particle systems, accurate wave functions are difficult to express because of the correlated (paired attractive and repulsive terms in H) motions of particles. The implication from this correlation is that no particle moves independently of all the others. The Born-Oppenheimer approximation simplifies this issue to a degree.

Typically the electrons of a system move significantly quicker than the nuclei, owing to the significant difference in their respective masses. Practically, electronic ‘relaxation’ is instantaneous with respect to nuclear motion, therefore it is convenient to decouple these

two motions, and compute the electronic energies for fixed nuclear positions instead, and thereby the kinetic energy term of the nuclei is independent of the electrons, the attractive nuclear- electronic potential energy term is eliminated, and the repulsive nuclear-nuclear potential energy term becomes a constant for a specific geometry, giving rise to the electronic Schrödinger equation:

(2.3)

$$(H_{el} + V_N)\psi_{el}(q_i; q_k) = E_{el}\psi_{el}(q_i; q_k)$$

The 'el' subscript in this equation highlights that the Born-Oppenheimer approximation has been taken into account. V_N is the nuclear-nuclear repulsion energy, q_i defines electronic coordinates which are independent variables, but the nuclear coordinates (q_k) are parameters. The eigenvalue of this equation is defined as the 'electronic energy', and V_N is constant for a fixed set of nuclear coordinates. In practice Equation 2.3 can almost always be solved without the V_N term because constants in wave functions have no effect. When this is done, the eigenvalue is defined as the 'pure electronic energy' and V_N can then be added to this to get E_{el} .

The Born-Oppenheimer approximation is an essential concept as it allows the potential energy surface (PES) to be defined. The PES is the surface over all possible nuclear coordinates, as defined by E_{el} . Without a PES, the concepts of equilibrium and transition state geometries would be lost, as they are by definition critical points on a PES; thus the advantages of incorporating the Born-Oppenheimer approximation are self-evident.

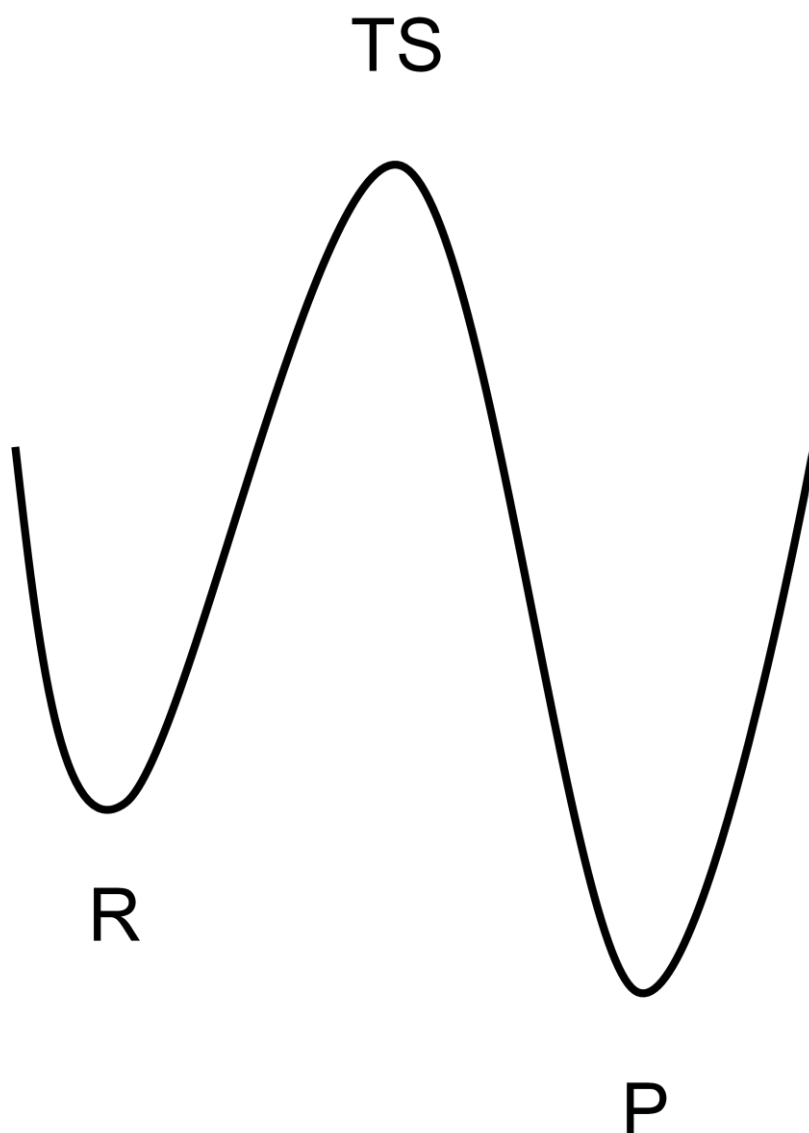


Fig. 2.1 – An illustrative reaction profile showing that a transition state exists at the peak of a potential energy curve, as opposed to a minimum which would exist in a potential energy trough.

2.2-2 *Ab Initio* Hartree-Fock Molecular Orbital Theory [40]

If the Schrödinger equation is taken in the context of a single-electron Hamiltonian, the only terms accounted for are the single-electron kinetic energy and nuclear attraction terms, and the operator itself is able to be separated:

(2.4)

$$H = \sum_{i=1}^N h_i$$

In the above equation, N is the total number of electrons and h_i is the single electron Hamiltonian:

(2.5)

$$h_i = -\frac{1}{2} \nabla_i^2 - \sum_{k=1}^M \frac{Z_k}{r_{ik}}$$

Eigenfunctions of Equation 2.5 must satisfy the following single-electron Schrödinger equation:

(2.6)

$$h_i \psi_i = \epsilon_i \psi_i$$

The many-electron eigenfunctions of the single-electron Hamiltonian can be constructed as products of single-electron eigenvalues ($\psi_{HP} = \psi_1 \psi_2 \dots \psi_N$) because it is separable, leading to the ‘Hartree Product’ (HP) wave function. The eigenvalue of ψ can be determined by the repeated application of Equation 2.6, which would demonstrate that the energy eigenvalue of the many-electron wave function is the sum of the single-electron energy eigenvalues.

The single-electron Hamiltonian does not incorporate inter-electronic repulsions, and thus the computation of these repulsions does not depend on a single electron, but rather on all

the possible simultaneous paired interactions. The goal is to locate orbital ψ which minimise the integral $\int \psi_{HP} H \psi_{HP} d\tau$. It can be demonstrated mathematically that each such orbital ψ_i is actually an eigenfunction of its own operator, h_i :

(2.7)

$$h_i = -\frac{1}{2}\nabla_i^2 - \sum_{k=1}^M \frac{Z_k}{r_{ik}} + V_i(j)$$

Equation 2.7 is the Hartree Hamiltonian operator, which includes an extra term to incorporate an interaction potential with all of the other electrons occupying orbitals (J):

(2.8)

$$V_i(j) = \sum_{j \neq i} \int \frac{\rho_j}{r_{ij}} d\mathbf{r}$$

Within this extra term, ρ_j is the charge (probability) density of electron j , which means that the repulsive third term on the right in Equation 2.7 is analogous to the attractive second term, except that the nuclei are regarded as point charges and the electrons as wave functions with diffuse charge, thereby making an integration over all space necessary. However, since the point of computation is to determine individual ψ , and since $\rho_j = |\psi|^2$, the problem arises of how these can possibly be used in the single-electron Hamiltonian before they have even been determined.

The solution comes from the implementation of an iterative ‘self-consistent field’ (SCF) method, the first step of which involves an initial guess of the ψ for all of the occupied molecular orbitals of a system and uses them to develop the required single-electron operators, h . Resolving each differential of the single-electron Schrödinger equation

(Equation 2.6) provides a new set of ψ , which are presumed to be both different to the initial guess and more accurate, consequentially new single-electron Hamiltonians are derived from them to determine each required ρ . This process is repeated to obtain an even better set of ψ each time, until the difference between the new set of wave functions and the ones from the preceding cycle falls below a certain threshold criterion, at which time the final set are taken to be ‘converged’.

In the *ab initio* Hartree-Fock method, the SCF approach is extended such that the interactions of an electron with the static field of all other electrons in a system now includes the exchange effects on the Coulomb repulsion, whilst ignoring all other electron correlation. The Density Functional Theory approach is coupled with the Hartree Fock method precisely because it does account for electron correlation, thus overcoming a major limitation of the Hartree Fock method.

The Fock operator for a single-electron system can be defined as:

(2.9)

$$f_i = -\frac{1}{2}\nabla_i^2 - \sum_{k=1}^M \frac{Z_k}{r_{ik}} + V_i^{HF}(j)$$

This Fock operator is essentially a modification of the Hartree Hamiltonian (Equation 2.7) and for this reason, it is sometimes referred to as the Hartree-Fock equation. The final term is the HF potential and equates to $2J_i - K_i$ where both J_i and K_i are operators from integrations to determine the Coulomb and inter-electronic repulsions respectively, the details of which are not outlined here.

It is computationally expensive to perform a single electronic structure calculation, and even more so to perform a molecular geometry optimisation, as the aim is to find stationary points on a PES as opposed to arbitrary structures. In order to locate such points efficiently, many algorithms use derivatives of the energy with respect to nuclear motion – and when those derivatives are available analytically, rather than numerically, the rates of convergence are significantly improved.

This is especially true when the stationary point to be located is a transition state structure. Unlike standard molecular mechanics, the HF energy has no apparent bias for minima compared to transition states – the key point is that molecular orbital theory provides an energy functional from which reasonable transition state structures can be identified, regardless of whether the methodology employed is computationally more expensive but more accurate (*ab initio*), or alternatively less expensive and less accurate (semi-empirical) in nature.

2.2-3 Density Functional Theory [40]

The Hamiltonian operator is dependant only on the atomic numbers and the positions of the nuclei and the total number of electrons. This latter dependence on the number of electrons suggests that a useful physical observable would be the electron density, ρ , which gives the total number of electrons, N , when integrated over all space:

(2.10)

$$N = \int \rho(\mathbf{r}) d\mathbf{r}$$

The nuclei can effectively be regarded as point charges, and thus their positions correspond to local maxima in the electron density, and these maxima are also cusps. Now all that remains to be specified in the Hamiltonian is the assignment of the nuclear atomic numbers, also readily available from the density since for each nucleus, A, located at an electron density maximum, r_A :

(2.11)

$$\left. \frac{\partial \bar{\rho}(r_A)}{\partial r_A} \right|_{r_A=0} = -2Z_A \rho(r_A)$$

The terms of Equation 2.11 are defined thus, Z is A's atomic number, r_A is the distance from nucleus A and $\bar{\rho}$ is the spherically averaged density.

All that has been presented so far in this section does not provide a simple formalism for finding the energy, it just simply indicates that for a known density, the Hamiltonian operator could be constructed, the Schrödinger equation solved, and the wave functions and energy eigenvalues determined. Nevertheless, this does suggest some simplifications may be possible.

In the context of density functional theory (DFT), electrons interact with one another and with an 'external potential'. In a uniform electron gas, this external potential is the evenly distributed positive charge, and in a molecule, the external potential is the attraction of the nuclei, expressed classically as:

(2.12)

$$V_{ne}[\rho(r)] = \sum_k^{nuclei} \int \frac{Z_k}{|r - r_k|} \rho(r) dr$$

In order to establish a dependence of the energy on the density in the Hohenberg-Kohn theorem, the ground-state density is utilised, and it is sufficient to show that this density determines the Hamiltonian operator. Equation 2.10 shows that the integration of the electron density gives the total number of electrons, thus all that is needed to define the operator is a determination of the external potential, i.e. the charges and positions of the nuclei. Proof that the ground-state density determines the external potential comes when it is shown that an assumption to the contrary leads to an impossible result.

The Hohenberg-Kohn Existence Theorem

For example, if we assume that two *distinct* external potentials, v_a and v_b (each present in its own Hamiltonian, H_a and H_b respectively) can be consistent with the same *non-degenerate* ground-state density, ρ_0 , then each Hamiltonian will also be associated with a ground-state wave function, ψ_0 and its related eigenvalue, E_0 . The variational theorem of MO theory dictates that the expected value of H_a over $\psi_{0,b}$ must be higher than $E_{0,a}$:

(2.13)

$$E_{0,a} < \langle \psi_{0,b} | H_a | \psi_{0,b} \rangle$$

Equation 2.13 can be rewritten to give Equation 2.14:

(2.14)

$$\begin{aligned}
E_{0,a} &< \langle \psi_{0,b} | H_a - H_b + H_b | \psi_{0,b} \rangle \\
&< \langle \psi_{0,b} | H_a - H_b | \psi_{0,b} \rangle + \langle \psi_{0,b} | H_b | \psi_{0,b} \rangle \\
&< \langle \psi_{0,b} | v_a - v_b | \psi_{0,b} \rangle + E_{0,b}
\end{aligned}$$

As the external potentials v are one-electron operators, the integral in the last line of Equation 2.14 can be expressed in terms of the ground-state density:

(2.15)

$$E_{0,a} < \int [v_a(r) - v_b(r)] \rho_0(r) dr + E_{0,b}$$

Since there is no difference being made between the a and b terms above, the indices in Equation 2.15 are interchangeable:

(2.16)

$$E_{0,b} < \int [v_b(r) - v_a(r)] \rho_0(r) dr + E_{0,a}$$

The addition of inequalities (Equations 2.15 and 2.16) gives:

(2.17)

$$\begin{aligned}
 E_{0,a} + E_{0,b} &< \int [v_b(r) - v_a(r)]\rho_0(r)dr + \int [v_a(r) - v_b(r)]\rho_0(r)dr + E_{0,b} + E_{0,a} \\
 &< \int [v_b(r) - v_a(r) + v_a(r) - v_b(r)]\rho_0(r)dr + E_{0,b} + E_{0,a} \\
 &< E_{0,b} + E_{0,a}
 \end{aligned}$$

where it is assumed that the ground-state densities associated with wave functions a and b are the same and this allows the integrals to be eliminated as they must sum up to zero.

However, this also leaves an impossible result whereby the sum of the two energies is less than itself, thus indicating that the initial assumption is incorrect. So, the non-degenerate ground-state density determines the external potential, and thus the Hamiltonian, and therefore the wave function. It is important to note here that the Hamiltonian determines not only the ground-state wave function, but also all of the excited-state wave functions as well, so there is an extensive amount of information that can be garnered from the density.

The Hohenberg-Kohn Variational Theorem

The above discussion briefly outlines the Hohenberg-Kohn existence theorem, as such in spite of its potential, this theorem is not very useful in determining how to make a prediction of the density of a system. As in MO theory, a means to optimise the fundamental quantity. A second theorem - the Hohenberg-Kohn variational theorem - demonstrates that the density obeys a variational principle, again just as in MO theory (see Equation 2.13).

If an ideal, well-behaved candidate density that integrates to the proper number of electrons, N , is assumed, then it can be shown that this density determines a candidate wave function and a candidate Hamiltonian, therefore the expected energy value can be determined:

(2.18)

$$\langle \psi_{cand} | H_{cand} | \psi_{cand} \rangle = E_{cand} \geq E_0$$

Equation 2.18 shows that the energy expectation value of the ideal candidate density, E_{cand} , must be greater than or equal to the ground-state energy, E_0 .

Essentially, different densities can continue to be chosen and those which result in lower energies, as determined by Equation 2.18, are closer to correct. This can prove to be rather an unsatisfying procedure because first there is no method for how to rationally select improved candidate densities, and second because the initial motivation for DFT was to avoid solving the Schrödinger equation, determining the energy as an expectation value is no improvement, as it is already known how to do that.

The Kohn-Sham Self-consistent Field (SCF) Methodology

The challenge lies in the nature of the functional itself. So far, it has been indicated that there are mappings from the density onto the Hamiltonian and the wave function, and thus the energy, but there has been no suggestion of a mechanical means by which the density can be utilised as an argument in a general, characteristic variational equation, to directly determine the energy without resorting to the wave function. The desire to determine the energy directly lead to the development of the Kohn-Sham Self-Consistent Field method in the mid 1960s.

It has already been heavily emphasised that density determines the external potential, which in turn determines the Hamiltonian, which then determines the wave function, and of course only then, with the determination of both the Hamiltonian and the wave function, can the energy be determined - however this approach provides no simplification over MO theory, as the last step still requires the Schrödinger equation to be solved, and this is difficult in most cases, and the difficulty arises from the electron-electron interaction term in the correct Hamiltonian. Therefore what is required is a Hamiltonian for a *non-interacting* system of electrons to simplify things, as such a Hamiltonian could be expressed as a sum of one-electron operators, would have eigenfunctions that are Slater determinants of the individual one-electron eigenfunctions, and would have eigenvalues that are the sum of the one-electron eigenvalues:

(2.19)

$$H^{(0)} = \sum_{i=1}^n f_i$$

where n is the number of basis functions and f_i is defined by Equation 2.9. Additionally, the HF wave function, which is a Slater determinant formed from the occupied orbitals, is defined by $\psi^{(0)}$. The eigenvalue of $H^{(0)}$ is the sum of the occupied orbital energies, when applied to the HF wave function:

(2.20)

$$H^{(0)}\psi^{(0)} = \sum_i^{occ.} \varepsilon_i \psi^{(0)}$$

where the orbital energies are the eigenvalues of the specific one-electron Fock operators, thus the sum of the right hand side of Equation 2.20 defines the eigenvalue.

The key to the Kohn-Sham SCF methodology is to take a starting point of a *theoretical* system of *non-interacting* electrons where the density of a *real* system of *interacting* electrons is identical to the overall ground-state density of the *theoretical non-interacting* system. At this point in the discussion, it is important to note that since the density determines the position of the nuclei, as per Equation 2.11, it is entirely necessary for these quantities to be the same in both the *non-interacting* and in the *real* systems. Having established a *theoretical* starting point of a *non-interacting* system of electrons, the energy functional must then be split into specific components for further analysis:

(2.21)

$$E[\rho(r)] = T_{ni}[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] + \Delta T[\rho(r)] + \Delta V_{ee}[\rho(r)]$$

The various terms on the right hand side of Equation 2.21 can be defined thus:

- T_{ni} is the kinetic energy of the *non-interacting* system of electrons, and is simply equal to the sum of all of the individual electronic kinetic energies.

- V_{ne} is the nuclear-electron attraction as defined by Equation 2.12.
- V_{ee} is the classical electron-electron repulsion term, where r_1 and r_2 are dummy integral variables running over all space. The term itself is defined by:

(2.22)

$$V_{ee}[\rho(r)] = \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2$$

- ΔT is the kinetic energy correction arising from the nature of the interactions of the electrons, and similarly ΔV_{ee} defines *all* of the non-classical corrections to the electron-electron repulsion energy.

An orbital expression for the density requires that Equation 2.21 be rewritten thus:

(2.23)

$$E[\rho(r)] = \sum_i^N \left(\left\langle \chi_i \left| \frac{1}{2} \nabla_i^2 \right| \chi_i \right\rangle - \left\langle \chi_i \left| \sum_k^{nuclei} \frac{Z_k}{|r_i - r_k|} \right| \chi_i \right\rangle \right) + \sum_i^N \left\langle \chi_i \left| \frac{1}{2} \int \frac{\rho(r')}{|r_i - r'|} \right| \chi_i \right\rangle + E_{xc}[\rho(r)]$$

Here N is the number of electrons and the density for a Slater determinantal wave function, which is an exact eigenfunction for the *non-interacting* system is:

(2.24)

$$\rho = \sum_{i=1}^N \langle \chi_i | \chi_i \rangle$$

In Equation 2.23, the ΔT and ΔV_{ee} terms of Equation 2.21 have been combined to give a new term, E_{xc} - the exchange correlation energy. This term incorporates the effects of quantum mechanical exchange and correlation as well as the correction for the classical electron-electron repulsion energy and the kinetic energy difference between the *theoretical non-interacting* system and a *real* one.

In the usual manner of finding the orbitals χ that minimise E in Equation 2.23, it can be seen that they satisfy:

$$h_i^{KS} \chi_i = \varepsilon_i \chi_i \quad (2.25)$$

The Kohn-Sham (KS) one-electron operator is given by:

$$h_i^{KS} = -\frac{1}{2} \nabla_i^2 - \sum_k^{nuclei} \frac{Z_k}{|r_i - r_k|} + \int \frac{\rho(r')}{|r_i - r'|} dr' + V_{xc} \quad (2.26)$$

V_{xc} is a functional derivative, best described as the one-electron operator for which the expected value of the KS Slater determinant equals the value of the exchange-correlation energy, E_{xc} , and is defined by:

$$V_{xc} = \frac{\delta E_{xc}}{\delta \rho} \quad (2.27)$$

The orbitals χ must give the *exact* density, meaning the minimum must correspond to reality, because the E of Equation 2.21 that is being minimised is exact. Furthermore it is these orbitals that construct the Slater determinantal eigenfunction for the separable non-interacting Hamiltonian, which is by definition the sum of the Kohn-Sham operators of Equation 2.26, therefore it can be seen that there is consistency within theorizing a *non-interacting* system of electrons with a density that is identical to that of a *real* system.

Within this context, it is justifiable to use the first term on the right hand side of Equation 2.23 to determine the kinetic energy of the *non-interacting* electrons, which is a large proportion of the kinetic energy of the actual system.

The determination of the KS orbitals is entirely analogous to a molecular orbital theory approach in that they are expressed within a basis set of functions, denoted ϕ , determining the individual orbital coefficients through solution of a secular equation, in the case of HF theory Equation 2.28, and in the case of DFT, equation 2.29.

(2.28)

$$F_{\mu\nu} = \left\langle \mu \left| -\frac{1}{2} \nabla^2 \right| \nu \right\rangle - \sum_k^{nuclei} Z_k \left\langle \mu \left| \frac{1}{r_k} \right| \nu \right\rangle + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - \frac{1}{2} (\mu\lambda|\nu\sigma) \right]$$

(2.29)

$$K_{\mu\nu} = \left\langle \phi_\mu \left| -\frac{1}{2} \nabla^2 - \sum_k^{nuclei} \frac{Z_k}{|r - r_k|} + \int \frac{\rho(r')}{|r - r'|} dr' + V_{xc} \right| \phi_\nu \right\rangle$$

While the mathematical details behind Equations 2.28 and 2.29 are not presented here, the important thing to understand is that $F_{\mu\nu}$ and $K_{\mu\nu}$ are general matrix elements of matrices \mathbf{F}

and \mathbf{K} respectively, similarly in Equation 2.28, $P_{\lambda\sigma}$ is a matrix element of a 'density matrix', \mathbf{P} . The convention followed here is that the basis functions are indexed by lower case Greek letters while the molecular orbitals are indexed by lower case Roman letters.

The similarities between HF theory and DFT are more extensive than just the mathematics of sharing a common variational principle. For example, the nuclear attraction and kinetic energy components of the matrix elements of \mathbf{K} are the same as those for matrix elements of \mathbf{F} . Moreover, if the density which appears in the classical electron-electron repulsion operator is expressed in basis functions which are identical to those used for KS orbitals, the result is that identical four-index electron-repulsion integrals appear in both \mathbf{K} and \mathbf{F} . The Kohn-Sham methodology must be employed as an iterative SCF procedure, as the density is required for computation of the secular matrix elements, but the density is determined using orbitals derived from the *solution* of the secular equation, in accordance with Equation 2.24.

The key difference between HF theory and DFT is that DFT contains no approximations, as described thus far - it is *exact*. All that remains to be determined is the exchange correlation energy, E_{xc} , as a function of the density, ρ . While Hohenberg and Kohn proved that a functional of the density has to *exist*, there is nothing to suggest its form. As a result, finding density functionals which may reasonably approximate E_{xc} has been a topic of considerable research.

The emphasis here is that, in the limit of an infinite basis set, HF is a deliberate *approximation*, whereas DFT is *exact*, with the caveat that the relevant equations can only be solved *approximately* due to a functional of the density having an unknown form.

It is also important to note that once approximations for the exchange correlation energy have been incorporated, exact DFT is no longer variational. The fundamental difference

between DFT and MO theory to remember is that whilst DFT optimises an electron density, MO theory optimises a wave function.

Exchange correlation functionals

The exchange correlation energy, E_{xc} , accounts for the difference between quantum mechanical and classical inter-electron repulsion, as well as the kinetic energy difference between the *theoretical* and the *real* system, at least in principle. However, most modern functionals do not explicitly compute this term, either ignoring it or attempting to construct a hole function, analogous to Equation 2.30 except that it includes the kinetic energy difference between the interacting and non-interacting systems.

(2.30)

$$\left\langle \psi \left| \sum_{i < j}^{electrons} \frac{1}{r_{ij}} \right| \psi \right\rangle = \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 + \frac{1}{2} \int \int \frac{\rho(r_1)h(r_1; r_2)}{|r_1 - r_2|} dr_1 dr_2$$

On the left hand side of Equation 2.30 is the exact quantum mechanical electron-electron repulsion, while the second term on the right corrects the errors in the classical expression (the first term on the right) using the hole function, designated h associated with the density. The notation $h(r_1; r_2)$ indicates that the hole is centred on the position of electron 1, and is evaluated as a function of the remaining spatial coordinates defining r_2 . It is important to note that as well as the value of the hole function varying as a function of r_2 for a defined value of r_1 , the form of h itself can vary as a function of r_1 .

In a one-electron system, the left hand side of Equation 2.30 must equal zero, however the first term on the right is not zero, since the density must be greater than or equal to zero throughout space, thus the hole function is simply the negative of the density. In a many-

electron system, the precise form of the hole function is not readily established. Hole functions in many-electron systems account for the exchange correlation energy as well as the self-interaction error.

Many functionals include empirical parameters which introduce some kinetic energy correction as a necessity, if they are experiment-based.

The functional dependence of E_{xc} on the electron density can be expressed as an interaction between the electron density and an 'energy density', ϵ_{xc} :

(2.31)

$$E_{xc}[\rho(r)] = \int \rho(r) \epsilon_{xc}[\rho(r)] dr$$

The energy density is the sum of individual exchange and correlation functionals. An important distinction between the two types of density involved here is that electron density is a per unit volume density and the energy density is a per particle density.

Conventionally the electron density is expressed in terms of an effective radius, whereby exactly one electron would be present within the sphere defined by that radius if it had the same density throughout its centre:

(2.32)

$$r_s = \left(\frac{3}{4\pi\rho(r)} \right)^{1/3}$$

Electron spin can also be handled within DFT, using individual functions of the α and β densities. At any position the spin densities are expressed in terms of the normalised spin polarisation, ζ :

$$\zeta(r) = \frac{\rho^\alpha(r) - \rho^\beta(r)}{\rho(r)}$$

Spin density α is simply one-half the product of the total ρ and $(\zeta + 1)$, whilst spin density β is the difference between that value and the total ρ .

The local density approximation (LDA) was originally used as an indication of any DFT approach whereby the value of the energy density, ε_{xc} , at some position, r , could be determined exclusively from the electron density at that position, essentially the only requirement then is that the electron density be single-valued at every position. This scenario only holds true in the case of the uniform electron gas, resulting in LDA often being used to imply that it is those exchange and correlation functions that are being used. Therefore it is reasonable to believe that in a molecular system, where the electron density is far from uniform, LDA has limitations.

The correlation functional could be improved by having it be not only dependent on the local value of the density, but also the gradient of the density, i.e. - the extent to which it is locally changing. Such improved correlation functionals, like the commonly used LYP or, as used in this work, PW91, which depend on the density and its gradient are 'gradient corrected, and the inclusion of such a correction is the definition of a generalised gradient approximation, or GGA. Most GGA functionals add the correction term to the LDA functional.

2.3 Computational Methodology

All of this work was done following an *ab initio* methodology [41-52], that is to say the calculations that were run, fully implemented the HF and as such computed all the terms in the Hamiltonian, as opposed to a semi-empirical method which often ignores some of the terms or estimates them. This approach is more computationally expensive and time consuming, but the reward is greater accuracy compared to semi-empirical methods.

2.3-1 The Hybrid DFT/HF Approach

The basis for all the calculations that were performed during this work is:

MPW1PW91/6-31g**

Breaking this down into its components, MPW1PW91 defines the density functional, in this case it is Barone's modified Perdew-Wang GGA correlation functional [53, 54] whilst 6-31g** is the basis set.

A basis set [40] is simply a set of mathematical functions from which ψ can be derived. The 6-31g [55] set implemented here is among the most widely used split-valence basis sets available. The first number, i.e. 6 corresponds to the number of primitives used in the contracted core functions, the numbers after the hyphen, i.e. 3 and 1 correspond to the number of primitives used in the valence functions – here the two numbers indicate a

double- ζ basis [40, 55]. The split-valence basis sets use sp basis functions having common exponents just as the simpler basis sets do. The * notation denotes the polarisation function – a set of d functions added to polarise the p functions of this basis set, and the second * denotes a set of p functions on H and He [55]. Thus this basis set can be expressed as:

6-31g (d,p)

This alternative notation is readily observed in the output from the calculations that were performed in this work. It is also important to note that unless specified, Gaussian implements Restricted Hartree-Fock (RHF) which is sufficient for closed shell systems with no unpaired electrons, but for open shell systems that do have unpaired electrons, the **#UHF** keyword precedes the HF/DFT basis set and requests Unrestricted Hartree-Fock for calculations.

2.3-2 The Transition State Search

All of calculations in this study were performed in Gaussian 03 [56] on BlueBEAR [37]. The search for the transition states detailed in the Results and Discussion section of this report was intensive both in terms of labour and computation because the method adopted, required all of the initial guesses for all of the theoretically possible TS structures to be constructed from scratch within Insight II [10] and run directly in Gaussian 03 .

In order to perform the search in Gaussian 03, the following keywords were used [57]:

OPT – performs a geometry optimisation by adjusting the geometry until a stationary point on the PES is located using gradients and the default Berny optimisation algorithm for optimising both TS and local minima. This algorithm uses redundant internal coordinates [58, 59].

Associated keywords within this option:

TS – optimises a (first order) transition state rather than a local minimum.

CalcFC – calculates force constants at the first point

CalcAll – an alternative used when the number of steps for optimisation is exceeded, it is potentially more time consuming and more computationally expensive but has a greater efficiency and accuracy because force constants are calculated at every point.

Noeigentest – suppresses testing the curvature in Berny optimisations, useful because the initial TS structures are only guesses and testing the curvature halts the optimisation at each saddle point if it is not a TS.

Saddle = N - optimises at a saddle point of order N, useful for optimising second order TS where N=2, or higher (for a first order TS, saddle =1; for a local minimum, saddle =0)

Freq – calculates vibrational frequencies (only useful at stationary points on the PES) by determination of second derivatives of the energy with respect to the nuclear Cartesian coordinates, then changing these into mass-weighted coordinates. The interest lies in the value of the single imaginary frequency – the transition vector (TV) for an optimised TS,

which is typically large and negative. For a successfully optimised stationary point, six values (five for linear molecule) relatively close to zero are also observed as illustrated below:

```
Full mass-weighted force constant matrix:
Low frequencies --- -767.3632    0.0002    0.0006    0.0009    6.2067
7.0699
Low frequencies ---    12.3892    35.3178    47.4322
*****    1 imaginary frequencies (negative Signs) *****
```

For clarity the six values have been highlighted in bold, and the TV frequency has also been italicised.

Geom=checkpoint - uses coordinates from the checkpoint file generated by the calculation to define the geometry. The checkpoint file typically records the final geometry before an optimisation either completes or terminates for some reason, so it is useful as it can restart the optimisation from the last point.

2.3-3 Intrinsic Reaction Coordinate Calculations

Intrinsic Reaction Coordinate (IRC) calculations were performed on Gaussian 03 [56] to verify that the optimised TS structures were true transition states for the hydride transfer reaction. The **IRC** keyword ensures that a reaction path is followed [59, 60]. The starting geometry must be a transition state, from which point the path in one, or both directions

can be followed. The forward direction is the one in which the transition vector (TV) is pointing when the largest component of the phase is positive [57]. The geometry is optimised at each point along the reaction path so that the segment between any two adjacent points along the path is described by a circular arc. By default, the path is followed at six points in each direction, in steps of $0.1 \text{ amu}^{1/2} \text{ bohr}$ along the path [57]. Initial force constants are required for this type of calculation and for this, the **CalcFC** keyword was used (or **CalcAll** if there were optimisation problems)

An example of the output achieved from this type of calculation is presented below:

Summary of reaction path following:
(Int. Coord: Angstroms, and Degrees)

	ENERGY	RX.COORD	x1	x2	x3
1	-913.97035	-0.59093	4.82046	6.02418	7.21961
2	-913.96967	-0.49122	4.82073	6.02302	7.22205
3	-913.96898	-0.39142	4.82099	6.02163	7.22469
4	-913.96830	-0.29185	4.82126	6.02016	7.22722
5	-913.96772	-0.19293	4.82162	6.01850	7.22970
6	-913.96733	-0.09705	4.82129	6.01736	7.23226
7	-913.96709	0.00000	4.82100	6.01600	7.23500
8	-913.96727	0.09560	4.82186	6.01401	7.23692
9	-913.96757	0.18958	4.82257	6.01238	7.23834
10	-913.96804	0.28855	4.82247	6.01089	7.24045
11	-913.96858	0.38807	4.82248	6.00914	7.24253
12	-913.96915	0.48791	4.82227	6.00738	7.24446
13	-913.96971	0.58768	4.82191	6.00567	7.24625

The calculation by default steps six points in either direction from the optimised TS and the TS can easily be identified from this output as point **7** along the path as highlighted in bold, primarily because the reaction coordinates of this point are **0.0000**, for further confirmation the energy reported here (to five decimal places) can be compared to the final energy from the TS optimisation to see if they coincide as they should.

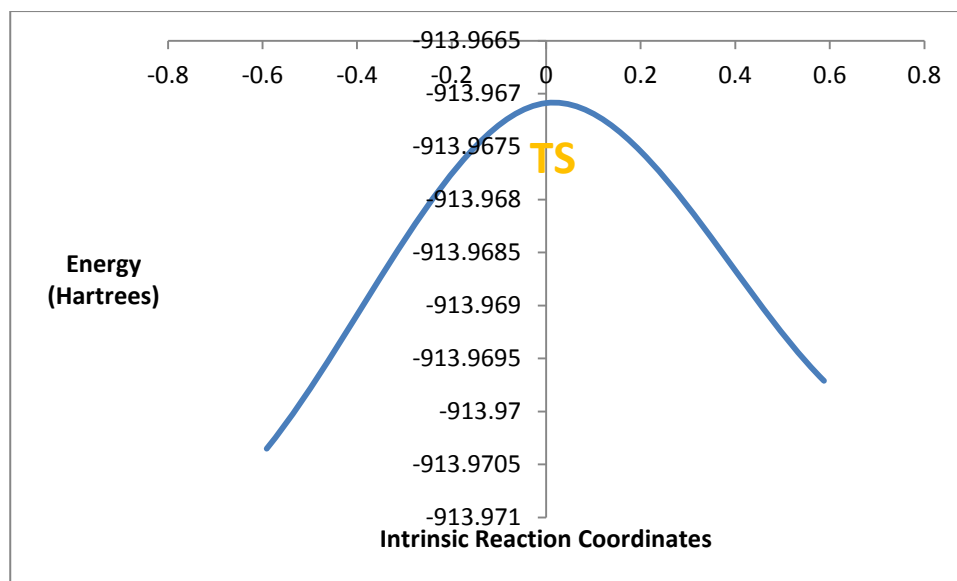


Figure 2.1 –An illustrative sketch of IRC results that can be plotted to show the TS at the top of an energy hill and the reaction path in either direction.

2.4 Hybrid QM/MM (ONIOM) Methodology [61]

The "our own n-layered integrated molecular orbital molecular mechanics" (ONIOM) [62-68] methodology is a hybrid QM/MM approach which allows different levels of theory to be applied to different parts (layers) of a molecular system. For a two-layer ONIOM, the small (model) system is calculated at both the high and low levels of theory while the large (real) system is calculated only at the low level of theory. The result for the real system at the level of high theory is estimated by addition of the difference between the high and low theoretical levels for the model system to the low level results for the real system thus:

(2.34)

$$E_{ONIOM}(\textit{real}, \textit{high}) = E_{high}(\textit{model}) - E_{low}(\textit{model}) + E_{low}(\textit{real})$$

2.5 Force Fields

The Universal Force Field (UFF) [69] was implemented within the PCM methodology used to investigate transition state structures within a solvated environment (see Section 2.6). However, the ONIOM study on model dimers and tetramers that is the focus of Chapter 6 implemented the AMBER (Assisted Model Building with Energy Refinement) force field [70]. The respective force fields are not discussed in any great detail here, except to present the generalised form of each in Equations 2.35 and 2.36 respectively.

(2.35)

$$V(r^N) = E_R + E_\theta + E_\phi + E_\omega + E_{vdw} + E_{el}$$

Equation 2.35 shows the general form of the UFF force field. Here the potential energy of a system, $V(r^N)$, is expressed as the sum of all bonded and non-bonded interactions. In this expression, the bonded interactions are defined as bond stretches (E_R), dihedral torsion angles (E_θ), angle bends (E_ϕ) and inversions (E_ω), whilst the non-bonded terms are van der Waals (E_{vdw}) and electrostatic interactions (E_{el}) [69].

A similar equation represents the general form of the AMBER force field:

(2.36)

$$\begin{aligned}
V(r^N) = & \sum_{bonds} k_b(l - l_0)^2 + \sum_{angles} k_a(\theta - \theta_0)^2 \\
& + \sum_{torsions} \frac{1}{2} V_n[1 + \cos(n\omega - \gamma)] \\
& + \sum_{j=1}^{N-1} \sum_{i=j+1}^N \left\{ \epsilon_{i,j} \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} \\
& + \sum_{improper} k_{imp} (\phi - \phi_0)^2
\end{aligned}$$

Equation 2.36 shows the general form of the AMBER force field and thus defines the potential energy of the system in much the same way as for the UFF force field. The first term on the right summing over bonds is representative of covalent bond energy. This harmonic force is a good approximation near equilibrium bond length but fails upon increasing bond length and atomic separation. The term which sums over the angles represents the energy of the electronic orbital geometry involved in covalent bonding. The third term on the right hand side, a summation over torsions is representative of the energy of bond rotating due to bond order (double, triple etc) and neighbouring bonds of electronic lone pairs. It is important to note here that a single bond can have more than one of these torsion terms such that a total torsion energy can be expressed as a Fourier series. The double summation term over i and j defines the non-bonding energy between all atomic pairs and can be separated into van der Waals dispersion energy (first summation) and electrostatic energy (second summation) [70].

The van der Waals energy is computed using equilibrium distance, r_{0ij} and well depth, ϵ . The factor of two ensures that the equilibrium distance equates to r_{0ij} and when soft-core

potentials are utilised, the van der Waals energy is expressed in terms of σ where the equilibrium distance equates to $2^{1/6}(\sigma)$ [70].

The electrostatic energy as defined here operates under the assumption that proton and electron charges in an atom can be expressed as a single point charge, or as a small number of point charges in the case of electron lone pairs [70].

Additionally the planarity of aromatic rings and other conjugate systems is enforced by the term which sums over all the improper torsions.

2.6 Polarizable Continuum Model Calculations

Polarizable Continuum Model (PCM) calculations within Gaussian 03 [56] allow calculations to be performed on transition state structures within a solvated environment using the continuum models approximations, of which PCM is the most popular methodology. This approach provides a good balance between computational time and accuracy.

Most of the keywords associated with this type of calculation are the same as outlined in Section 2.3-2, particularly those associated with the **OPT** keyword. Keywords which are specific to this type of calculation are defined below [57]:

SCRF - requests a calculation to be performed in the presence of a solvent.

Keywords associated with this option:

PCM - Polarizable Continuum Model calculations using models where the cavity is created using overlapping spheres, devised by Tomasi *et al.*. The work implemented within Gaussian was done by Barone *et al.*. Utilises a reaction field calculation using the IEF-PCM model [71-77].

Solvent=Water - selects the solvent to be water, $\epsilon=78.39$.

Read - requests the inclusion of additionally specified parameters in the PCM, such as **radii=UFF** to define the set of radii used to define the spheres, in this case from the UFF force field [69], where hydrogen atoms have individual spheres and are therefore treated explicitly.

Additional to the SCRF keyword, the **SCF=Tight** keyword requests normal, tight convergence criteria of the SCF.

An example of the type of output achieved from a PCM calculation is shown below:


```

SCF Done: E(RmPW+HF-PW91) = -914.070313134 A.U. after 16 cycles
      Conv = 0.7517D-08 -V/T = 2.0091
      S**2 = 0.0000

-----
Variational PCM results
=====
<psi(f)| H |psi(f)> (a.u.) = -913.988991
<psi(f)|H+V(f)/2|psi(f)> (a.u.) = -914.070313
Total free energy in solution:
with all non electrostatic terms (a.u.) = -914.040492
-----
(Polarized solute)-Solvent (kcal/mol) = -51.03
-----
Cavitation energy (kcal/mol) = 43.28
Dispersion energy (kcal/mol) = -26.19
Repulsion energy (kcal/mol) = 1.62
Total non electrostatic (kcal/mol) = 18.71
-----

```

The total energy in solution is the sum of the SCF energy and all the non electrostatic terms as highlighted. These results also include the dipole moment in both the gas phase and in solution (not shown) as well as the various components of the predicted SCRF energy and the ΔG of solvation.

CHAPTER

THREE

COMPUTATIONAL MODELLING

AND SIMULATION

This chapter provides an overview of the computational modelling techniques implemented to build nicotinamide hydride transfer transition states and also the use of molecular dynamics and ONIOM simulation on models of *R. rubrum* transhydrogenase

3. Computational Modelling

This chapter covers the nomenclature and computational modelling techniques implemented to build models of the nicotinamide hydride transfer transition states and also modelling of *R. rubrum* transhydrogenase as well as the use of molecular dynamics and ONIOM simulations to investigate the hydride transfer reaction.

3.1 Modelling Transition States For Hydride Transfer

This section outlines the specifics of the hydride transfer transition states modelled for the hydride transfer reaction.

3.1-1 Generating N-methyl Pyridine Transition States

For an energetic comparison to the nicotinamide TS structures, ‘core’ transition states were also generated, i.e. C_2 , C_{2v} , and C_{2h} TS structures transferring hydride between N-methyl pyridine and N-methyl dihydropyridine, as illustrated below:

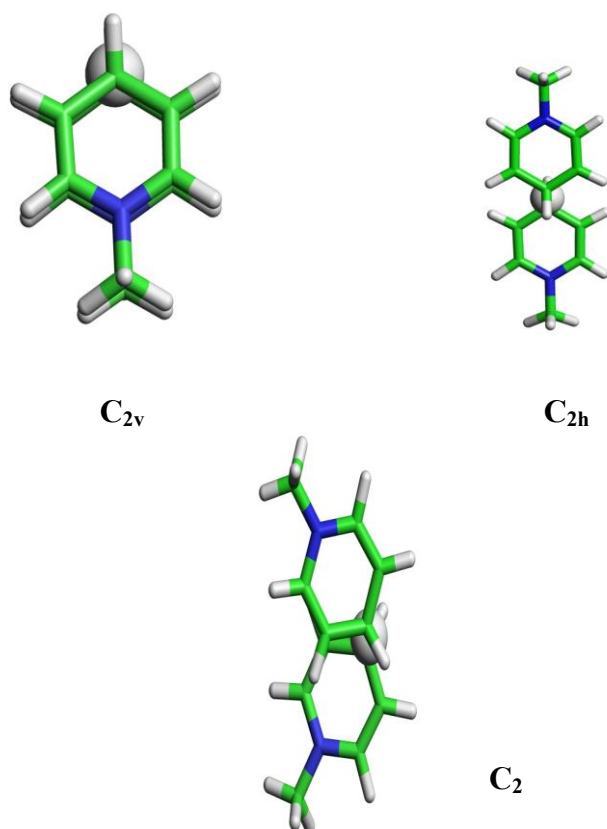


Fig. 3.1 N-methyl Pyridine 'core' structure transition states

As these core TS structures have true two-fold symmetry, they were optimised by ensuring that the transferring hydride C-H bond distances remained identical (initially set to 1.40 Å in the Z-matrix) so as to allow a simple energy optimisation to a first order TS using **OPT=Z-matrix** in Gaussian 03. Subsequently this constraint was removed and the TS optimised in the more conventional manner discussed in section 2.3-2, however this did not modify the optimised structure.

3.1-2 Generating Nicotinamide Hydride Transfer Transition States

Since the methodology employed here to optimise the TS geometries relies on a reasonable estimate as the initial geometry, and because the CPU usage time on BlueBEAR is restricted to 240 hours per calculation, it is important that the starting geometry of each TS is reasonably close to the final optimised TS.

In order to generate all the possible theoretical nicotinamide hydride transfer TS structures to be investigated, the amide position and orientation and the nicotinamide ring geometry were varied. All of the possible variants are illustrated below:

Possible Amide Orientations

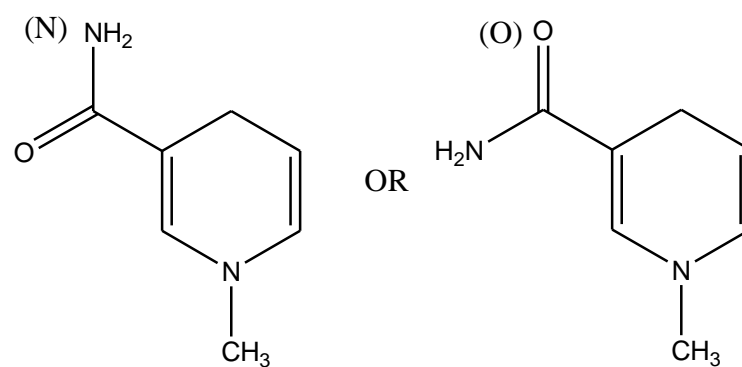


Fig. 3.2 Possible amide orientations: (N) or (O)

Possible Amide Positions

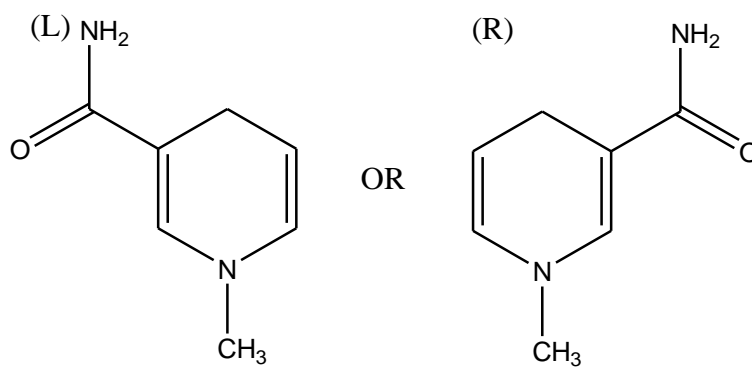


Fig 3.3 Possible amide positions: left (L) or (R) right

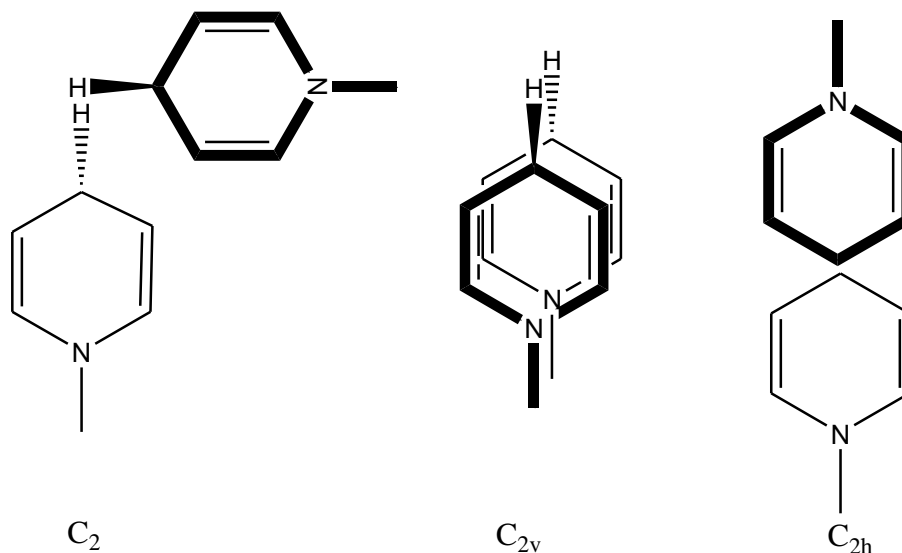


Fig 3.4 Possible ring geometries: C₂, C_{2v} and C_{2h}. These illustrations have been simplified for clarity, such that the amide position is not shown and in the C_{2v} case, the ring overlap has been slightly offset.

In the C_{2h} case the sp² character of the adjacent ring H atoms has not been identified. The translocating proton is also not shown, but the top ring in each orientation is in bold.

When all these variations are considered together, 64 theoretical TS for hydride transfer between the two nicotinamide rings exist, because for each of the three possible ring geometries above, the amide orientations (Figure 3.2) can be **NN**, **NO**, **ON** or **OO**, similarly the amide positions (Figure 3.3) can be **LL**, **LR**, **RL** or **RR**. This means 16 possibilities with C_{2h} geometry, 16 with C_{2v} and 32 with C₂ (as the rotation can be clockwise or anticlockwise). Each C₂ structure has an exact mirror image in the set so it is only necessary to compute 48 transition state structures, since mirror images would have the same energy to the original TS structures. In fact, the true number is a little lower than

48 as some of these structures would form symmetry related pairs. It can be difficult to reliably identify which structures are symmetry related from nomenclature alone. It is important also to emphasise here that the symmetry designations (C_2 , C_{2h} and C_{2v}) refer exclusively to the underlying symmetry of the pyridine rings and **not** the symmetry of the whole system.

3.2 Transition State Nomenclature

The naming conventions adopted to describe the nicotinamide hydride transfer TS structures are as follows:

3.2-1 Common descriptors

The common descriptors in all of the TS structure designations describe the ring geometry and the amide orientation respectively, and here each of them are outlined in turn.

Descriptors for the Ring Geometries

$C_2(x)$ describes the ring geometry. C_2 indicates solely two-fold symmetry, and (x) denotes the presence of a mirror plane, when (x) = **v**, this mirror plane is coincidental with the rotational axis, and when (x) = **h**, this mirror plane is perpendicular to the rotational axis. It is important to note that this symmetry descriptor does not strictly hold when the amides are considered as part of the TS structure, nonetheless it is a convenient descriptor for ring geometry, hence its implementation in this work.

Descriptors for the Amide Orientations

The descriptors **N** and **O** designate the amide orientation by which atom of the amide group is pointing *trans* in relation to the N1 nitrogen atom of the nicotinamide ring, see Figure 3.1.

3.2-2 The C₂ Transition State Descriptors

The C₂ TS descriptors **E** and **X** designate the position of the amide, relative to the adjacent ring. Each amide group can be either **E** - *endo*' - whereby it overlaps onto the adjacent ring, or **X** -*exo* - meaning instead of overlapping the amide group protrudes outward, relative to the adjacent ring. The ordering of rings in the nomenclature is quite arbitrary but it is important that the order is consistent between describing the rotation of the amides and their positions. If the order is reversed the result is a symmetry related structure so the energy should remain unchanged. The C₂ designations also have an extra descriptor to describe the dihedral (torsion, θ) angle which can either be (+) or (-), depending on whether this angle is positive or negative (Figure 3.5).

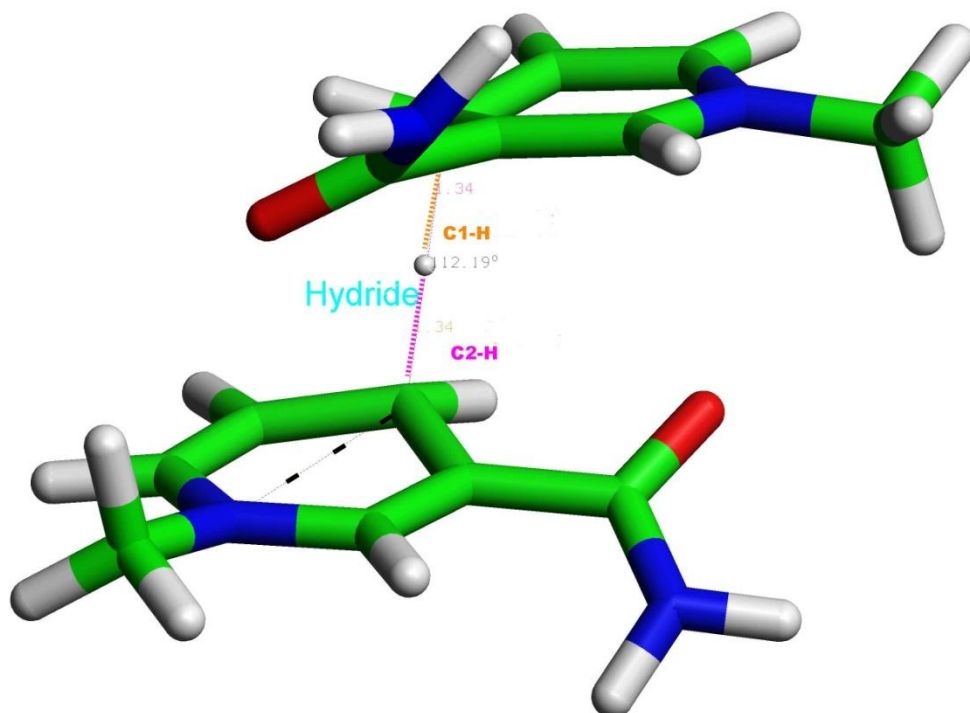


Fig. 3.5 A 3D representation of a nicotinamide hydride transfer TS, illustrating the bottom part of the N1-C4-C4-N1 θ torsion angle (black dashed line). This angle is positive ($+112.19^\circ$) thus the full structural designation for this TS would be C2-EE-OO (+). Also shown are the two C-H bond distances (orange C1-H and purple C2-H). both are 1.34 Å and in this case means that this TS is fully symmetric.

3.2-3 The C_{2h} and C_{2v} Transition State descriptors

The structural designations of both the C_{2h} and C_{2v} TS are simpler, similar, and ultimately more familiar in that the amide positions are prefixed (**E**) or (**Z**) based on whether both amides are on the same side (*zusammen*) or on opposite sides (*entgegen*) in relation to the adjacent ring - an isomer naming convention borrowed straight from organic chemistry.

Figure 3.6 below illustrates each of the concepts explained throughout the entirety of section 3.2 through an example of each of the three possible ring geometries, C_2 , C_{2v} and C_{2h} .

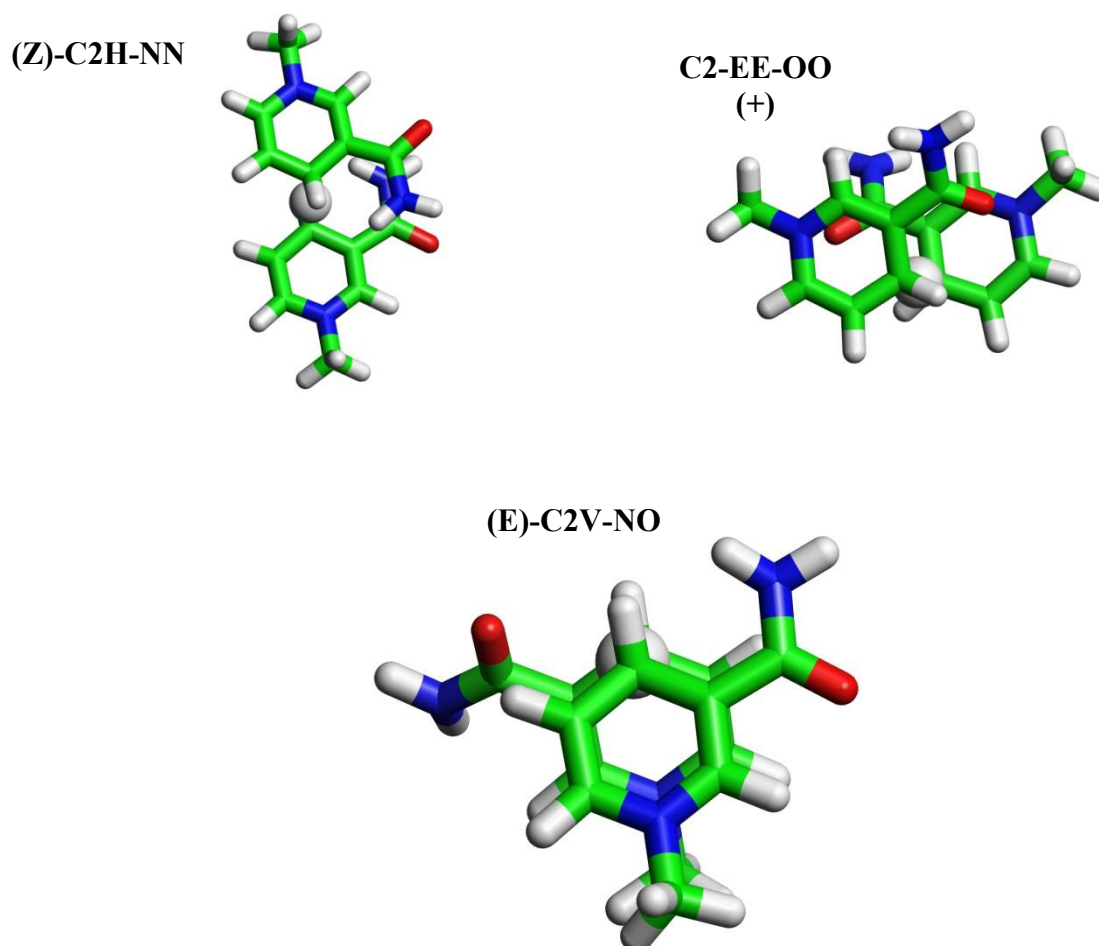


Fig. 3.6 Illustrating the full naming convention for a C_2 , C_{2h} and a C_{2v} TS structure, in each case the amide orientation and position is defined with respect to the top ring first.

3.3 Modelling *Rhodospirillum rubrum* Transhydrogenase

3.3-1 Preparing The Cofactors

Before the protein models could be built, the NAD(H) and NADP(H) cofactors within each of the models, had to be prepared separately in Antechamber within Amber so that it could be recognized and handled correctly during the simulation. Unlike the protein model, each of the cofactors requires explicit hydrogen atoms to begin with. These cofactors were prepared in separate ligand MOL2 files. Any bonding, valence or hybridization issues were also corrected within Insight II [10]. These ligand files were used to generate prep files with net charges for each cofactor being determined as summarised in Table 3.1. A semi-empirical model (bcc) was used to include the charges within Antechamber. In addition to prep files, frcmod parameter files were also generated. Both prep and frcmod files for each cofactor were then loaded into xleap for incorporation of the cofactors into each of the models.

Cofactor	Net Charge	Net Charge Derivation
NAD ⁺	-1	+1 for Nicotinamide ring, -1 for each phosphate linker
NADH	-2	-1 for each phosphate linker
NADP ⁺	-3	+1 for Nicotinamide ring, -1 for each phosphate linker, -2 for phosphate group
NADPH	-4	-1 for each phosphate linker, -2 for phosphate group

Table 3.1 Net Charges of the cofactors

3.3-2 Preparing The Protein Models

It is important to note here that in the crystal structure of transhydrogenase, the dI (B) domain contained 379 residues, the last of which was a leucine, and because the dI (A) domain only contained 378 residues, the 379th leucine residue was added to the dI (A) domain within Insight II [10] to make each binding site model domains dI (A) and dI (B) completely structurally equivalent. The structural PDB files in each case were prepared for use in Amber by editing them such that they contained no hydrogen atoms. Having only the heavy atoms specified when the PDB file is initially read into xleap avoids any issues with incompatible hydrogen atom names which can lead to any number of bonding issues being flagged up by the program. It is therefore a great deal simpler to strip all of the hydrogen atoms from the protein structure initially and let xleap add them automatically.

Within xleap, sodium (Na^+) ions were added to balance the negative charges of each model system and hence bestow an overall neutrality to the model before it was solvated in a rectangular TIP3P box, with 15 Å of water molecules around it .

All the protein models for this work were generated using Insight II [10] running on a UNIX-based SGI Workstation [38]. As already noted in Section 1.3, the intact transhydrogenase enzyme from *R.rubrum* is a dimer and consists of two dI+dII+dIII (monomers), but the crystal structure (PDB: 2OO5) is a dI₂dIII complex [8].

3.3-3 Modelling dIdIII Dimers

The dI₂dIII trimer (PDB: 2OO5) contains one complete active site (dimer) which is proximal in nature. In order to model a distal dimer, the tetramer needs to be built as described below. Table 3.2 below summarises the dimer configurations that were modelled in this work.

Binding Site Conformation	Binding Domains	Bound Cofactors
Distal	dI(A), dIII	NADH, NADP ⁺
Distal	dI (A), dIII	NAD ⁺ , NADPH
Proximal	dI (B), dIII	NADH, NADP ⁺
Proximal	dI (B), dIII	NAD ⁺ , NADPH

Table 3.2 The modelled single dIdIII dimer configurations

3.3-4 Modelling dI₂dIII₂ Tetramers

Tetrameric dI₂dIII₂ were modelled by constructing the missing dIII domain by overlaying one of the two dI domains from two equivalent dI₂dIII transhydrogenase crystal structures on top of each other, such that the dIII domain from the second crystal structure is positioned correctly into the first trimeric crystal structure (See Section 1.3). Tetrameric dI₂dIII₂ were modelled, by constructing the missing dIII domain by overlaying one of the two dI domains from two equivalent dI₂dIII transhydrogenase crystal structures on top of each other, such that the dIII domain from the second crystal structure is positioned correctly into the first trimeric crystal structure (See Section 1.3). The tetramer is constructed by overlaying the dI (A) of the second dI₂dIII over the dI (B) of the first dI₂dIII (see Chapter 1, Figure 1.2). Within this set up, the two dI (B) domains are deleted whilst the two dI (A) domains are retained alongside both dIII domains to generate the distal/distal tetramer. Similarly, one dI (A) and one dI (B) is deleted whilst the two dIII domains are retained to generate the distal/proximal tetramer. To generate a proximal/proximal tetramer, the two dI (A) domains are deleted, whilst the two dI (B) domains and two dIII domains are retained.

This extra NADP(H) binding domain resulted in many more conformational possibilities than the simple dIdIII dimers discussed above, primarily because the conformation of the second binding site, whether it is proximal or distal in nature is unknown. Therefore all possible configuration combinations have to be considered when investigating the reaction within these models. For the models where both sites are in the same conformation, i.e. both distal or both proximal, only one combination of sites where one site binds NADH (and NADP⁺) has to be modelled. In the case where both sites have different

conformations, one distal and one proximal, both combination of sites have to be considered. That is to say the distal site with NADH (and NADP⁺) bound and the proximal site with NAD⁺ (and NADPH) bound or vice versa.

For the preliminary ONIOM study (see Sections 3.5 and 3.6), each of the two active sites present within a tetrameric structure was treated separately, such that each tetramer yielded two 10 Å model snapshots per structural conformation. Table 3.3 below summarises these modelled conformations.

Binding Site Conformations	Bound Cofactor (dI, dI)	Bound Cofactor (dIII, dIII)
Distal/Distal	NADH, NADH	NADP⁺, NADP⁺
Distal/Distal	NADH, NAD⁺	NADP⁺, NADPH
Distal/Distal	NAD⁺, NAD⁺	NADPH, NADPH
Distal/Proximal	NADH, NADH	NADP⁺, NADP⁺
Distal/Proximal	NADH, NAD⁺	NADP⁺, NADPH
Distal/Proximal	NAD⁺, NADH	NADPH, NADP⁺
Distal/Proximal	NAD⁺, NAD⁺	NADPH, NADPH
Proximal/Proximal	NADH, NADH	NADP⁺, NADP⁺
Proximal/Proximal	NADH, NAD⁺	NADP⁺, NADPH
Proximal/Proximal	NAD⁺, NAD⁺	NADPH, NADPH

Table 3.3 - The modelled dI₂dIII₂ tetrameric configurations

3.3-5 The Amide Orientation of NAD(H)

In addition to each of the binding site conformation models listed in Tables 3.2 and 3.3, the amide orientation observed in the crystal structure of the enzyme is ambiguous [6] – arising from the fact that in X-ray crystal structures, protons do not appear and both nitrogen and oxygen are indistinguishable – thus both possibilities have to be considered in each of our protein models. The amide orientation of NAD(H) can be either **N** or **O** (see Figure 3.1) and as such we end up with double the number of models listed in each case, i.e. 8 dimers (4 with **N**, 4 with **O** orientations), and 20 tetramers (10 with **N** and 10 with **O**).

3.4 Molecular Dynamics Simulations

A dynamics simulation requires a set of initial velocities and coordinates and an energy function (interaction potential). This interaction can be considered constant for a short time step, allowing an updated set of velocities and coordinates to be estimated and thus the calculation of a new interaction. The behaviour of a system over time can be shown by taking a large number of small time step. A simulation typically explores only the region close to the starting point due to a short fundamental time step and a large phase space. As a result different simulations with different starting conditions can estimate the stability of the results [49].

Each of the protein models discussed in Section 3.3-3 were set up for MD simulations using Sander within Amber 8. The three stages of these MD simulations involve initial minimisation and equilibration at a temperature of 300 K for 500 ps followed by a set of

data gathering runs, each of 500 ps at 300 K for a total simulation time of around 3500 ps. For the ONIOM simulations, a snapshot was taken every 500 ps. Three MD simulations were run for each enzyme complex, whether it is a reactant or product state and which NAD(H) orientation is under consideration in each complex. Each 500 ps snapshot was taken and minimised to generate a starting structure for ONIOM calculations.

For the preliminary ONIOM study of the tetramers discussed in Section 3.3-4, only one MD simulation per each snapshot was run, with each snapshot being 250 ps owing to the larger system simulations running slower at each MD stage, so every second snapshot was taken to give comparable 500 ps snapshots.

3.4-1 Stages of an MD Simulation

Minimization

The input file for the minimization step was set up as follows:

```
title
&cntrl
imin=1, maxcyc=100000, NTB=1, cut=12
/
```

The key parameters here are `imin=1` which designates that this is a minimization step, not part of the MD simulation and `NTB=1` which designates that the minimisation was carried out at a constant volume.

An example of a typical output from a minimization is shown below:

NSTEP	ENERGY	RMS	GMAX	NAME	NUMBER
26750	-3.6244E+05	3.0709E-02	3.7601E+00	C16	8329
BOND	=	29680.6720	ANGLE	=	1110.5311
			DIHED	=	
5175.1245					
VDWAALS	=	79808.7794	EEL	=	-504675.2354
			HBOND	=	
0.0000					
1-4 VDW	=	1736.5810	1-4 EEL	=	24723.4514
			RESTRAINT	=	
0.0000					
.... RESTARTED DUE TO LINMIN FAILURE ...					
***** REPEATED LINMIN FAILURE *****					
FINAL RESULTS					
NSTEP	ENERGY	RMS	GMAX	NAME	NUMBER
26774	-3.6244E+05	2.5742E-02	1.8727E+00	C16	8329
BOND	=	29680.5033	ANGLE	=	1110.3926
			DIHED	=	
5175.2131					
VDWAALS	=	79809.3444	EEL	=	-504675.8801
			HBOND	=	
0.0000					
1-4 VDW	=	1736.4961	1-4 EEL	=	24723.5549
			RESTRAINT	=	
0.0000					

The 'linmin failure' is typical of a minimization, as it indicates that though the gradient is larger than the default minimization criterion, no structure of lower energy could be generated. However a final RMS value of E-02 and/or a GMAX value of E-01 constitutes a successful minimization.

Equilibration

The input file for the 500 ps equilibration step was set up as follows:

```
title NVT (canonical ensemble) dynamics - 500 ps equilibration run
&cntrl
imin=0, ntave=1000, ntwx=2500, ntwv=2500, ntwr=2500, ntf=2, NTB=1,
cut=12.0
dt=0.002, nstlim=250000, ntt=1, temp0=300.0, tempi=300.0, ntc=2
/
```

Since this is no longer a minimization, imin=0, but is still carried out at a constant volume, NTB=1. It can also be noted that the initial temperature, tempi and the final temperature, temp0 are the same, and ntt=1, collectively ensuring a run over a constant temperature of 300 K.

Data Gathering

The input file for the 500 ps data gathering dynamics runs was set up as follows:

```
title NVT (canonical ensemble) dynamics: data gathering 2ns
&cntrl
imin=0, ntave=1000, ntwx=5000, ntwv=5000, ntwr=5000, ntf=2, NTB=1,
cut=12.0,
dt=0.002, nstlim=250000, ntt=1, temp0=300.0, tempi=300.0, ntc=2,
ntx=5, irest=1
/
```

All the previously defined key parameters are unaltered from the equilibration step, $imin=0$, $NTB=1$, and $temp0$ and $tempi$ are both set to 300 K, and $ntt=1$ to ensure simulation runs over 500 ps at a constant volume and temperature on each run. Other important parameters unique to this file is ntx , which when set to 5 as in this case, and when $irest=1$ picks up the calculation from the restart file of the previous step to continue the run to generate successive 500 ps snapshots after each run.

Described above are the input files for the smaller dimeric systems, for the larger tetrameric systems, the equilibration and data gathering input files had the number of cycles halved - $nstlim=125000$ instead of $nstlim=250000$ in order to allow completion over a maximum run time of 240 hours per calculation, this meant that each run was essentially a 250 ps snapshot as opposed to 500 ps and therefore the snapshots of interest were every second one which respectively gave 500 ps snapshots.

In order to ensure a good stability of the final results, each conformational protein structure was run with three different simulations after initial minimisation. This was done by adding an extra parameter to the input files to use a different random 'seed' to spawn the

velocities. The default is IG=72177, this was changed to IG=70000 for the second runs and IG=75000 for the third runs.

Each stage of an MD simulation is limited to 240 hours and since we have many calculations at many stages for many structures to run, they are all run as serial jobs together utilizing the multi-processors of BlueBEAR [37].

3.5 Generating ONIOM Input Files: Development and Implementation of New Programs

As the protein dimers and tetramers were too large to be used directly to generate the input files within Antechamber for ONIOM calculations, a new program was developed and implemented to generate the required 10 Å spheres around the QM region as defined around in Figure 3.7. This results in each of the model systems typically being between 1300-1800 atoms in size. The structure files for each conformation were generated using Ambpdb within Amber after molecular dynamics. The partial charges on each atom were retained within the PDB files of these structures using the **-pqr** option within Ambpdb. so that the new program could use them for generating input files with and without these charges.

The program required as its input the QM region and the MM region as separate PDB files for each structural snapshot. Overall four variations of this program were developed using Fortran 90. First for rigid and uncharged MM layer structures, second for rigid and charged MM layer structures, third for 5 Å partially flexible and uncharged MM layer structures, and finally for 5 Å partially flexible and charged MM layer structures. In

conjunction with each of these programs, a sed script using a sed file that contained all the required Amber atom types within an Amber 03 force field [78, 79] for each amino acid was also used to create the input file.

Part of such an input file is given below:


```
# ONIOM(MPW1PW91/6-31g(d,p):Amber=HardFirst)=EmbedCharge opt
```

```
Transhydrogenase oniom, dimeric complex
```

```
0 1 1 1 1 1
```

```
O-O 0 47.169 56.400 59.213 H
C-C 0 47.617 55.689 60.087 H
N-N 0 47.360 54.393 60.008 H
H-H 0 47.667 53.750 60.716 H
H-H 0 46.754 54.077 59.267 H
C-CA 0 48.388 56.194 61.171 H
C-CA 0 48.677 55.445 62.245 H
H-H4 0 48.367 54.409 62.360 H
C-CA 0 48.831 57.462 61.146 H
H-HA 0 48.620 58.100 60.286 H
C-CA 0 49.524 57.970 62.171 H
H-HA 0 49.871 58.998 62.136 H
C-CA 0 49.763 57.224 63.246 H
H-H4 0 50.299 57.610 64.114 H
N-N* 0 49.346 55.942 63.309 H
C-CT 0 49.094 55.341 64.646 H
H-H2 0 49.779 55.773 65.389 H
O-OS 0 49.321 53.920 64.639 L H-H2 16
C-CT 0 47.627 55.444 65.123 L H-H2 16
```

This part of the file shows part of the QM layer with link atoms (see Figure 3.8). It is worth noting that the QM region does not require explicit partial charges on each atom, and as such the **EmbedCharge** keyword applies to treatment of the MM layer with regards

to charges. The six integers before the structural information define the charge/multiplicity information for the various parts of the system - see Section 2.4.

For clarity, part of the MM layer of this input file is shown below.

N-N--0.3009	0	47.824	65.120	68.953	L
H-H-0.2337	0	48.517	65.829	68.757	L
C-CT--0.1314	0	47.987	63.831	68.277	L
H-H1-0.0533	0	47.015	63.527	67.887	L
C-CT-0.0367	0	48.935	64.018	67.086	L
H-HC-0.0280	0	49.665	64.792	67.319	L
H-HC-0.0280	0	49.488	63.099	66.903	L
C-CT-0.0125	0	48.173	64.373	65.805	L
H-HC-0.0030	0	47.542	63.532	65.515	L
H-HC-0.0030	0	47.551	65.252	65.962	L
C-CT-0.1263	0	49.191	64.657	64.706	L
H-H1-0.0681	0	49.717	65.582	64.939	L
H-H1-0.0681	0	49.918	63.843	64.673	L

The general required input requires each atom specification to follow *atom-type-charge*. It is worth noting here that the Amber atom type for the link atoms treated as hydrogen atoms is also specified, i.e. H2. The digit which follows can either be 0 to denote a flexible atom, or -1 to denote a rigid (frozen) atom. The letter which follows the coordinate information donates the layer that the atom is in, **H** for High (QM region) or **L** for Low (MM layer). The **HardFirst** keyword employs the use of known angle bend and torsion Amber force field parameters between atoms of particular atom types. Some of these parameters have to be manually defined in the input file of each structure, particularly for the NAD(H) and NADP(H) cofactors. The Amber atom type specifications for the

cofactors and their associated Amber force field parameters used in this work were developed by Ryde [80-82] and are outlined below as the last part of the example input file as generated by the programs developed for this work:

HrmStr1	CA	N*	448.00	1.344
HrmStr1	*	IP	0.00	2.800
HrmBnd1	CA	C O	80.00	119.80
HrmBnd1	CA	C N	70.00	115.70
HrmBnd1	C	CA CT	63.00	120.00
HrmBnd1	CA	CA N*	70.00	119.50
HrmBnd1	CT	CA HA	35.00	120.00
HrmBnd1	H4	CA N*	35.00	114.00
HrmBnd1	CA	CT CA	63.00	120.00
HrmBnd1	CA	N* CA	70.00	121.20
HrmBnd1	CA	N* CT	70.00	120.60
HrmBnd1	*	* IP	0.00	120.00

It should also be noted that included within each 10 Å sphere are any sodium ions (amber type IP) and water molecules from the solvation of these models, and *HrmStr1* defines 'bond lengths' between atoms while *HrmBnd1* defines torsions.

3.6 Protein Modelling For ONIOM Study

In this work, the QM region (High Layer) is defined as the nicotinamide rings of both cofactors. Protein residues within 10 Å around this define the MM region (Low Layer). Additionally in these models, the QM layer is flexible, whilst the MM layer can be fully

rigid or partially flexible (protein residues within 5 Å of the QM layer) and either charged or uncharged. In regards to the hydride transfer reaction, this allows investigation of the influences of either the electrostatic effects of the surrounding protein environment, as in the charged cases, or the protein as a simple steric block, where geometry alone is the influencing factor, as in the uncharged cases. The ONIOM calculations for the fully rigid cases typically run faster to completion, taking anywhere between 10 hours to 2 days, whereas the flexible cases can take longer, typically between 5 days and 4 weeks for some of the snapshots.

Figure 3.7 below shows a representation of the ONIOM layers.

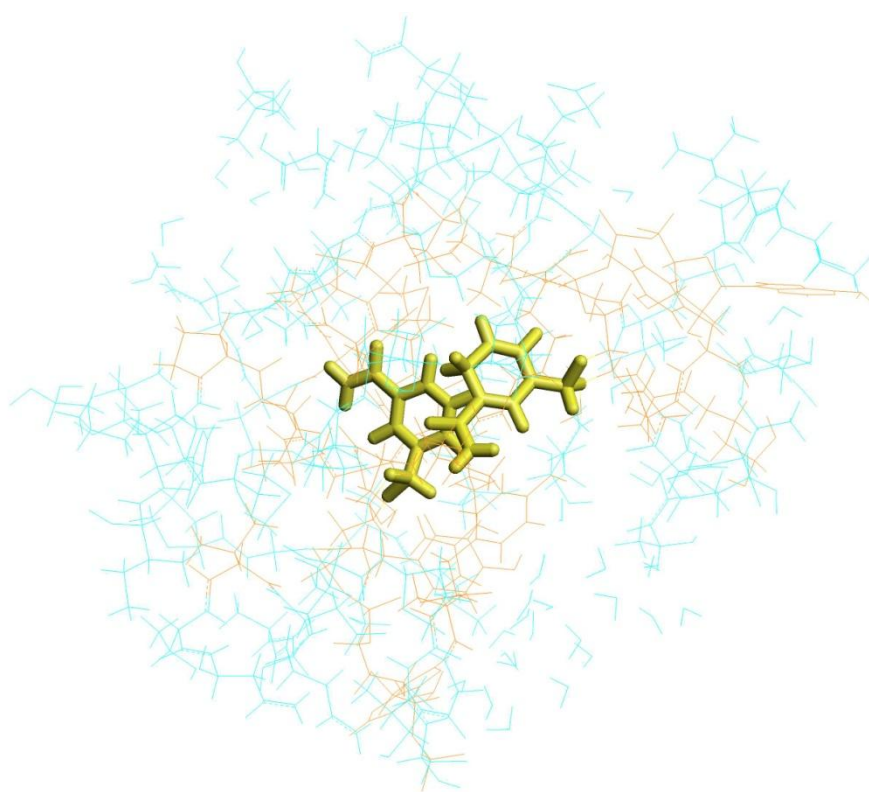


Fig. 3.7 A representation of ONIOM layers. The nicotinamide rings in the QM layer are shown in yellow, the flexible part of the MM (residues within 5 Å of the QM layer) is shown in orange, and the outer rigid part of the MM layer is shown in blue

The QM and MM layers are bonded through so called link atoms (designated * in Figure 3.8). As illustrated the link carbon and oxygen atoms in the MM (Low) layer are treated as hydrogen atoms in the QM (High) layer.

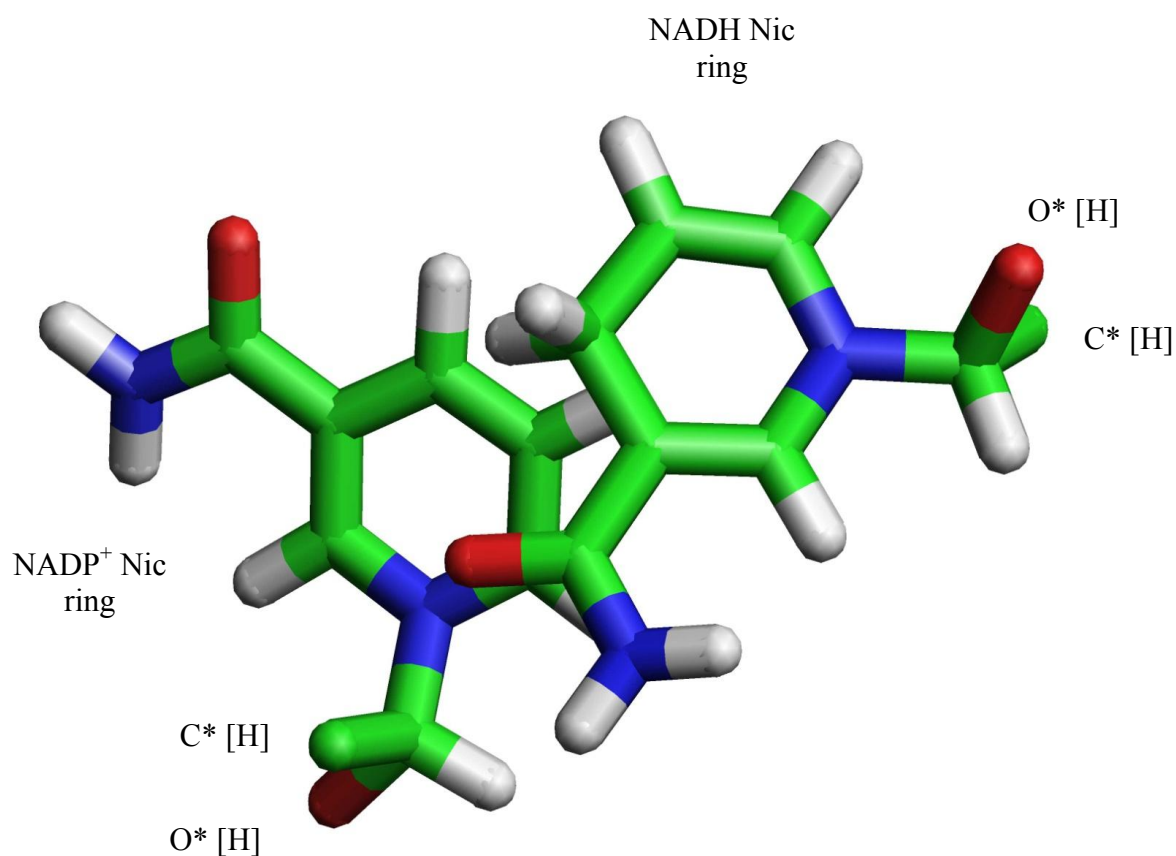


Fig. 3.8 - Illustrating the concept of link atoms which connect the QM and MM layers in ONIOM. The link atoms C* and O* in the Low (MM) layer are treated as hydrogen atoms [H] in the High (QM) layer

CHAPTER FOUR

PRELIMINARY WORK

This chapter provides a summary of the precursor work that underpins much of this research

4. Preliminary Work

A true hydride transfer mechanism that is distinct from electron/proton/electron is yet to be proven, so as a precursor to this work, energy optimisations were performed on the reactants and products of the sequential electron/proton/electron transfer mechanism. As well as this, the energetics of C₂, C_{2h} and C_{2v} N-methyl pyridine hydride transfer transition states were investigated.

4.1 The Sequential Electron/Proton/Electron Transfer Mechanism

Table 4.1 shows that the one- and two-electron transfer products modelled in the gas phase are relatively high energy intermediates, supporting the observations from previous studies in solution [33].

It is also important to note that an electron/proton/electron reaction order only utilises the first step as the intermediates both before and after proton transfer are mirror images, but an electron/electron/proton reaction requires both steps and the final product is a mirror image of the reactant.

It is important to consider the possibilities of the sequential electron and proton transfer mechanisms in the preliminary work for investigating the hydride transfer reaction in transhydrogenase as alternatives for direct hydride transfer because some quantum mechanical calculations of similar reactions in protein systems, such as for the reduction of CB1954 by wild-type NTR [83] show that an electron transfer mechanism is preferred

as opposed to net hydride transfer, although gas-phase reactions show that a hydride transfer mechanism is still feasible thermodynamically for this reduction reaction. It is important to note also that NTR uses flavin as the reducing agent which is known to facilitate one-electron reductions [83], compared to NAD(H) which only ever facilitates two-electron reductions.

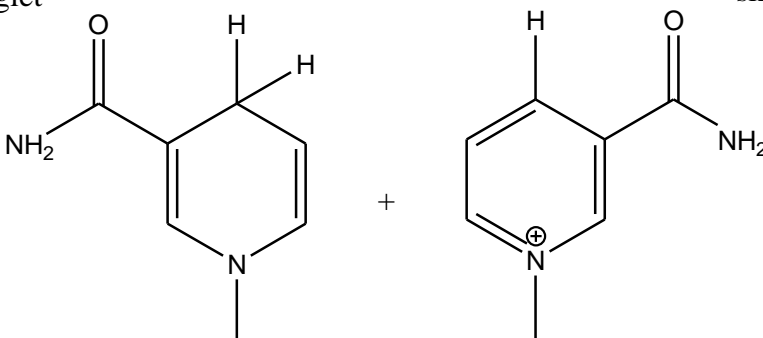
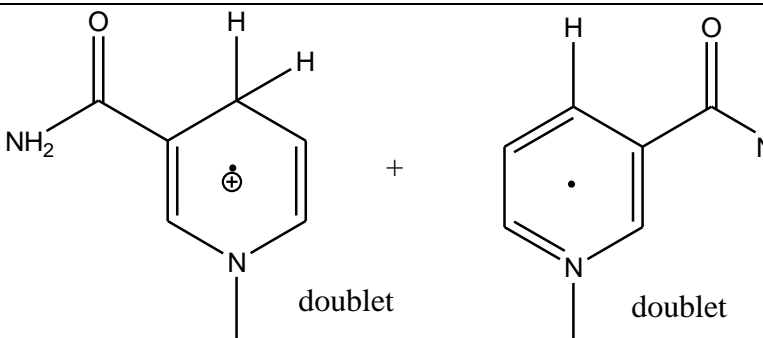
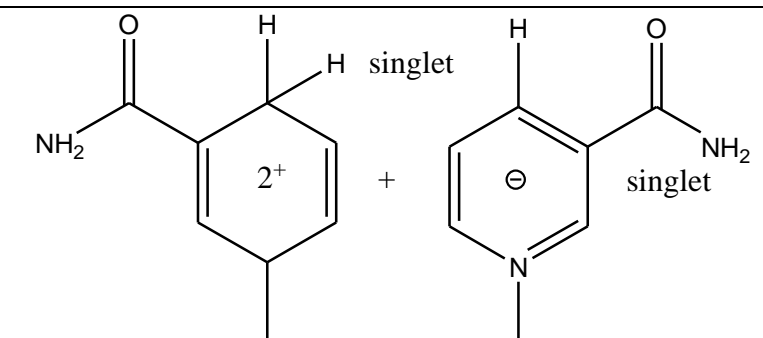
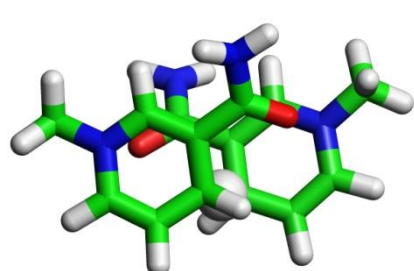
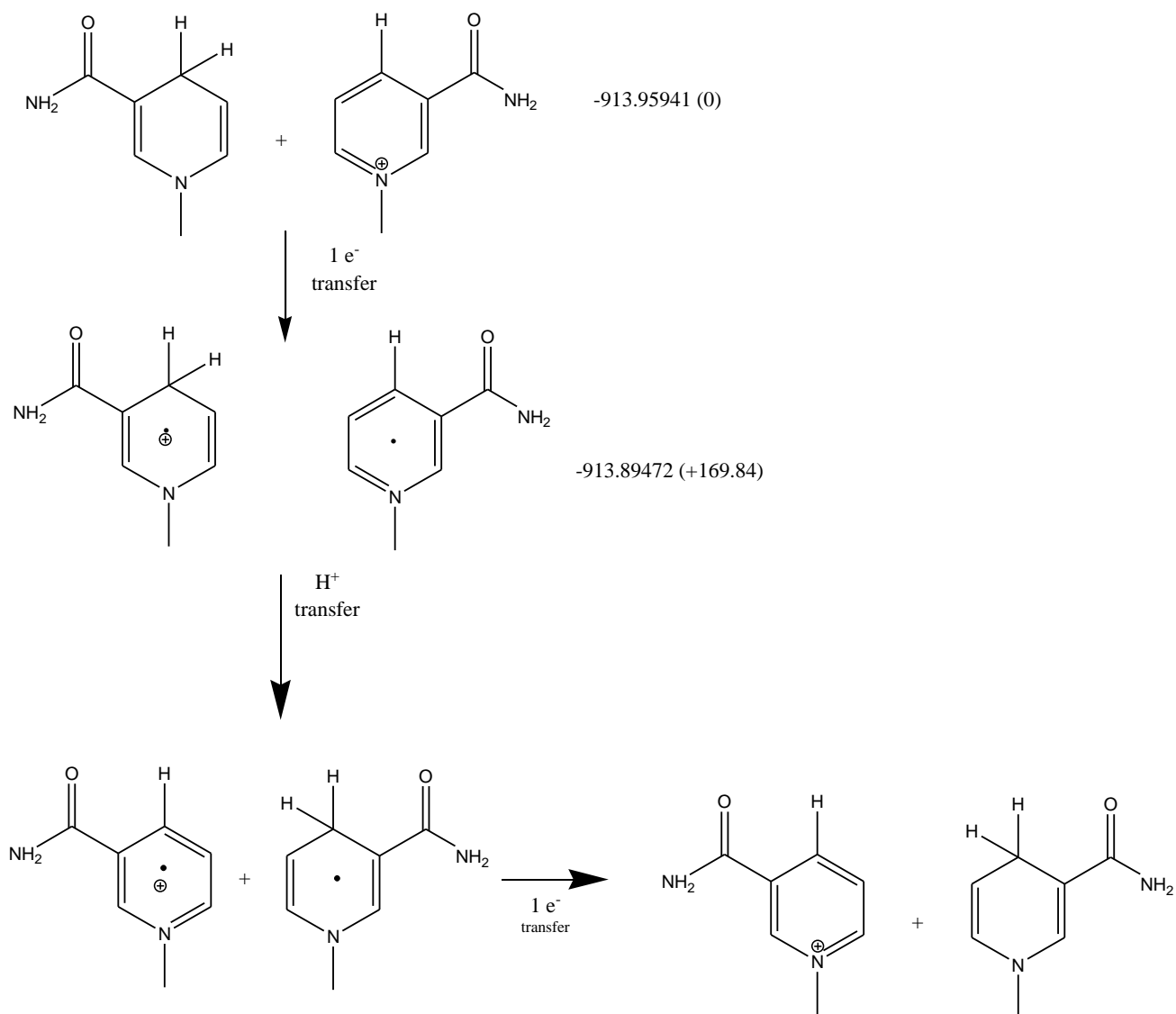
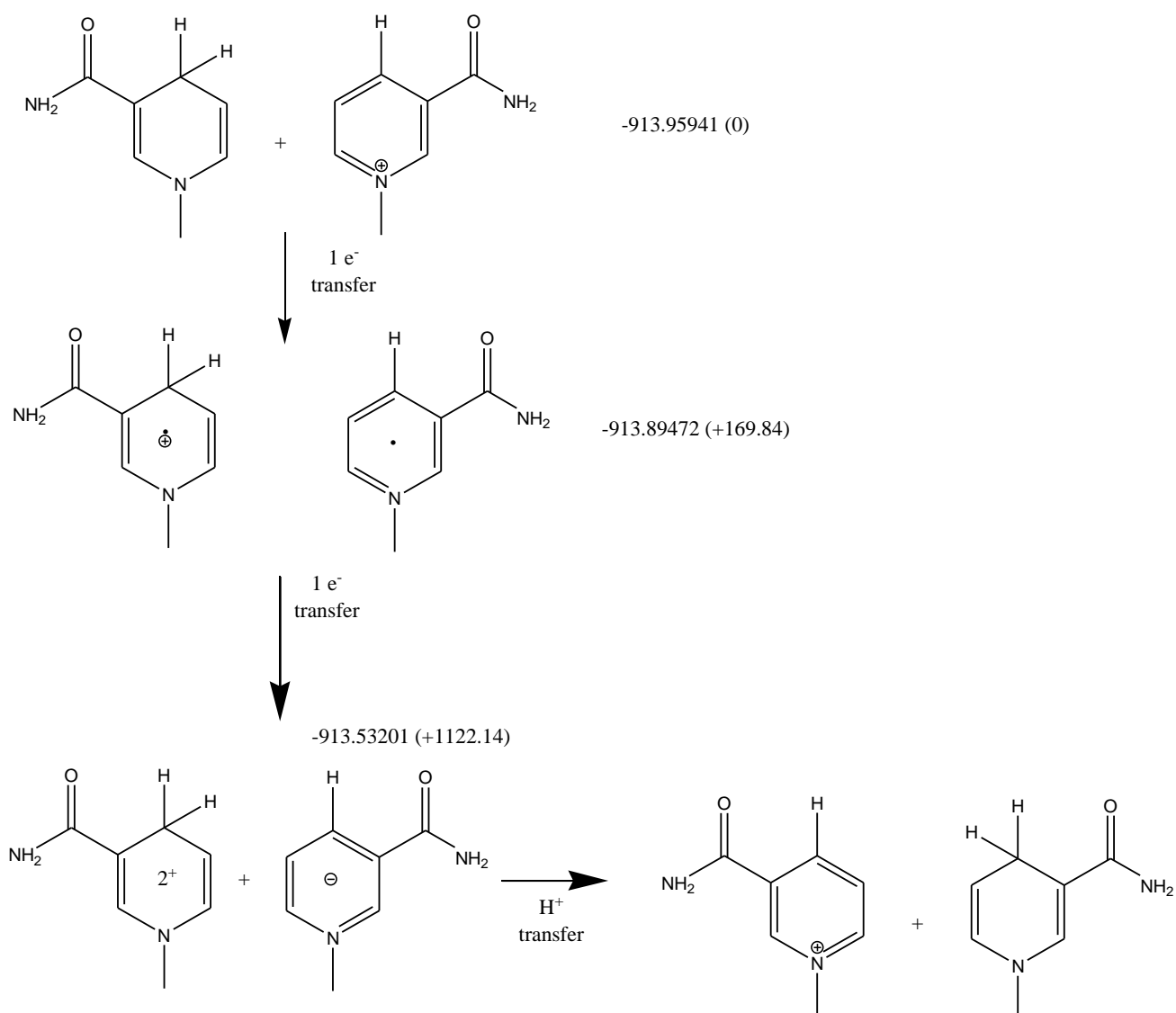
	Total Energy / Hartree	Relative Energy / kJ mol⁻¹
<p>singlet</p>  <p>+</p> <p>singlet</p>	-913.95941	0
 <p>doublet</p> <p>+</p> <p>doublet</p>	-913.89472	+169.84
 <p>singlet</p> <p>2⁺</p> <p>+</p> <p>singlet</p> <p>⊖</p>	-913.53201	+1122.14
 <p>Lowest energy TS for direct hydride transfer</p>	-913.97172	-32.32

Table 4.1 Intermediates of sequential electron and proton mechanisms



Scheme 4.1 The Sequential Electron/Proton/Electron Transfer Reaction.

The total energies of the reactant and intermediate steps are shown in Hartree (relative energies in kJ mol in parentheses).



Scheme 4.2 The Sequential Electron/Electron/Proton Transfer Reaction.

The total energies of the reactant and intermediate steps are shown in Hartree (relative energies in kJ mol in parentheses).

Schemes 4.1 and 4.2 above illustrate the two distinct electron/proton/electron and electron/electron/proton transfer mechanisms. Comparison of the energies of the intermediates, as shown in Table 4.1 to the lowest energy transition state for direct hydride transfer, which has a energy of 913.97172 Hartree confirm the nature of these intermediates as high energy. For the electron/proton/electron transfer mechanism (Scheme 4.1), the highest energy intermediate is the one following the first single electron transfer step, having an energy of -913.89472 Hartree. For the electron/electron/proton transfer mechanism (Scheme 4.2), the highest energy intermediate is the one following the second single electron transfer step, having an energy of -913.53201 Hartree.

The high energy gas-phase intermediates discussed above suggest that direct hydride transfer is preferred to electron and proton transfer mechanisms because the lowest energy transition state for hydride transfer is significantly lower in energy than the intermediates, by $\sim 32 \text{ kJ mol}^{-1}$ when compared to the lowest of the high energy intermediates, and thus this observation validates the comprehensive investigation of the hydride transfer transition states as detailed in the next chapter.

4.2 Energetics of N-methyl Pyridine Hydride Transfer Transition States

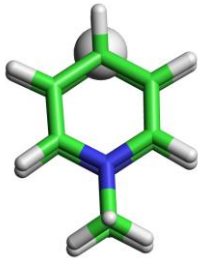

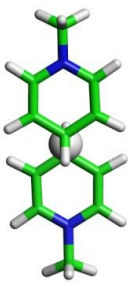
As a precursor to the comprehensive study of the nicotinamide hydride transfer transition states discussed at length in the next chapter, N-methyl pyridine hydride transfer transition states for each of the three possible ring geometries, C_2 , C_{2h} and C_{2v} , were optimised. Optimisation of these transition states repeat the earlier work of Wu *et al.* [84] albeit with two important improvements over their approach. Firstly, our structures incorporate an N-

methyl group into the ring providing an improved model analogue to the natural biological molecule ring system, and secondly our computational methodology employs a higher level of theory (DFT/6-31 G**) compared to their simple Hartree-Fock methodology (HF/6-31G*) [84]. These transition states are also important as upon optimisation they provide simple starting 'core' transition state structures for each possible ring geometry onto which the amide groups can be added to generate the nicotinamide hydride transfer transition states.

Table 4.2 shows the results from the TS optimisations of the 'core' structures, and demonstrates that whilst all three ring orientations exist here as first order transition states with optimal C-H distances for hydride transfer being 1.35 Å (and 1.34 Å for C_{2h}) and dihedral angles of 180° for C_{2h}, 0.00° for C_{2v} and 103.09° for C₂.

Table 4.3 summarises the results of Wu *et al.* for their pyridine hydride transfer transition states. These energies are once again relative to the lowest energy first order transition state they isolated, which had a C_{2v} ring geometry and an energy of -494.85051 Hartree with the 6-31 G**/3-21G methodology and an energy of -496.45623 Hartree when a post Hartree-Fock MP2 electronic correlation [85-90] is included. As seen in Table 4.3 below, using the 6-31 G**/3-21G basis set, the energy of the C₂ orientation was observed to have intermediate energies lying somewhere between the two extremes of their lowest energy C_{2v} and highest energy C_{2h} transition state structures [84], a trend reflected in the N-methyl pyridine transition states that we isolated. Wu *et al.* also noted that their method of incorporating electronic correlation resulted in a significant difference between the results when compared to the HF results, as the MP2 correlation is known to overestimate the stabilisation of a transition state [84]. Thus the relative energies seen from our DFT

calculations (Table 4.2) are much closer to those seen from the simple HF results from Wu *et al.* [84] than to their MP2-based results - a spread of 6 kJ mol⁻¹ comparing well with 7 kJ mol⁻¹ rather than 25 kJ mol⁻¹.

Structure	C1- H / Å	C2- H / Å	Relative Energy / kJ mol ⁻¹	TV Frequency (imaginary)
C_{2v}				
	1.35	1.35	0	767.6244
C₂				
	1.35	1.35	+0.76	785.1876
C_{2h}				
	1.35	1.34	+6.01	781.9820

θ (Opt) = Dihedral Angle (Optimised Structure)

Table 4.2 - N-methyl Pyridine Hydride Transfer Transition States

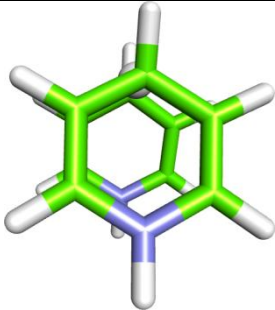
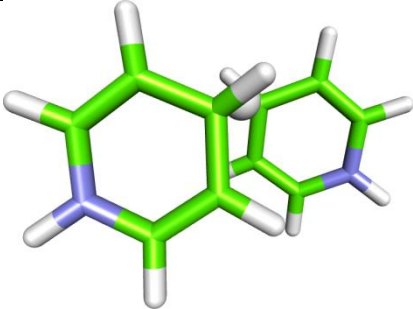
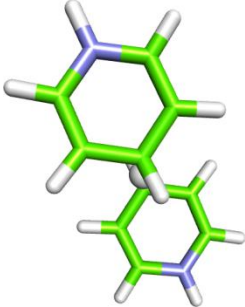
Structure	Relative Energy / kJ mol ⁻¹ [HF 6-31G*//3- 21G] [84]	Relative Energy / kJ mol ⁻¹ [MP2/6-31G*//3-21G] [84]
C_{2v} 	0	0
C_2 	+3.35	+14.23
C_{2h} 	+7.11	+25.10

Table 4.3 Wu *et al.* [84] Pyridine Hydride Transfer Transition States

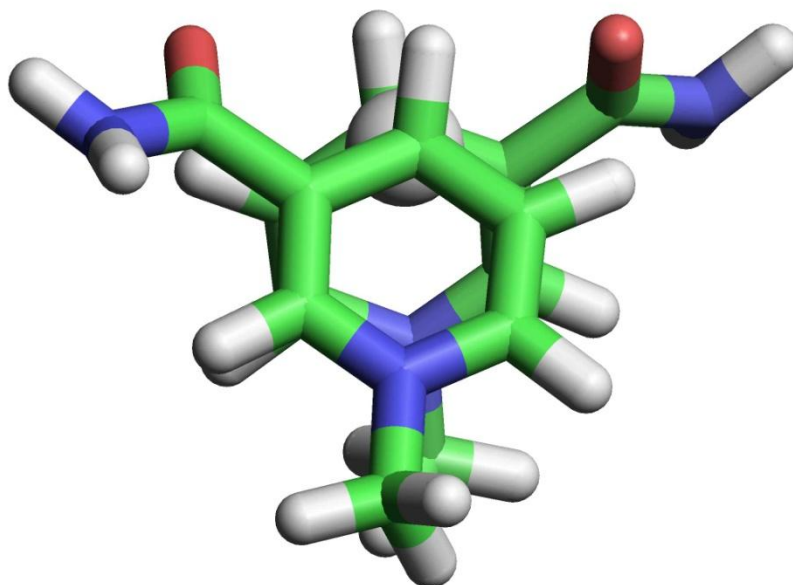
4.3 Disproving the Symmetrical Transition State Hypothesis

In 1995, Wu *et al.* [84] proposed that the hydride transfer reaction proceeded through a symmetrical transition state with a C_{2v} ring geometry. It is important to note here that this proposition was made firstly without the benefit of improved model systems (incorporating N-methyl groups into the nicotinamide rings) and secondly without a higher level of theory (using DFT/6-31G** and including an electronic correlation contribution) compared to *ab initio* calculations which implemented HF/6-31G* [84]. Perhaps the most important consideration to make when comparing the evidence for an asymmetric transition state for the reaction as discussed in Chapter 5 to the findings of Wu *et al.* is that their conclusions were made in the absence of a crystal structure, and formulated solely on the reasoning that a symmetrical ($A+B \rightarrow B+A$) reaction proceeds through a symmetrical transition state [6, 84] and that based on this reasoning, their model transition states for the hydride transfer reaction were constrained to a symmetrical C_{2v} ring geometry [84]. These nicotinamide hydride transfer transition states as well as all those discussed in greater detail in the next chapter were subject to no such conformational restraint.

The observed C_2 symmetry of the nicotinamide rings in the crystal structure, does not fit the earlier assumption made by Wu *et al.* in 1995 which put forward the notion that a symmetrical C_{2v} ring symmetry with its stacking interactions between the rings is preferred for hydride transfer [84]. Extrapolating from their assumption of a symmetrical transition state, they isolated two possible C_{2v} transition states for hydride transfer, the N-methyl analogue equivalents of which we also isolated and are shown in Figure 4.1. Neither of these high energy symmetrical transition states resemble the arrangement of the rings found in the X-ray crystal structure. Whilst they are discussed here in the context of

preliminary work and in the context of the work of Wu *et al.* [84], they were found as part of the comprehensive transition state search presented in the next chapter.

(a) (E)-C2V-OO



(b) (E)-C2V-NN

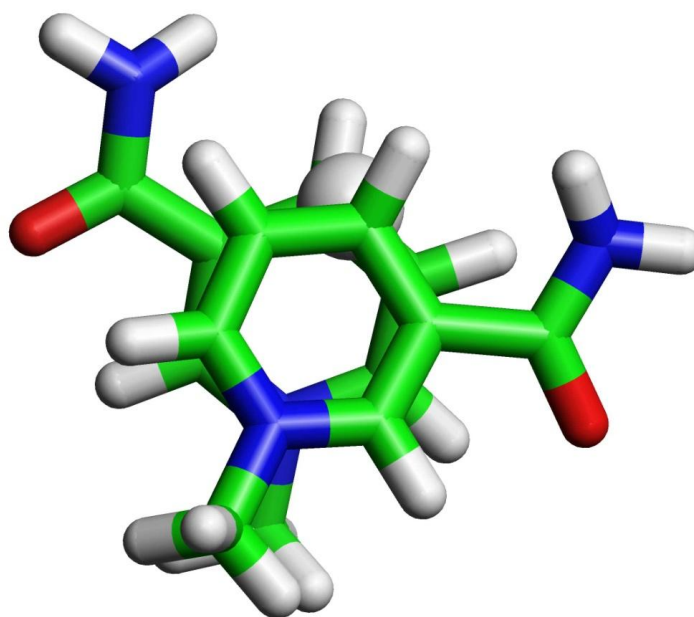
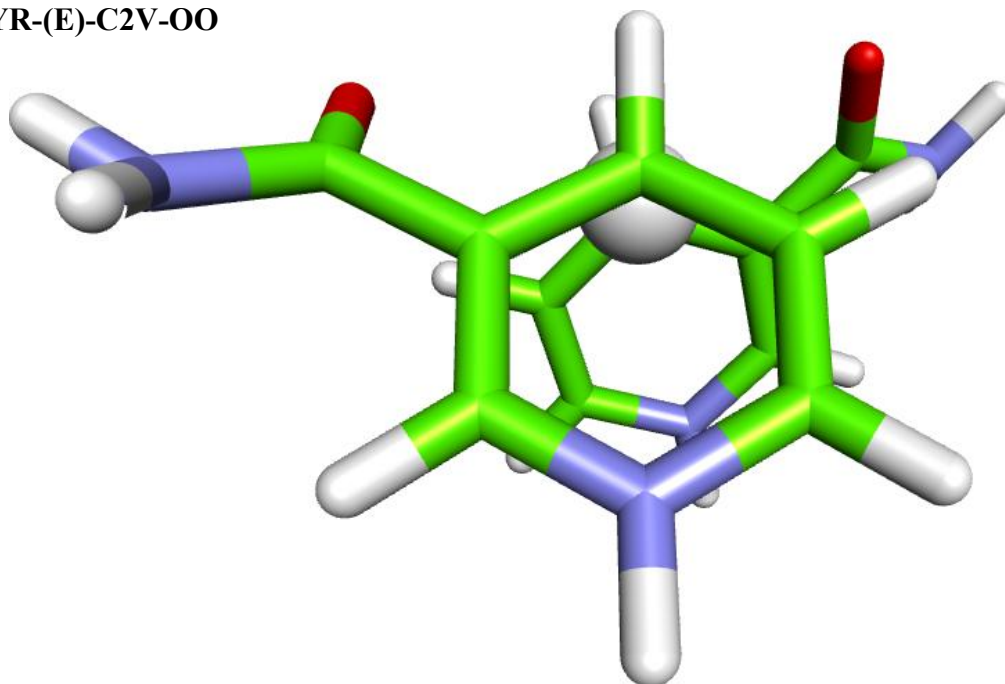


Fig. 4.1 The two equivalents of the two C_{2v} transition state structures proposed by Wu *et al.* [84] that have a two-fold rotation symmetry *not* observed in the enzyme: (a) (E)-C2V-OO and (b) (E)-C2V-NN (see section 3.2 for nomenclature details).

The observation that the nicotinamide rings in the crystal structure of transhydrogenase have no symmetry except the one fold rotation axis as opposed to at least a two-fold symmetry rotation axis (a C_2 arrangement) is perhaps even more surprising when you consider that in the case of the N-methyl pyridine analogues (without the amide groups attached), a symmetrical C_{2v} ring arrangement with stacking and full ring overlap is energetically favoured. These conflicting observations highlight the dangers of oversimplifying a model system to study a reaction and underline the importance of electrostatic interactions between the rings and the respective amide groups when they are introduced into the model system, therefore the role of the amides themselves in influencing hydride transfer should not be overlooked. The incorporation of the amide groups into the model transition states seems to activate hydride transfer by rather substantially lowering the energy of the reaction, an observation supported by Wu *et al.* and reflected in the energies for their transition states (see Figure 4.2) [84].

It is important to note here that in describing the transition state structures shown in Figures 4.1 and 4.2 respectively, the naming conventions detailed in Section 3.2 of the previous chapter are followed and thus for the structures have the same designations when the amides are included, so to distinguish between the two sets of structures discussed here, the Wu *et al.* transition states have the prefix 'PYR' added to their designation (Figure 4.2).

(a) PYR-(E)-C2V-OO



(b) PYR-(E)-C2V-NN

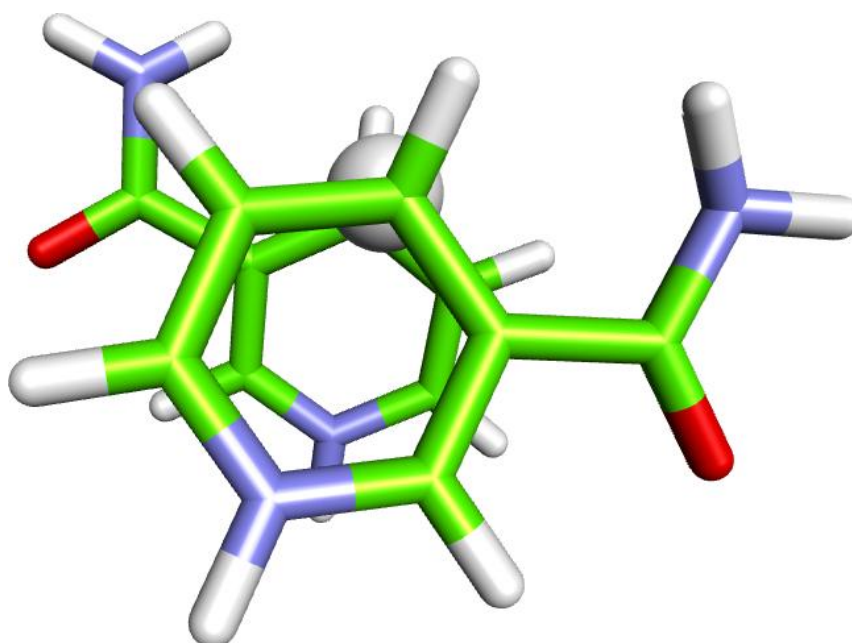


Fig. 4.2 The two C_{2v} transition state structures proposed by Wu *et al.* [84] that have a two-fold rotation symmetry *not* observed in the enzyme described with our naming system: (a) PYR-(E)-C2V-OO and (b) PYR-(E)-C2V-NN (see text and section 3.2 for nomenclature details).

The rotation of the amides as in PYR-(E)-C2V-NN (Figure 4.2 (b)) results in an increase in the energy of the transition state by approximately $19.25 \text{ kJ mol}^{-1}$ compared to PYR-(E)-C2V-OO [84]. Our results for the N-methyl analogues show a comparable transition state energy increase of approximately $13.47 \text{ kJ mol}^{-1}$ when the energy of (E)-C2V-NN (Figure 4.1 (b)) is compared to the energy of the lowest energy C_{2v} transition state we isolated, (E)-C2V-OO (Figure 4.1 (a)).

For the sake of comparison with the findings of Wu *et al.* [84], we have presented here some of our nicotinamide hydride transition states and given that there is crystallographic evidence [6] that the ring orientation found in transhydrogenase is *not* fully symmetrical, it follows that the transition state for the hydride transfer reaction is also *not* fully symmetrical and perhaps even more importantly *does not* have a C_{2v} ring geometry as proposed by Wu *et al.* [84] but rather it has a C_2 ring geometry. The next chapter comprehensively details the transition state search and contextualises our findings and explores their implications for the biology of the hydride transfer reaction in *R. rubrum* transhydrogenase.

CHAPTER FIVE

NICOTINAMIDE HYDRIDE

TRANSFER TRANSITION STATES

This chapter comprehensively covers the study of the energetics of all the theoretical nicotinamide hydride transfer transition states and ranks them from lowest to highest energetically

5. Nicotinamide Hydride Transfer Transition States

The active site of *R. rubrum* transhydrogenase seems to have naturally evolved to optimise the positions of the respective nicotinamide rings to facilitate rapid hydride transfer between the bound nucleotides, NADH and NADP⁺ as indicated by the structural similarity between the ring arrangement found in the crystal structure and the two relatively low energy transition state possibilities, C2-XE-OO and C2-XE-ON. The proximal-distal switch that needs to occur during the hydride transfer reaction by moving the nicotinamide rings to either allow or block the transfer of hydride is necessary for the efficiency of coupling to proton translocation [6, 84], and it appears that this switch between hydride transfer being blocked and hydride transfer being allowed occurs without the need to compromise on the optimised position of the nicotinamide rings that facilitate the redox chemistry of hydride transfer [6].

As detailed in Section 3.1-2, a complete theoretical set of nicotinamide hydride transfer transition states for hydride transfer would contain 48 unique structures (64 when the chirality of the C₂ ring configuration is also considered) due to all the different combinations of ring orientations (C₂, C_{2v} and C_{2h}), amide positions (L or R) and amide orientations (N or O) that are possible. As noted in Chapter 3, the true number of theoretical transition state structures is a little lower as some of these 48 structures would form symmetry related pairs. It is difficult to determine which structures would be symmetry related, because this is dependent on the naming system and are not always easily identified. The nomenclature used herein to describe these transition state structures is discussed earlier in Section 3.2.

5.1 The Transition State Search

The search for nicotinamide hydride transfer transition states completes a previous 2007 study [6] by comprehensively considering the C_{2v} transition state structures and investigating further the lack of first order C_{2h} transition states. The comprehensive results from the extensive search for transition states for hydride transfer between nicotinamide and dihydronicotinamide are reported below in Table 5.1. Perhaps the most striking observation is that of the 48 possible first order transition state structures for the hydride transfer reaction that could in theory be isolated, only 15 structurally and energetically distinct first order TS structures were found during the course of this study. All the possible permutations of amide position, amide orientations and ring orientations have been considered suggesting that these are the only these 15 transition states that exist for this reaction.

Altogether, including mirror images and duplicates, the 48 starting structures yielded 43 optimised first order TS structures from which 15 distinct transition state structures could be identified. The use of default gradient thresholds in Gaussian result in optimised TS structures within a specific group which do not possess exactly the same energy so only the lowest energy TS within each group is of interest. This unique set of 15 TS structures are ranked energetically and presented for discussion in this chapter. For interest, the full details of the complete set of 43 can be found in Table S1 in the appendix.

Overall these results show that the C_2 based structures are lowest in energy and therefore the mechanism for hydride transfer between nicotinamide and dihydronicotinamide most likely depends on the first order transition state having the two rings in a C_2 conformation.

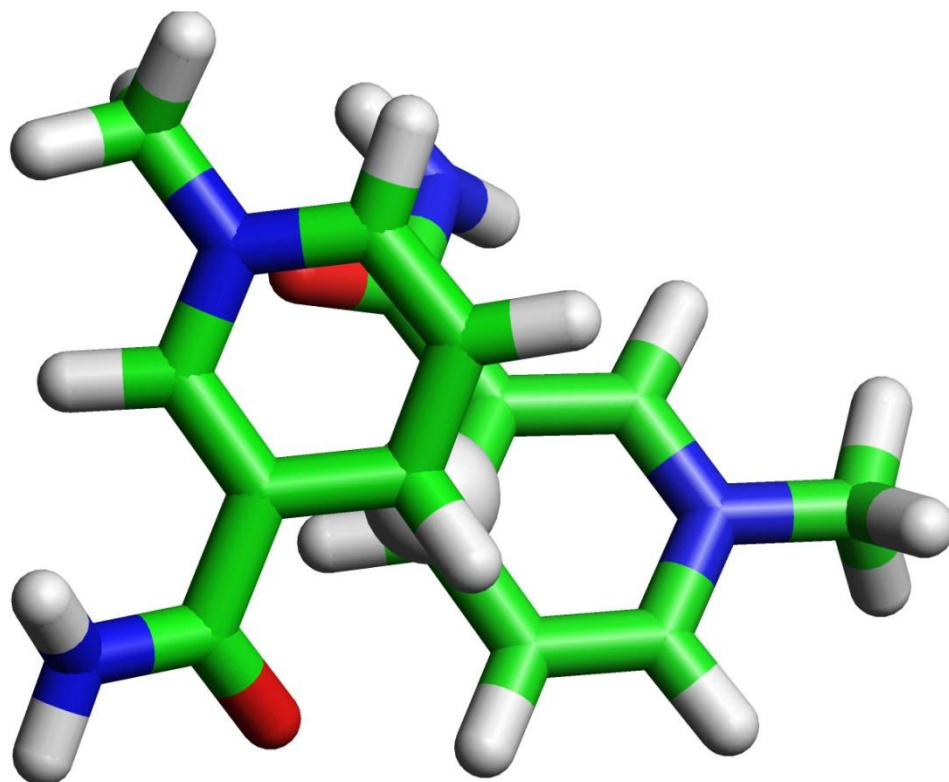
The C_{2v} structures are generally higher in energy and fewer in number whilst no first order C_{2h} TS structures exist at all. The lack of first order C_{2h} TS structures could be rationalised based on two reasons. Firstly the fact that even the N-methyl pyridine analogues discussed in the previous chapter show a C_{2h} ring orientation to be highest in energy. Secondly, the nature of the electrostatic interactions that are introduced into the system when the amide groups are included on each ring so that in both cases, there is greater scope for C_{2h} to be an energy maximum for rotation rather than a minimum. The additional electrostatic interactions due to the amide groups may also explain why so few C_{2v} TS structures were isolated as there are cases where a starting C_{2v} geometry optimises to a C_2 geometry (see Section 5.5).

Generally the energetic trends observed by Wu *et al.* [84] for the C_{2v} transition states are reflected in our results when the energies of our N-methyl pyridine transition states are compared with the energies of our nicotinamide transition states. (E)-C2V-NN with both amides being rotated N results in an approximate 18.2 kJ mol^{-1} increase in energy when compared to this lowest asymmetric TS structure, C2-XE-OO overall. The significance of amide orientation in the hydride transfer reaction is therefore readily apparent even in these C_{2v} transition states where a simple switch from both of the carbonyl oxygen atoms being orientated O to both being orientated N results in an energy increase of almost 13.5 kJ mol^{-1} . These results demonstrate therefore that the OO amide orientated form of the C_{2v} transition state is more stable than the NN amide orientated form, an observation supported by the work of Wu *et al.* [84].

Without amides, the C_{2v} transition state analogue has an energy of -576.622992593 Hartree, compared to the lowest energy C_{2v} transition state with amides, (E)-C2V-OO, which has an energy of -913.967508069 Hartree. Wu *et al.* [84] also found that for the

hydride transfer transition state structures, the activation energy is reduced by overlap of the LUMOs of the two nicotinamide rings, with the most important stabilising interactions arising from the overlap of the π orbitals of the two C4 ring atoms, and reasoned that the most favourable way of maximising this stabilising interaction was to have the two nicotinamide rings configured in a *syn* arrangement (where the transferring hydride and the carbonyl oxygen are on the same side), preferably in a C_{2v} geometry as opposed to a C_{2h} geometry where this orbital overlap is poorer. Because their hypothesis and model systems were constrained to a proper C_2 symmetry, they made no consideration for an asymmetrical C_2 ring arrangement, but these results show that even in an asymmetric transition state this overlap is still fairly substantial and that the *syn* preference of the carbonyl oxygen is maintained (Figure 5.1).

(a) C2-XE-OO(+)



(b) *Syn* preference for hydride transfer

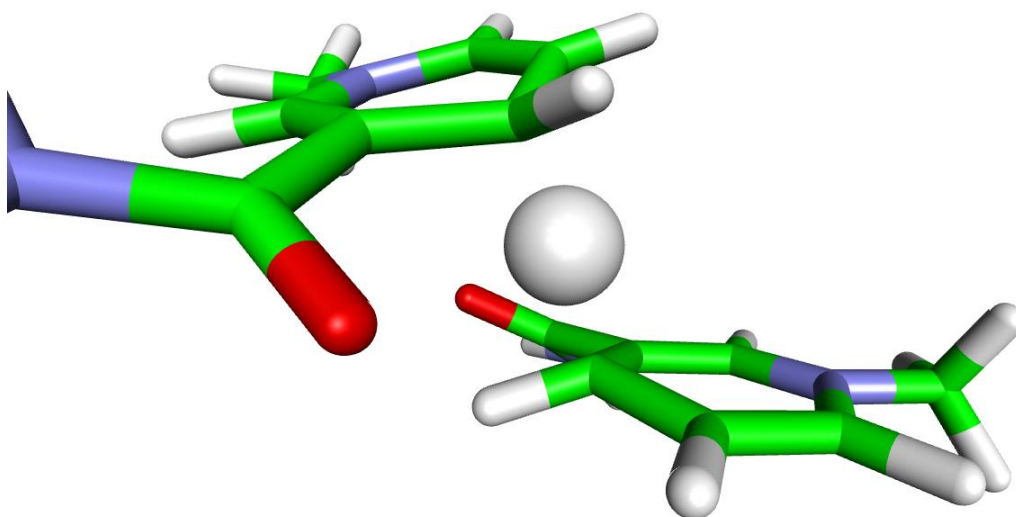


Fig. 5.1 The *syn* preference for hydride transfer: (a) showing reasonable overlap of the two C4 atoms in an asymmetric transition state C2-XE-OO, therefore reasonable overlap of the π -orbitals can be inferred maintaining the stabilising interaction and (b) showing the preference of the amide carbonyl groups to point towards the transferring hydride as well as a non-linear C-H-C angle.

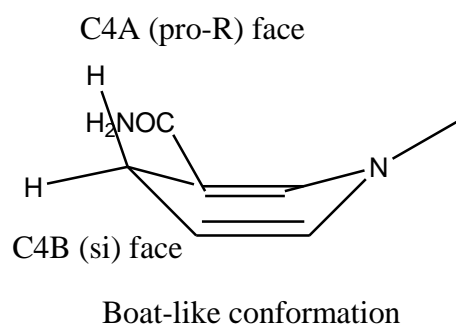


Fig. 5.2 Nicotinamide ring puckering in the hydride transfer reaction [1, 84, 91].

It is also interesting to consider that the *syn* preference also causes the nicotinamide rings to pucker slightly into boat-like conformations as illustrated in Figure 5.2. This *syn* preference is thought to arise from the electrostatic attraction of the negative carbonyl oxygen and the slightly positive hydride acceptor (polarization of the C4 atom) [84, 92]. This supposition seems to explain why the tilt of the carbonyl oxygens towards the transferring hydride is observed in both the lowest energy C_{2v} symmetric (E)-C2V-OO and the lowest energy asymmetric C2-XE-OO transition states. It can be reasoned that electrostatic attraction is the strongest in the lowest energy fully symmetric transition state C2-EE-OO where further electrostatic attraction is introduced by the proximity of the carbonyl oxygens to the positive nitrogen atom of each nicotinamide ring achieved by the *endo/endo* amide positioning of this structural configuration. Conversely, this governing attractive force is lost when the amides are rotated N as opposed to O, and therefore this goes some way in explaining why the highest energy hydride transfer transition state, C2-XX-NN, has such opposing structural features when compared to the structure of C2-EE-OO and why an electrostatic repulsion is much more prominent in this transition state structure. Wu *et al.* [84] also conceded that whilst alkylation of the N₁ position (incorporation of a methyl group) would improve their model transition states, it would

have no impact on the *syn* preference of the hydride transfer reaction, a postulation supported by this work. As well as the *syn* preference of hydride transfer, Wu *et al.* [84] also observed that a non-linear C-H-C angle was influential in hydride transfer [84], this again is supported by these results where this angle is bent to 172.61° in the lowest energy asymmetric transition state, C2-XE-OO.

It should also be noted here that, just as for the N-methyl pyridine first order hydride transfer transition states as discussed in Section 4.2 of the previous chapter, all energies for the nicotinamide first order hydride transfer transition states as reported in Table 5.1 are relative to the lowest energy transition state structure which in this case has a C₂ ring orientation and an energy of -913.971715621 Hartree (1 Hartree = 2625.498 kJ mol⁻¹).

Table 5.1 - Nicotinamide Hydride Transfer Transition States

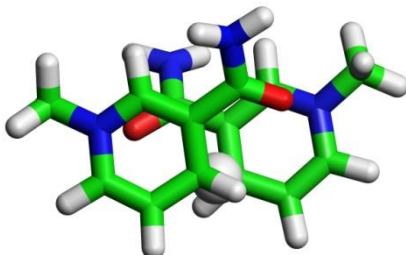
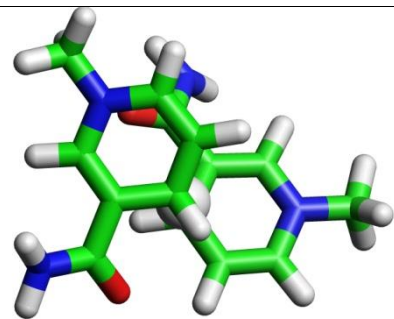
Rank	Designation	Structure	C1-H / Å	C2-H / Å	Relative Energy / kJ mol ⁻¹	TV Frequency (imaginary)	Bhakta <i>et al.</i> 2007 [6] Structure
1	C2-EE-OO (+)		1.34	1.34	0	767.3075	1
2	C2-XE-OO (+)		1.29	1.38	+6.29	694.3529	7

Table 5.1 - Nicotinamide Hydride Transfer Transition States (Continued)

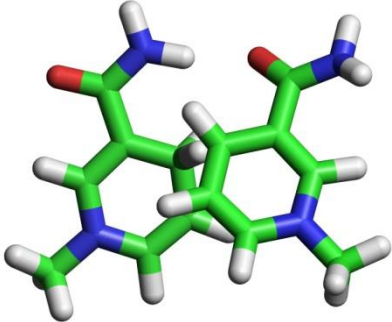
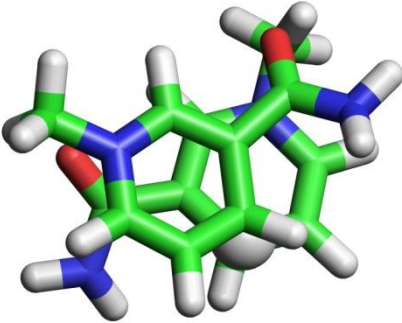
Rank	Designation	Structure	C1-H / Å	C2-H / Å	Relative Energy / kJ mol ⁻¹	TV Frequency (imaginary)	Bhakta <i>et al.</i> 2007 [6] Structure
3	C2-XX-ON (+)		1.48	1.24	+8.56	528.1576	6
4	C2-EE-NN (+)		1.36	1.36	+9.72	998.1557	2

Table 5.1 - Nicotinamide Hydride Transfer Transition States (Continued)

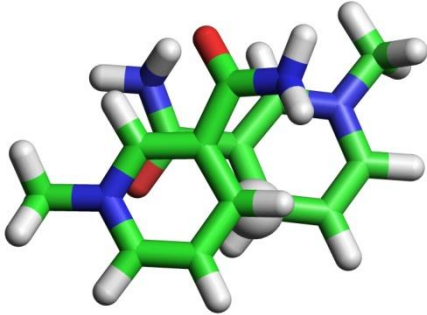
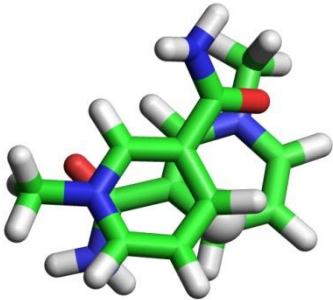
Rank	Designation	Structure	C1-H / Å	C2-H / Å	Relative Energy / kJ mol ⁻¹	TV Frequency (imaginary)	Bhakta <i>et al.</i> 2007 [6] Structure
5	C2-EE-NO (+)		1.36	1.34	+10.24	889.1855	<i>N/A</i>
6	C2-EE-ON (+)		1.38	1.33	+10.47	891.8612	3

Table 5.1 - Nicotinamide Hydride Transfer Transition States (Continued)

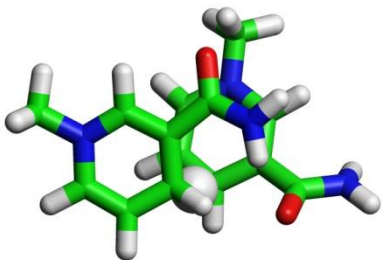
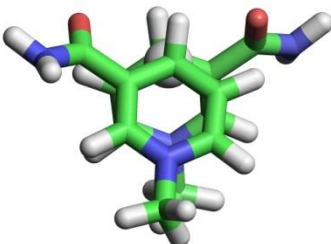
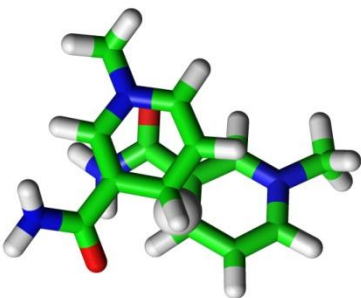
Rank	Designation	Structure	C1-H / Å	C2-H / Å	Relative Energy / kJ mol ⁻¹	TV Frequency (imaginary)	Bhakta <i>et al.</i> 2007 [6] Structure
7	C2-EX-NO (+)		1.30	1.38	+10.89	807.4717	<i>N/A</i>
8	(E)-C2V-OO		1.33	1.33	+11.05	716.0715	<i>11</i>
9	C2-XE-ON (+)		1.34	1.35	+11.33	837.2079	<i>10</i>

Table 5.1 - Nicotinamide Hydride Transfer Transition States (Continued)

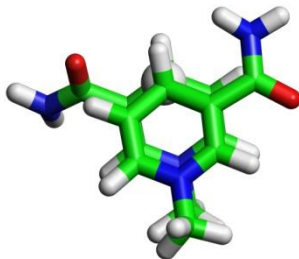
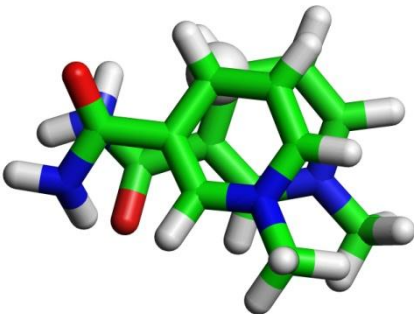
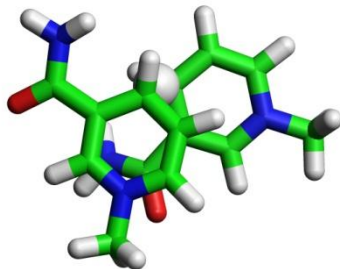
Rank	Designation	Structure	C1-H / Å	C2-H / Å	Relative Energy / kJ mol ⁻¹	TV Frequency (imaginary)	Bhakta <i>et al.</i> 2007 [6] Structure
10	(E)-C2V-NO		1.31	1.37	+17.33	745.4013	N/A
11	(Z)-C2V-ON		1.31	1.39	+23.18	824.7683	N/A
12	C2-XE-NN (-)		1.39	1.32	+23.55	878.8370	8

Table 5.1 - Nicotinamide Hydride Transfer Transition States (Continued)

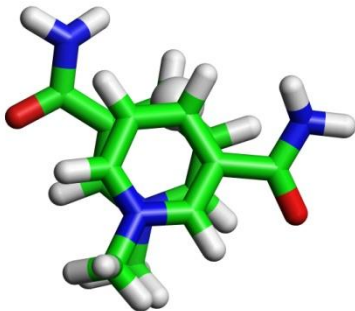
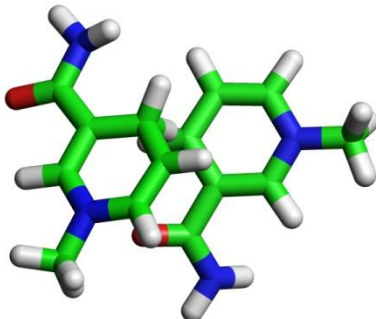
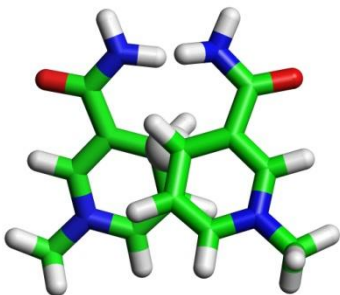
Rank	Designation	Structure	C1-H / Å	C2-H / Å	Relative Energy / kJ mol ⁻¹	TV Frequency (imaginary)	Bhakta <i>et al.</i> 2007 [6] Structure
13	(E)-C2V-NN		1.35	1.35	+24.52	874.4048	12
14	C2-XE-NO (-)		1.25	1.47	+26.02	588.8092	9
15	C2-XX-NN (+)		1.37	1.32	+31.17	820.1746	5

Table 5.1 lists the 15 unique first order nicotinamide hydride transfer transition states yielded from this search, they are ranked from the lowest to highest in energy. It is again important to emphasise here that though in the first instance this work completes the TS work from a previous study in 2007 [6], it differs by comprehensively covering C_{2v} transition states and also offering some justification as to why no first order C_{2h} transition states were isolated. Where applicable, the corresponding numerical structural designations of transition states found in that study are also reported in this table. The TS structures reported by Bhakta *et al.* were not ranked energetically as they have been here, and our results show that we found all 11 of their previously reported theoretical transition states plus four new ones (see Section 5.2 for details), resulting in a total of fifteen unique first order transition states for hydride transfer that were isolated. Neither this study or the previous one in 2007 found a C2-XX-OO TS, as illustrated in Figure 5.3 below, (Bhakta *et al.* TS structure **4** [6]), this supports the view that it does not exist. This structure is similar to the highest energy (15th) TS structure we found, C2-XX-NN (+31.17 kJ mol⁻¹) but with both amide groups in an O orientation, i.e. C2-XX-OO.

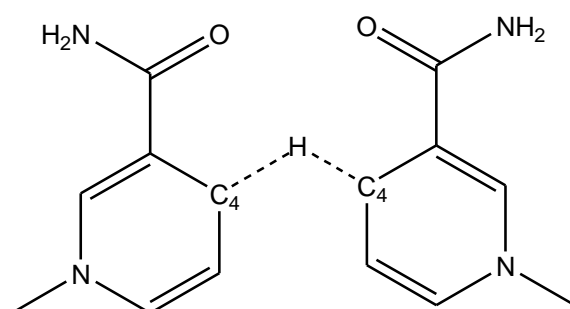


Fig 5.3 Theoretical TS structure, C2-XX-OO. This structure does not exist as shown by our work and the previous study by Bhakta *et al.* [6].

5.2 New Transition States

The search for nicotinamide hydride transfer transition states yielded fifteen unique structures, eleven of which had been previously reported by Bhakta *et al.* [6] and four of which are new. These new TS structures include two new examples of a C_2 ring orientation and two new examples of a C_{2v} ring orientation.

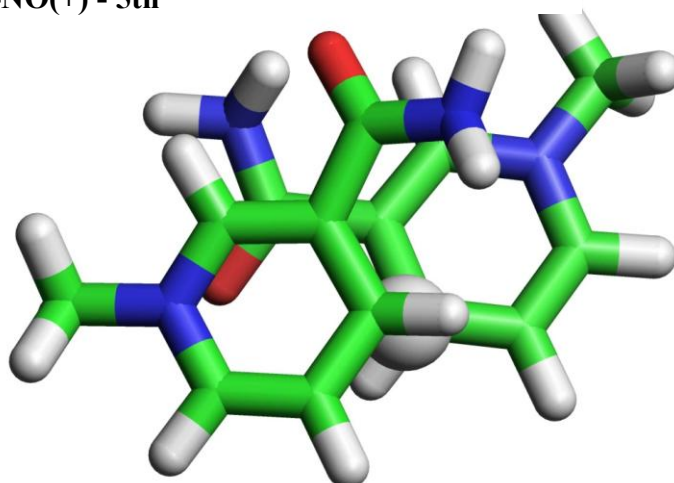
5.2-1 New C_2 Structures

The two new C_2 TS structures are shown in Figure 5.4 below. The first of the new C_2 TS, C2-EE-NO(+) is ranked 5th in Table 5.1. It is structurally similar to the overall lowest energy TS, C2-EE-OO(+), but it has the amide group of one of its nicotinamide rings has an N orientation, as opposed to both amides having an O orientation.. This small structural difference results in a slight asymmetry, due to the repulsion between the positive NH_2 group of the top ring and the positive N methyl nitrogen atom of the bottom ring. For this structure, the C-H bond distances for hydride transfer are 1.36 Å and 1.34 Å, compared to the full symmetry of C2-EE-OO (for which both C-H bond distances are 1.34 Å), and a rather significant energy difference of +10.24 kJ mol⁻¹, perhaps providing some further evidence for the influence of amide orientation upon hydride transfer.

The second new C_2 TS, C2-EX-NO(+) is ranked 7th (+10.89 kJ mol⁻¹), and is structurally similar to C2-EE-NO(+) but the amide position of the bottom ring is *exo* rather than *endo*, the relatively small difference in energy between the two new C_2 TS structures of 0.65 kJ mol⁻¹ may indicate that differences in amide position are not as influential upon hydride transfer as differences in amide orientation. Though its energy is quite similar, this

structure is considerably more asymmetric than C2-EE-NO(+), with C-H bond distances of 1.30 Å and 1.38 Å for hydride transfer.

(a) C2-EE-NO(+) - 5th



(b) C2-EX-NO(+) - 7th

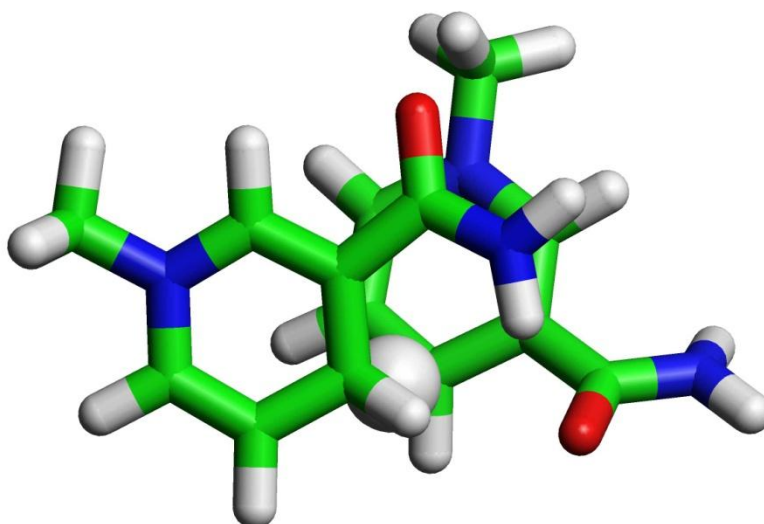


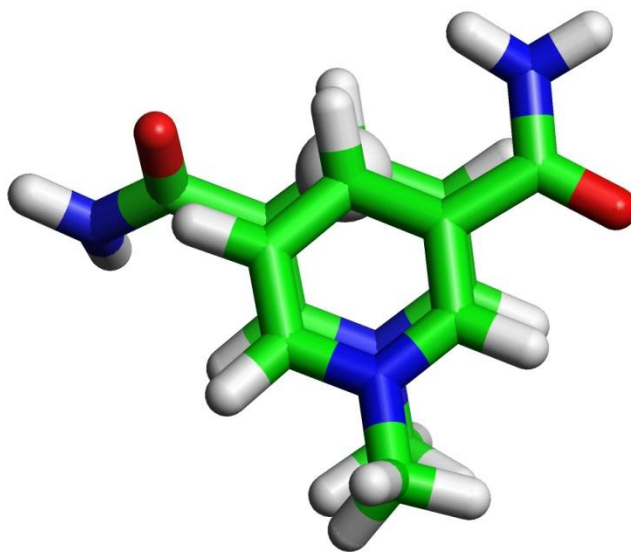
Fig. 5.4 New C_2 asymmetric transition states: (a) C2-EE-NO(+) and (b) C2-EX-NO(+)

5.2-2 New C_{2v} Structures

The two new C_{2v} TS structures are relatively high, but similar in energy, ranked 10th and 11th respectively, and shown in Figure 5.5 below. The first of these, (E)-C2V-NO, ranked 10th (+17.33 kJ mol⁻¹) is similar to the lowest overall C_{2v} TS structure, (E)-C2V-OO, which is ranked 8th (+11.05 kJ mol⁻¹) apart from the one of the amides having an N orientation as opposed to both amides having an O orientation. This change in amide rotation results in this structure being asymmetric for hydride transfer with C-H bond distances for hydride transfer being 1.31 Å and 1.37 Å. The influence of amide orientation is also seen in the highest overall energy (E)-C2V-NN, ranked 13th (+24.52 kJ mol⁻¹), which has both its amides in an N orientation rather than O as they are in (E)-C2V-OO. As its name suggests, (E)-C2V-NN is fully symmetric - with both C-H bond distances being 1.35 Å.

The second new C_{2v} TS structure, (Z)-C2V-ON, ranked 11th (+23.18 kJ mol⁻¹) is highly asymmetric (C-H bond distances for hydride transfer of 1.31 Å and 1.39 Å) and unique amongst the other C_{2v} TS structures that were isolated because it is the only TS where both amide groups are positioned *cis* (**Z**) as opposed to being positioned *trans* (**E**) as observed in the other three C_{2v} transition states that were isolated.

(a) (E)-C2V-NO - 10th



(b) (Z)-C2V-ON - 11th

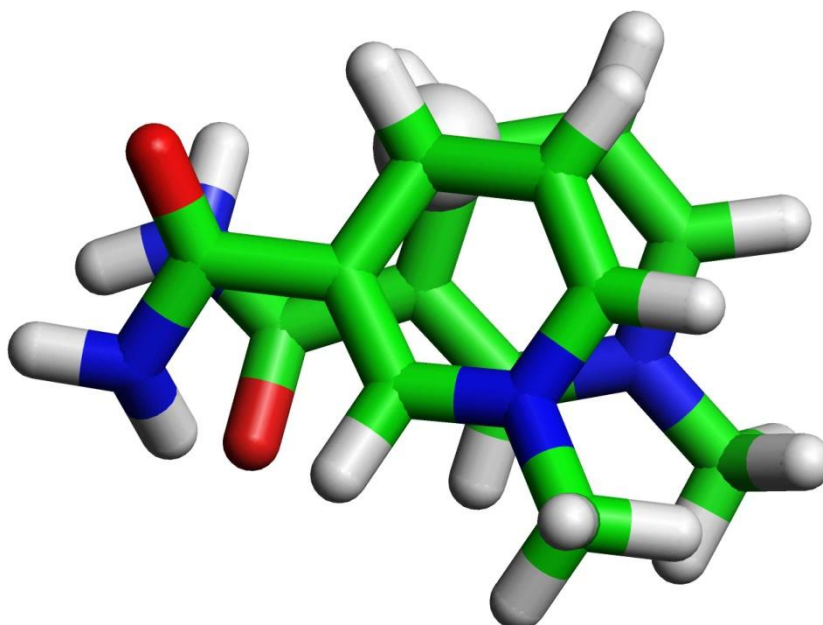


Fig. 5.5 New asymmetric C_{2v} transition states: (a) (E)-C2V-NO and (b) (Z)-C2V-ON

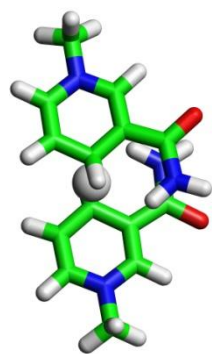
5.3 Second Order C_{2h} Transition States

First order transition state structures with a C_{2h} ring geometry were not isolated. Earlier work by Bhakta *et al.* [6] made no consideration of transition states with a C_{2h} ring geometry owing to its high energy as seen in the N-methyl pyridine analogues, and as the crystal structure ring configuration is known to be C₂. Wu *et al.* [84] optimised a first order C_{2h} transition state by employing a geometric constraint of 180° to the two rings prior to optimisation and addition of amides, whereas we put no constraints on the system prior to optimisation, and let the rings optimise with the amides already attached and thus did not force the system to adopt a specific ring geometry.

Although no first order C_{2h} TS structures exist, several second order TS structures with a C_{2h} ring geometry were isolated. Second order transition states are not considered true transition states. Generally these second order transition state structures are considered of no chemical significance because they are not accessed during a normal chemical reaction.

There is no unique IRC from a second order saddle point and therefore it is unclear as to which reactant/product that this saddle point connects. Theoretically a second order transition state IRC calculation, following each of the two vectors with negative eigenvalues would result in four structures, the reactant, the product and two first order transition states, one of which connects to the reactant and the other which connects to the product, however in practice, as the structure diverges from the second order transition state, the eigenvalues mix and attempts to perform IRC calculations on second order transition state structures often fail. In this case they failed to complete even a single cycle of convergence.

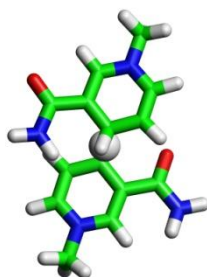
Figure 5.6 below shows the three second order transition states with a C_{2h} ring geometry that were isolated.



(a) (Z)-C2H-NN



(b) (Z)-C2H-NO



(c) (E)-C2H-NO

Fig. 5.6 Examples of second order transition states with a C_{2h} ring geometry.

Table 5.2 below shows the nature of these three transition states as second order by showing the two negative eigenvalues of each structure. Comparison of the eigenvector components shows the first of these negative eigenvalues corresponds to a hydride transfer, while the second is a rotation approximately around the C4-C4 axis.

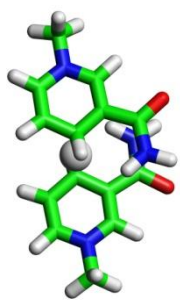
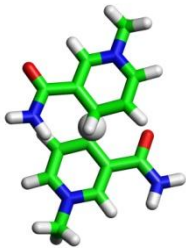
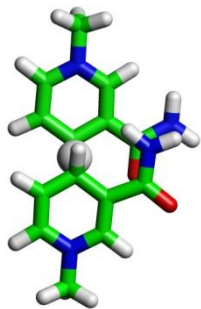
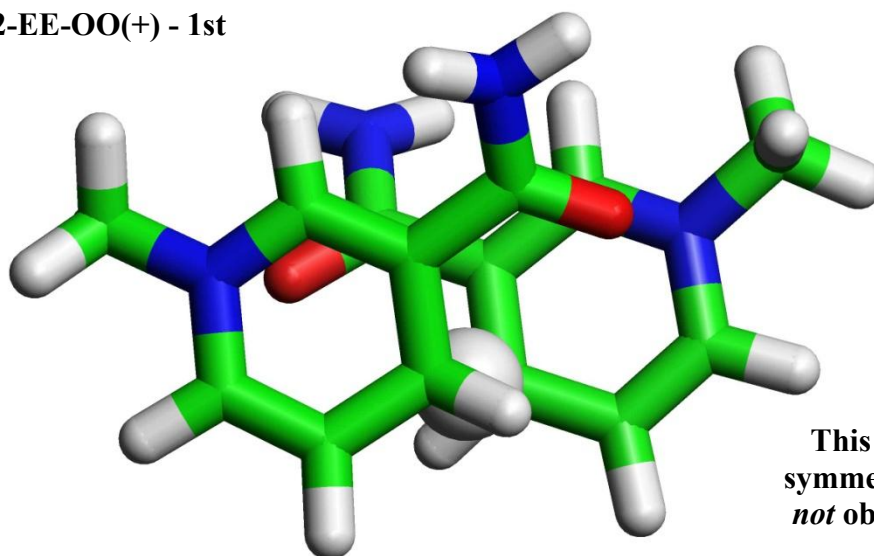
Designation	Structure	Hydride Transfer TV Freq. (imaginary)	Rotational Frequency (imaginary)
(Z)-C2H-NN		898.1208	67.0610
(E)-C2H-NO		841.7640	32.7579
(Z)-C2H-NO		746.6607	125.9847

Table 5.2 Second order C_{2h} transition states

5.4 Transition State Energies

Eleven of the fifteen unique first order transition states isolated in this study were previously reported [6] and in this section, we cover the energies of these structures. C2-EE-OO(+) is observed to be the lowest energy TS structure, ranked first with a total energy of -913.971715621 Hartree (Figure 5.7). This structure is fully symmetrical with both C-H bond distances for hydride transfer being 1.34 Å. The most notable features of this TS are the electrostatic attractions between the highly electronegative oxygen atoms of the amide groups in one ring to the positive nitrogen atom in the other ring – a tilt of the amides to this effect can be seen from both the front and the back of this structure.

C2-EE-OO(+) - 1st



This lowest energy, fully symmetric transition state is *not* observed in the enzyme

Fig 5.7 The lowest energy TS, C2-EE-OO(+), with a total energy of -913.971715621 Hartree (1 Hartree = 2625.498 kJ mol⁻¹). This structure is fully symmetric, with both C-H bond distances for hydride transfer being 1.34 Å. This TS structure is *not* observed in the enzyme.

The key observation to take particular note of with regards to this TS structure is that in spite of it being the lowest energy and fully symmetric, the amide position of one of the rings is *not* found within the enzyme. The crystal structure of transhydrogenase [6] actually shows one of the amide groups to be in an *exo* (X) position with respect to one nicotinamide ring, and the other amide to be in an *endo* (E) position with respect to the other nicotinamide ring (see Section 3.2-2, and Figure 5.8 below).

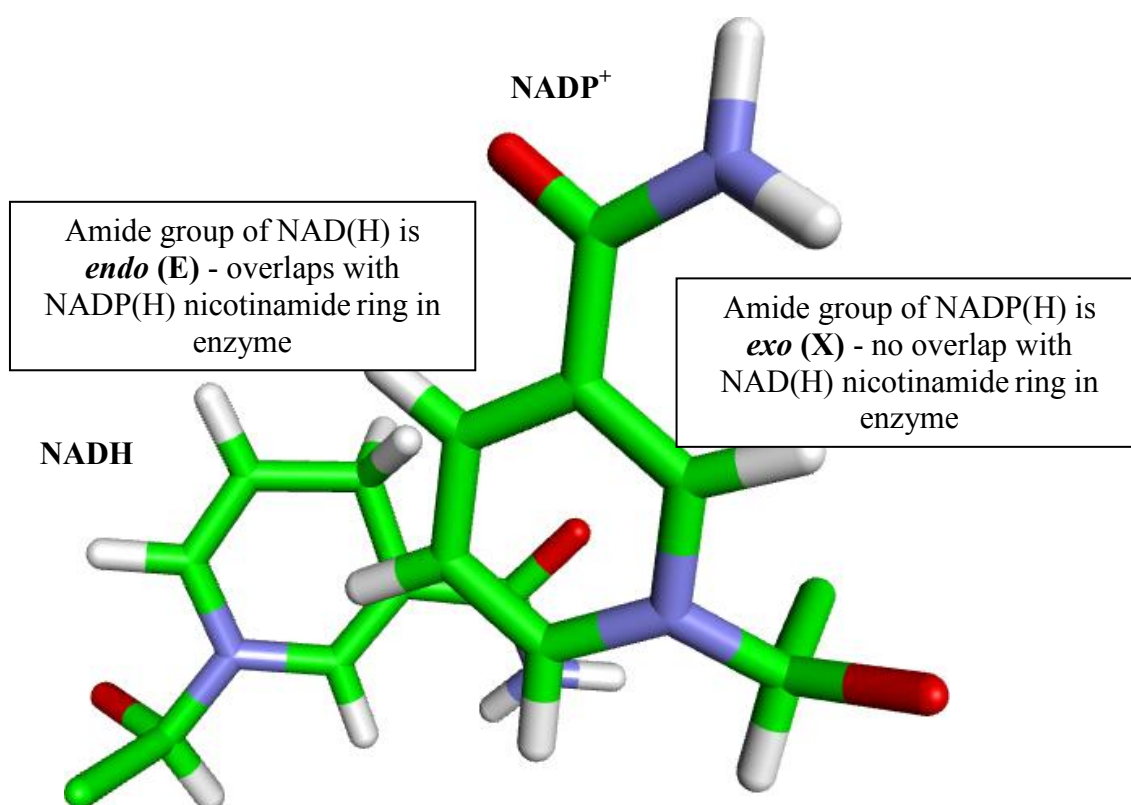


Fig. 5.8 A simplified illustration of the C₂ ring orientation and nature of the amide positions found within the crystal structure of transhydrogenase [6]. In this example NADH and NADP⁺ are bound (reactant conformation) and the nicotinamide rings are held in a proximal configuration, i.e. close enough to allow hydride transfer. For clarity and ease of visual comparison with the structural TS analogues discussed herein, only the nicotinamide rings have been shown, the rest of the respective cofactor molecules and the surrounding protein environment are not shown.

The structure that best fits the published crystal structure [6] as shown in Figure 5.8 would be designated C2-XE-OO. It is also worth recalling here that the ambiguity of the amide orientation of NAD(H) leads to two possible transition states that could fit the enzyme crystal structure, namely C2-XE-OO and C2-XE-ON (see Section 5.8).

Thus, though the underlying model reaction here is symmetrical, many of the transition states are asymmetrical in terms of amide position, and it turns out, asymmetrical in terms of hydride position and therefore in position along the reaction pathway. The nature of this asymmetry is illustrated through example by Figure 5.9.

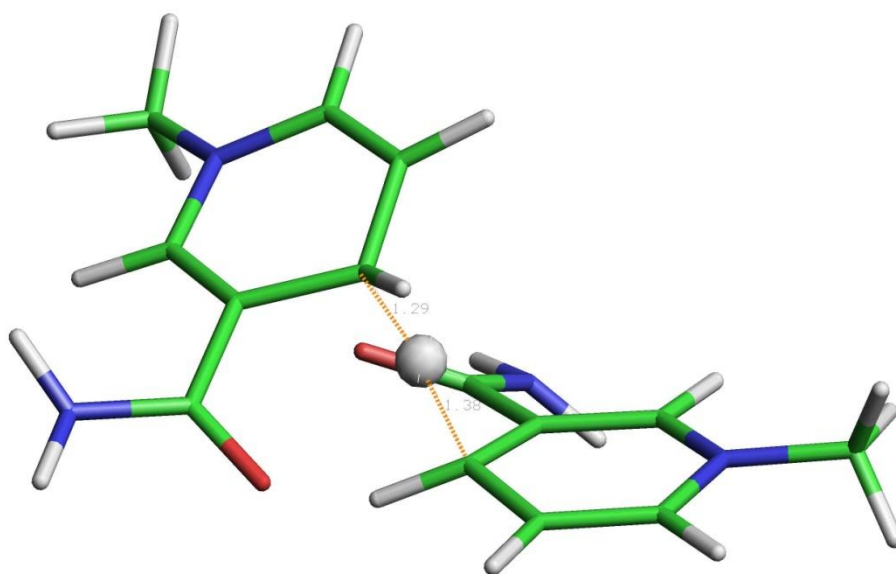


Figure 5.9 Illustrating the nature of an asymmetrical transition state, C-H bond distances of 1.29 Å and 1.38 Å respectively.

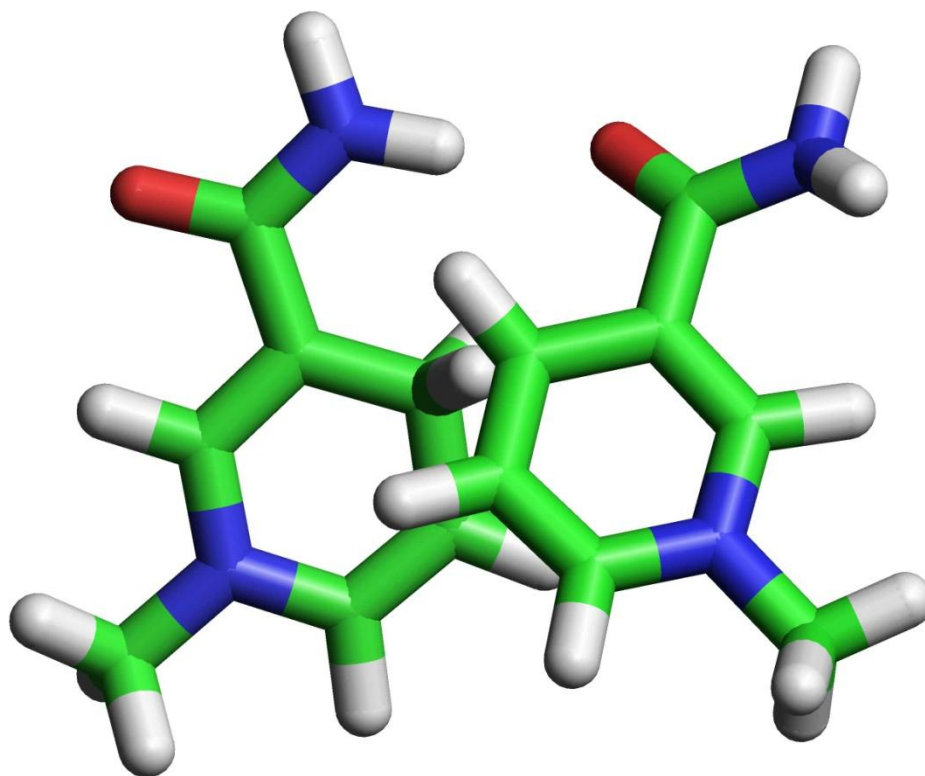
Figure 5.9 shows the asymmetric nature of a transition state which closely corresponds to the crystal structure. The asymmetry of this transition state, which is not the lowest in energy, is readily apparent as the two C-H bond distances between the C4 of each ring and

the transferring hydride are 1.29 Å and 1.38 Å respectively. It was this observation that lead to the proposal of the asymmetric hypothesis as outlined in Section 1.7.

Figure 5.10 below shows two examples of TS structures yielded by this search that are highly asymmetrical in nature, C2-XX-ON(+) and C2-XE-NO(-). As well as illustrating that highly asymmetrical TS structures can be both relatively low and relatively high in energy, these examples also serve as good illustrations of the C₂ ring orientation N1-C4-C4-N1 torsion angle (θ) consideration that is made explicit in the full structural designations of these transition states as these particular TS structures have opposite torsion angles, (+) and (-) respectively.

C2-XX-ON(+) has C-H bond distances of 1.48 Å and 1.24 Å, and so is the most asymmetric of all the transition states, but is comparatively low energy and is ranked 3rd (+8.56 kJ mol⁻¹). There is hydrogen bonding interaction between the amide hydrogen atoms on the bottom ring and the oxygen atom of the amide of the top ring. The higher energy example, C2-XE-NO(-) has similar C-H bond distances for hydride transfer of 1.25 Å and 1.47 Å, but is ranked 14th (+26.02 kJ mol⁻¹), the second highest in energy overall.

(a) C2-XX-ON(+) - 3rd



(b) C2-XE-NO(-) - 14th

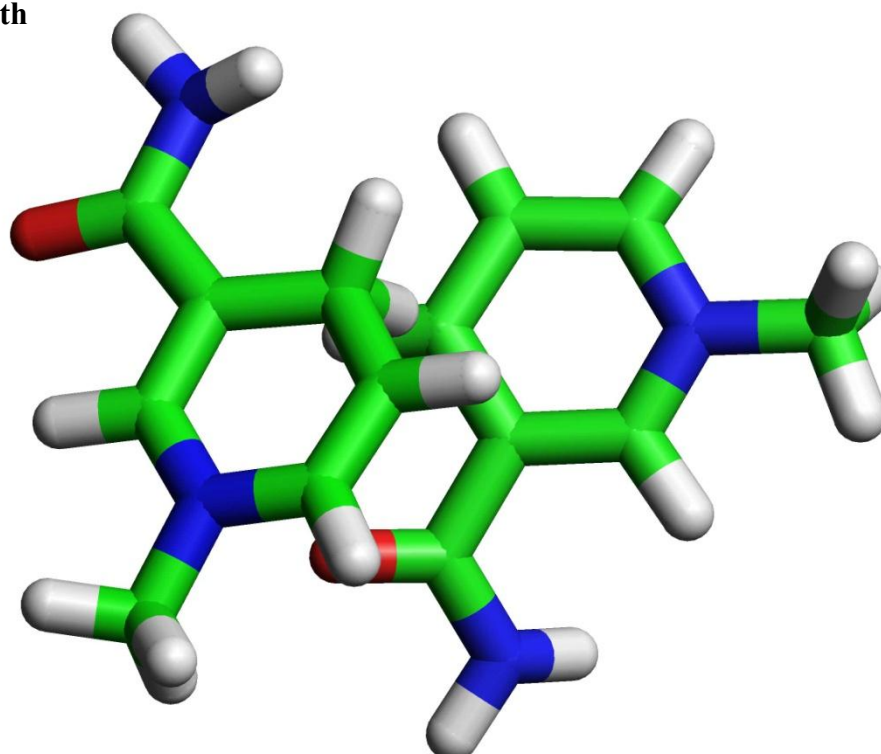
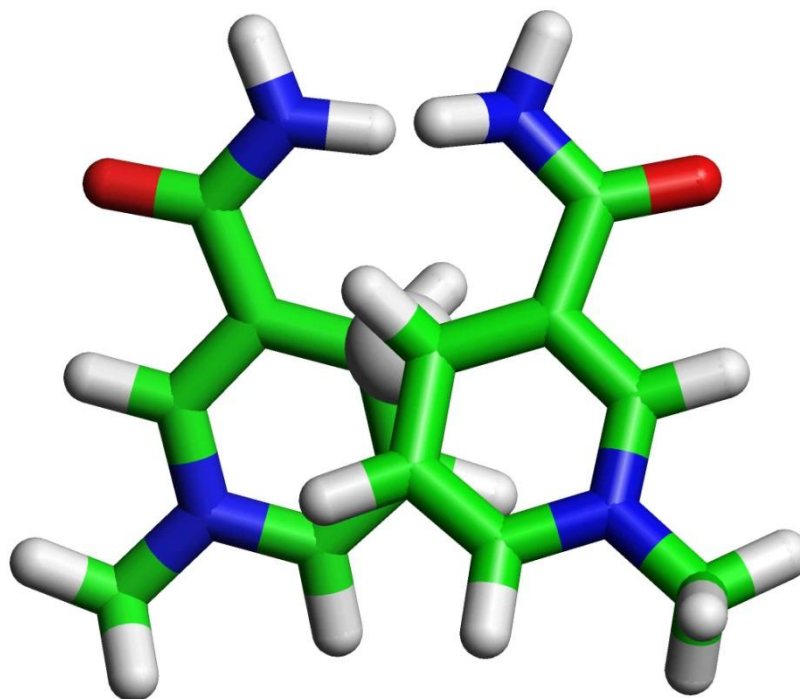


Fig 5.10 Highly asymmetric transition states: (a) C2-XX-ON(+), C-H bond distances of 1.48 Å and 1.24 Å, and (b) C2-XE-NO(-), C-H bond distances of 1.25 Å and 1.47 Å for hydride transfer, respectively.

The opposite torsion angles of the two TS structures, C2-XX-ON(+) and C2-XE-NO(-) are readily apparent in Figure 5.10.

(a) C2-XX-NN(+) - 15th



**(b) Asymmetric tilt of amides
in C2-XX-NN(+)**

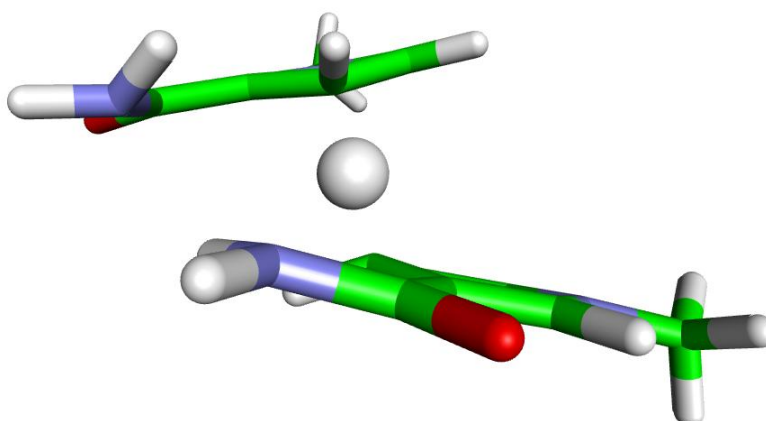


Fig. 5.11: (a) the highest energy TS, C2-XX-NN(+) (+31.17 kJ mol⁻¹), with C-H bond distances for hydride transfer of 1.37 Å and 1.32 Å and (b) the asymmetric tilt of amides in this structure

Of the 15 unique TS structures, C2-XX-NN(+) is the highest in energy, (Figure 5.11 (a)), with an energy difference of +31.17 kJ mol⁻¹, relative to the lowest energy TS, C2-EE-OO(+). Perhaps the most striking observation to be made here is that all the key structural features of this TS are the exact opposite of those structural features found in the lowest energy TS (Figure 5.7). C2-EE-OO(+) is fully symmetric in nature, with both C-H bond distances being 1.34 Å whilst C2-XX-NN(+) is asymmetrical with C-H bond distances of 1.37 Å and 1.32 Å despite its symmetrical nomenclature. The asymmetry of this TS is due to the asymmetric tilt of the amide groups (Figure 5.11 (b)) where it can be clearly seen that they are rotated slightly out of the plane with one NH₂ group pointing towards the transferring hydride and the other pointing away. The amide groups of C2-XX-NN(+) are both *exo* as opposed to *endo* as they are in the lowest energy transition state, C2-EE-OO(+). As well as being oppositely positioned, the amide groups of the highest energy TS are also both rotated N compared to the lowest energy TS, in which they are both rotated O. Another key feature is that rather than a strong electrostatic attraction as seen in C2-EE-OO(+) being present in C2-XX-NN(+), a repulsion can be seen between the two δ⁺ NH₂ groups of the respective amides on each nicotinamide ring.

5.5 Mirror Images and Duplicates

There are a significant number of TS structures that are either mirror images or duplicate structures as listed in Table S1. Generally if a structure has an energy in Hartree corresponding to another structure to 6 or 7 decimal places, and a similar TV frequency it is considered a duplicate. Duplicates arise from changes in ring geometry that occur during optimisation - an initial C_{2h} TS structure optimising to a final C₂ TS structure, for example, and from symmetry related starting points..

Table 5.3 lists the initial (unoptimised) and final (optimised) torsion angles (θ) for each of these transition states. This information is most useful in identifying changes in ring geometry during optimisation. Figure 5.12 below illustrates by example, the C-H the bond distances and torsion angle for a fully symmetric TS, in this case, C2-EE-OO(+).

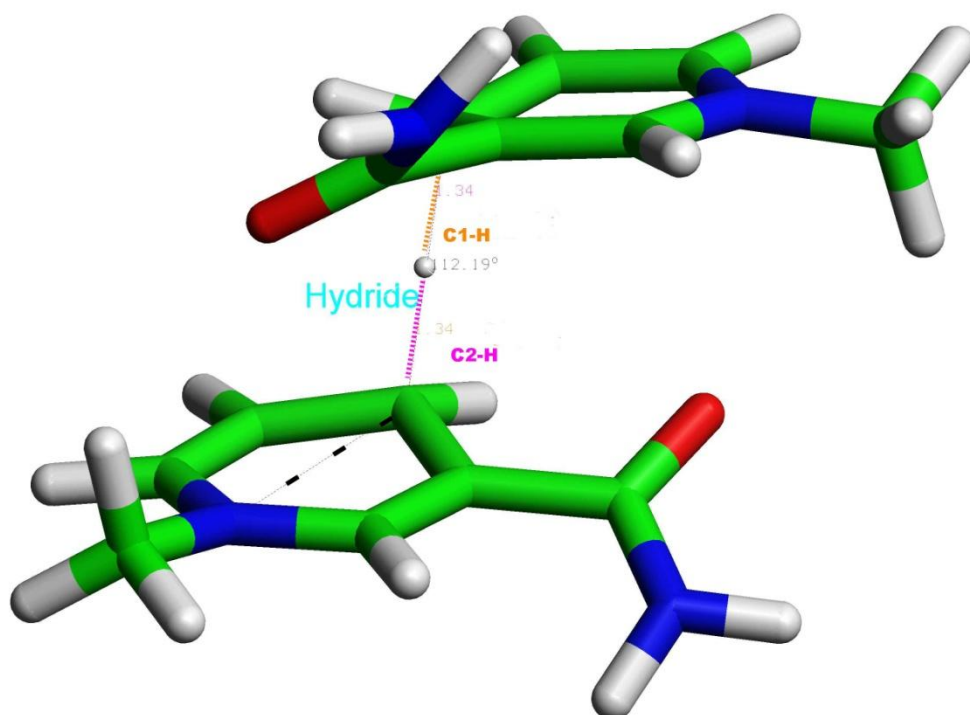


Fig. 5.12 A 3D representation of a nicotinamide hydride transfer TS, illustrating the bottom part of the N1-C4 (θ) torsion angle (black dashed line). These atoms have been explicitly identified for clarity, only the C4 atom of the top ring is obscured from view by the NH₂ of its amide group. The torsion angle is positive (+112.19°) thus the full structural designation for this TS would be C2-EE-OO(+). Also shown are the two C-H bond distances (orange C1-H and purple C2-H), both are 1.34 Å and in this case means that this TS is fully symmetric.

Table 5.3 - Torsion Angles of Nicotinamide Hydride Transition States

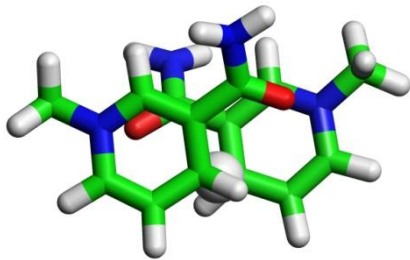
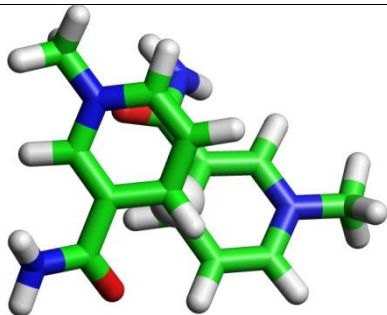
Rank	Designation	Structure	θ (Unopt.) / Degrees	θ (Opt.) / Degrees	Relative Energy / kJ mol^{-1}	TV Frequency (imaginary)
1	C2-EE-OO (+)		-179.98	+112.19	0	767.3075
2	C2-XE-OO (+)		-1.14	+120.67	+6.29	694.3529

Table 5.3 - Torsion Angles of Nicotinamide Hydride Transfer Transition States (Continued)

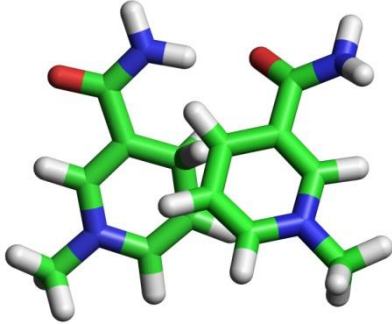
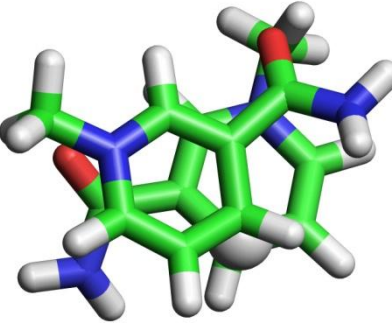
Rank	Designation	Structure	θ (Unopt.) / Degrees	θ (Opt.) / Degrees	Relative Energy / kJ mol^{-1}	TV Frequency (imaginary)
3	C2-XX-ON (+)		+73.50	+86.80	+8.56	528.1576
4	C2-EE-NN (+)		+77.49	+77.48	+9.72	998.1557

Table 5.3 - Torsion Angles of Nicotinamide Hydride Transfer Transition States (Continued)

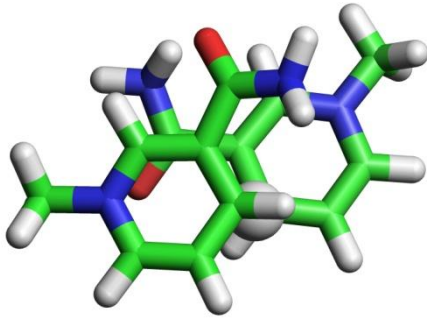
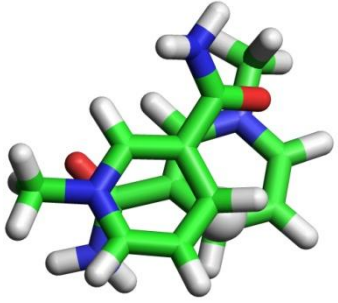
Rank	Designation	Structure	θ (Unopt.) / Degrees	θ (Opt.) / Degrees	Relative Energy / kJ mol^{-1}	TV Frequency (imaginary)
5	C2-EE-NO (+)		-174.86	+121.46	+10.24	889.1855
6	C2-EE-ON (+)		+97.17	+97.16	+10.47	891.8612

Table 5.3 - Torsion Angles of Nicotinamide Hydride Transfer Transition States (Continued)

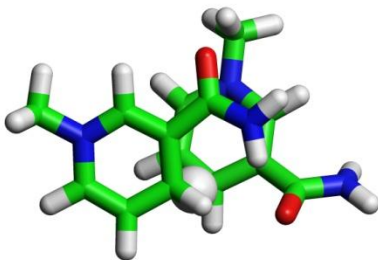
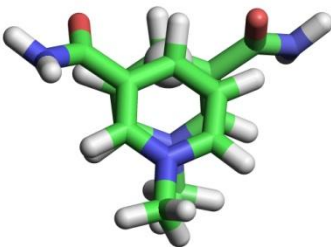
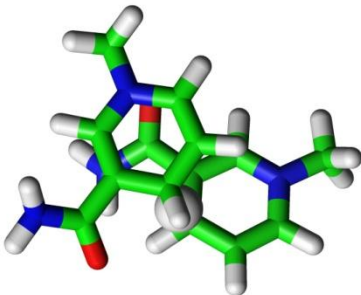
Rank	Designation	Structure	θ (Unopt.) / Degrees	θ (Opt.) / Degrees	Relative Energy / kJ mol^{-1}	TV Frequency (imaginary)
7	C2-EX-NO (+)		-179.98	+63.48	+10.89	807.4717
8	(E)-C2V-OO		-1.14	-15.86	+11.05	716.0715
9	C2-XE-ON (+)		-179.98	+79.16	+11.33	837.2079

Table 5.3 - Torsion Angles of Nicotinamide Hydride Transfer Transition States (Continued)

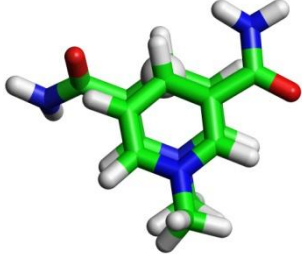
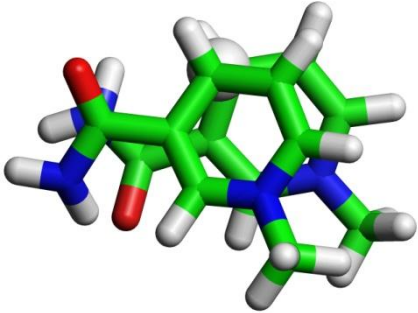
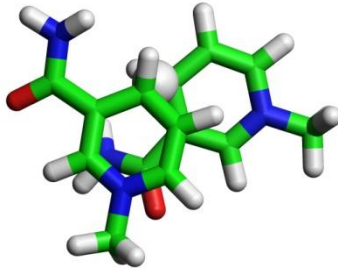
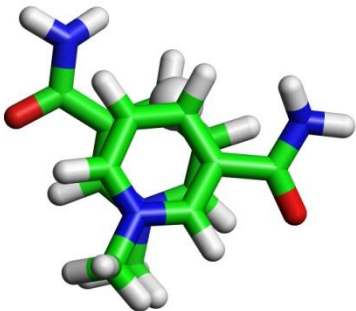
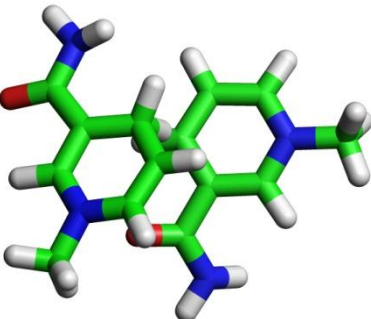
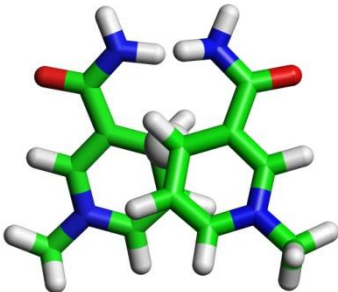
Rank	Designation	Structure	θ (Unopt.) / Degrees	θ (Opt.) / Degrees	Relative Energy / kJ mol^{-1}	TV Frequency (imaginary)
10	(E)-C2V-NO		+2.54	+1.13	+17.33	745.4013
11	(Z)-C2V-ON		-14.82	-14.82	+23.18	824.7683
12	C2-XE-NN (-)		-14.82	-78.74	+23.55	878.8370

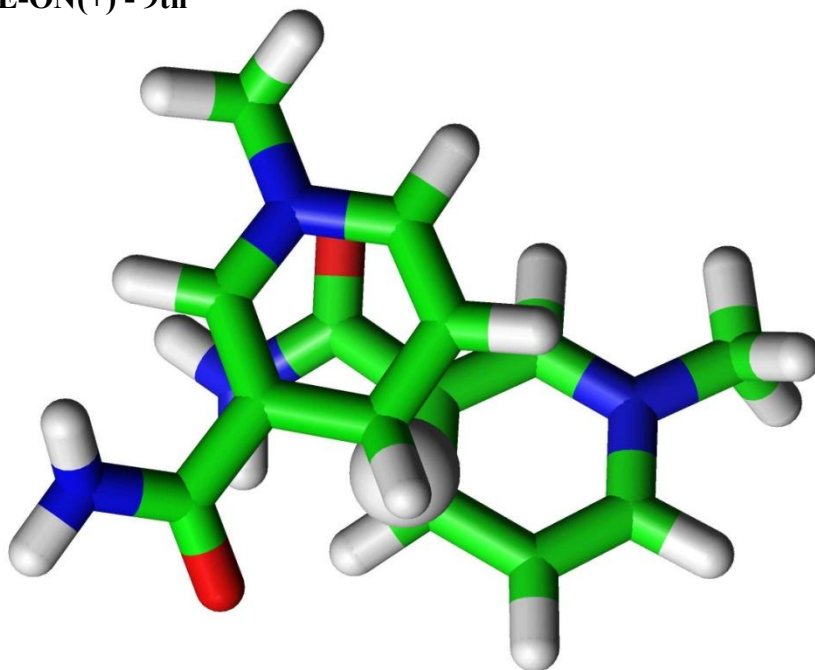
Table 5.3 - Torsion Angles of Nicotinamide Hydride Transfer Transition States (Continued)

Rank	Designation	Structure	θ (Unopt.) / Degrees	θ (Opt.) / Degrees	Relative Energy / kJ mol^{-1}	TV Frequency (imaginary)
13	(E)-C2V-NN		+2.56	-18.16	+24.52	874.4048
14	C2-XE-NO (-)		-179.98	-117.03	+26.02	588.8092
15	C2-XX-NN (+)		+100.80	+89.10	+31.17	820.1746

Mirror images exist because of the chirality of the C₂ ring orientation, whereby the torsion angle, θ , can either be positive (+) or negative (-), as illustrated through example in Figure 5.13 below.

Within these results, a corresponding mirror image TS structure was found for all but one of the unique transition state structures with a C₂ ring orientation, as detailed in Table S1. The exception being the highest energy C₂ TS structure, C2-XX-NN.

(a) C2-XE-ON(+) - 9th



(b) C2-XE-ON(-) - Mirror

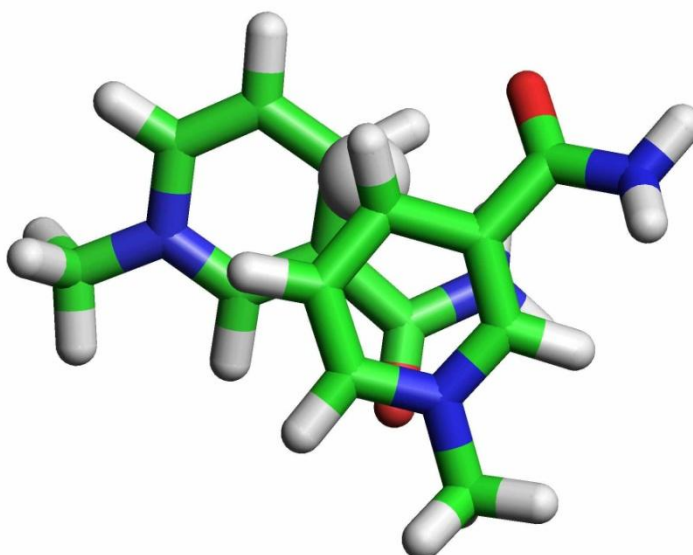


Fig. 5.13 An example of a C_2 transition state (a) C2-XE-ON(+) and its mirror image (b) C2-XE-ON(-) due to the chirality of the C_2 ring orientation.

In some cases, a clear change in ring symmetry was observed on optimisation, examples are shown in Figures 5.14 and 5.15 below. In these cases, optimisation was repeated with some variation in starting structure to verify that the change in ring geometry was valid and not an artefact of the specific optimisation.

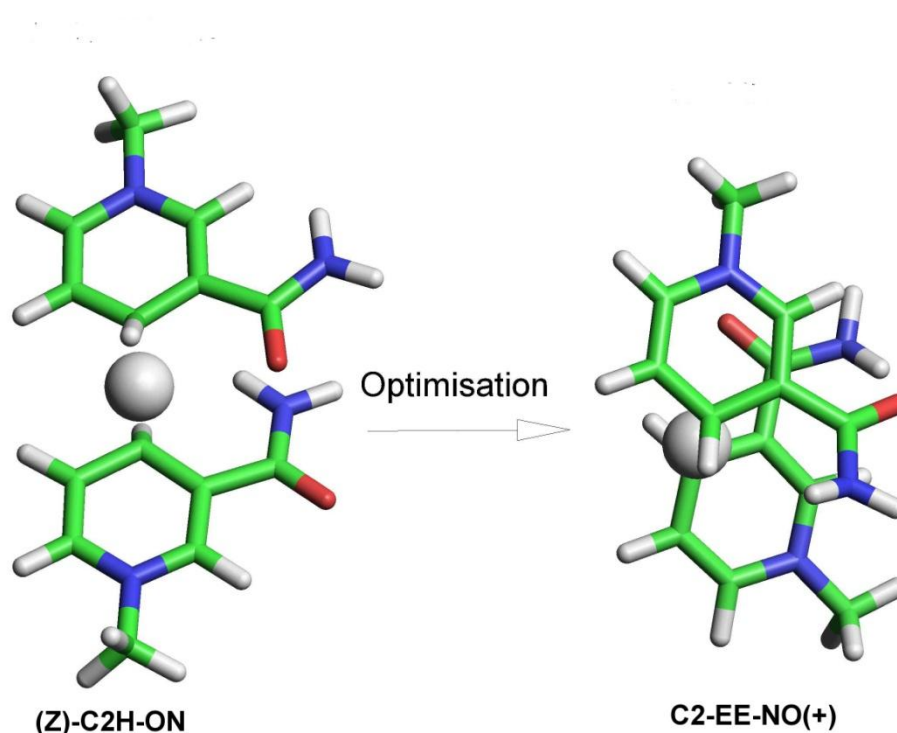


Fig. 5.14 Example of an initial C_{2h} ring geometry optimising to a final C_2 ring geometry.

/

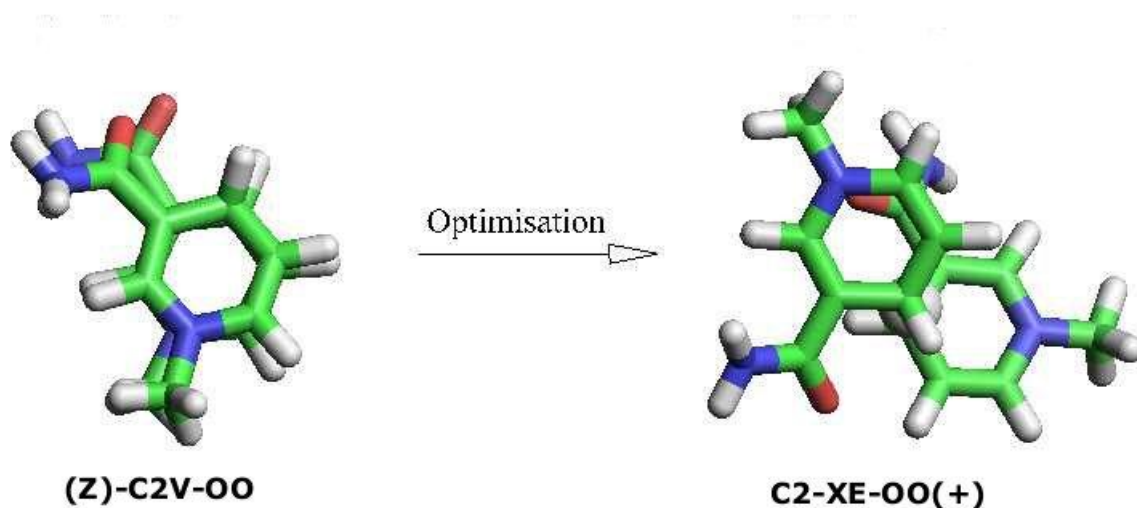


Fig. 5.15 Example of an initial C_{2v} ring geometry optimising to a final C_2 geometry.

Changes in ring geometry during structural optimisation of a transition state go some way in explaining why no first order TS structures with a C_{2h} ring geometry (Figure 5.15) and why so few first order TS structures with a C_{2v} ring geometry were isolated (Figure 5.14). In both of the cases illustrated above, the TS structure is optimised to give a first order TS with a C_2 ring orientation.

Though a true C_{2v} structure would have a torsion angle of 0° , the presence of amide groups introduces distortions so that any final optimised torsion angle between 0° and 20° was regarded as a TS structure with a C_{2v} ring geometry.

5.6 Reactants, Products and True Transition States

Intrinsic Reaction Coordinate (IRC) calculations were performed to verify that the optimised transition state structures are true transition states for hydride transfer. From a true transition state, reaction paths can be followed to determine which minima (reactant or product structures) are connected by a particular transition state.

Following a successful IRC calculation on a given transition state, an energy optimisation yields the reactant and product to which the transition state is connected.

A number of IRC calculations failed due to convergence failure. In these cases, as the IRC progresses, the interactions holding the molecules in close proximity decline rapidly allowing the molecules to drift apart and thus becoming prone to SCF failure.

Fortunately, the presence of duplicate transition state structures allowed several attempts at an IRC calculation so that reactant and products could be determined for most of the fifteen unique transition states and are shown in Table 5.4. All energies are reported relative to the lowest minimum energy structure that was isolated, in this case one connected to the C2-EE-NO transition state and has a total energy of -913.006399739 Hartree. IRC calculations on all 43 transition states that were found in total, including mirror images and duplicates, by this study are presented in Table S2 in the Appendix.

There is considerable range in the energy of these optimised reactants and products and it is clear from the structures that they bear little resemblance to the originating transition

state because once the C4-hydride interaction is lost for one of the nicotinamide rings, there is little holding the two rings in a particular orientation.

Table 5.4 Transition States and Global Minima

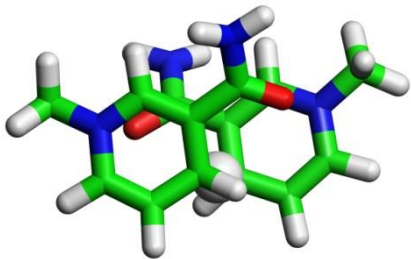
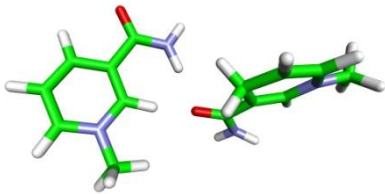
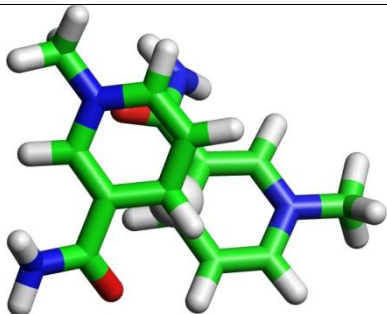
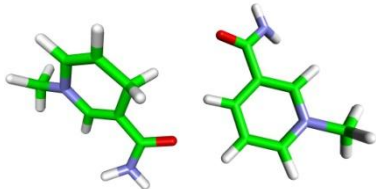
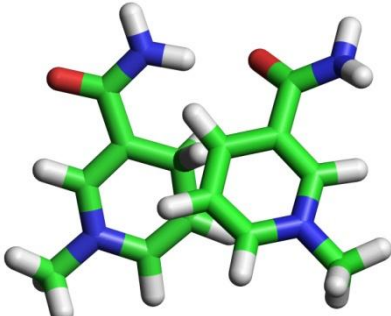
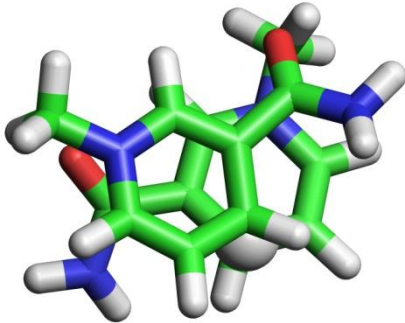
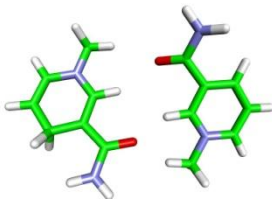
Transition State Structure	Reactant/Product Structure (Global Minimum)	Relative Energy / kJ mol^{-1}
 <p>C2-EE-OO(+)</p>		+9.43
 <p>C2-XE-OO(+)</p>		+52.14
 <p>C2-XX-ON(+)</p>	N/A	N/A
 <p>C2-EE-NN(+)</p>		+10.25

Table 5.4 Transition States and Global Minima (continued)

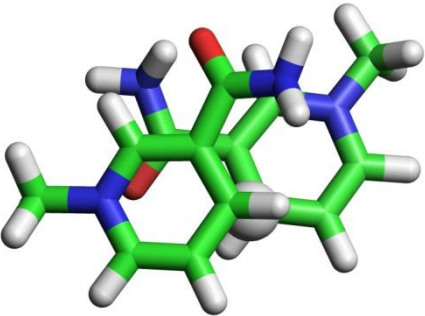
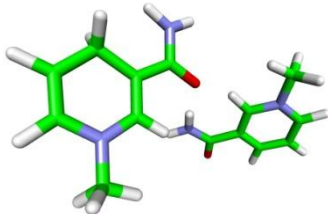
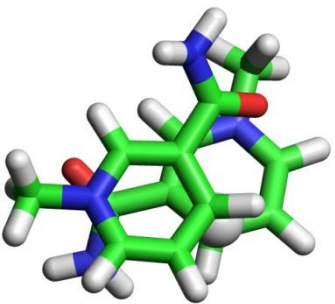
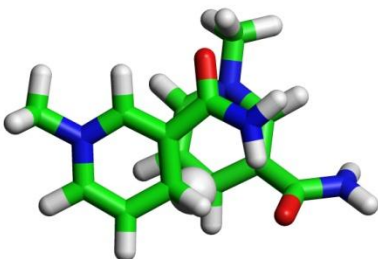
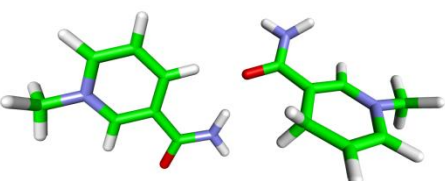
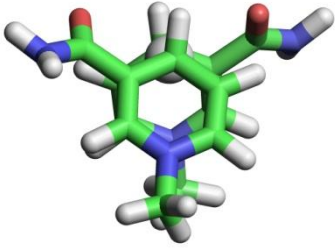
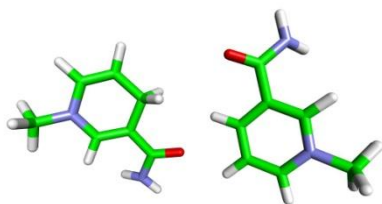
Transition State Structure	Reactant/Product Structure (Global Minimum)	Relative Energy / kJ mol^{-1}
 C2-EE-NO(+)	 C2-EE-NO(+)	0
 C2-EE-ON(+)	N/A	N/A
 C2-EX-NO(+)		+9.83
 (E)-C2V-OO		+52.13

Table 5.4 Transition States and Global Minima (continued)

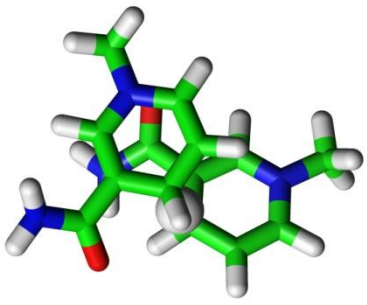
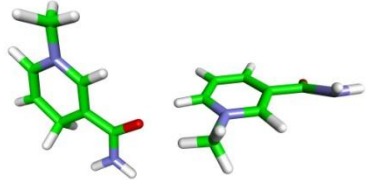
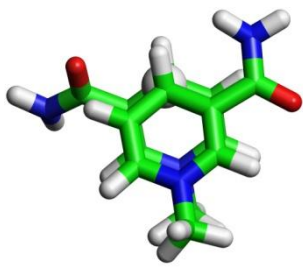
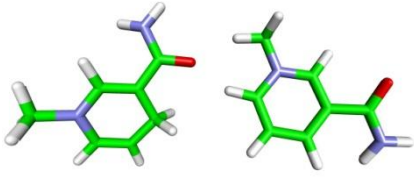
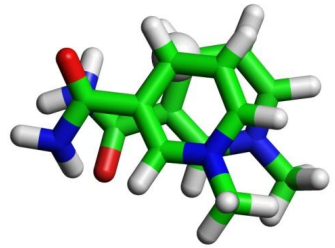
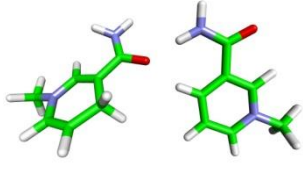
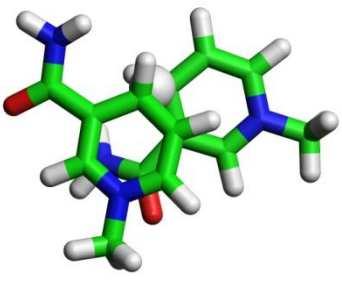
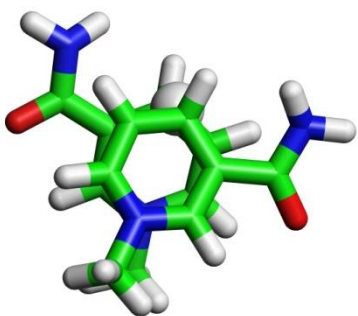
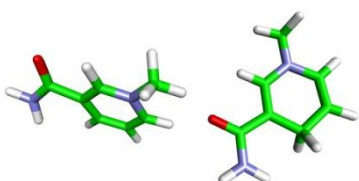
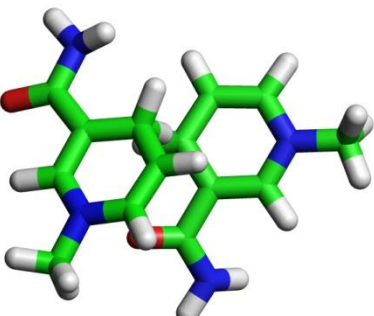
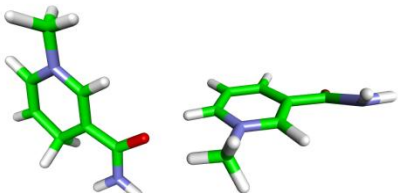
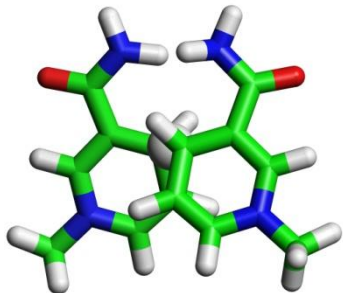
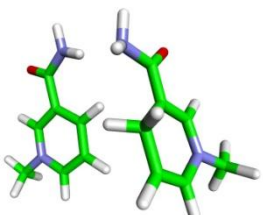
Transition State Structure	Reactant/Product Structure (Global Minimum)	Relative Energy / kJ mol^{-1}
 C2-XE-ON(+)		+22.02
 (E)-C2V-NO		+21.66
 (Z)-C2V-ON		+10.04
 C2-XE-NN(-)	N/A	N/A

Table 5.4 Transition States and Global Minima (continued)

Transition State Structure	Reactant/Product Structure (Global Minimum)	Relative Energy / kJ mol^{-1}
 (E)-C2V-NN		+14.09
 C2-XE-NO(-)		+0.017
 C2-XX-NN(+)		+52.42

5.7 Transition States in Solution

As full geometry optimisations of each of the 15 unique transition states using the PCM methodology to investigate the effects of solvation failed, single point PCM optimisations on the transition states were carried out to give an initial idea of the free energies of the transition states in solution (ΔG).

Little can be interpreted from these results apart from how far energetically these results may be from the fully optimised geometries in solution. The single point PCM total free energies in solution (ΔG of solvation) are reported below in Table 5.5 where the transition state structures are presented in their rank order from the gas-phase. All solvent-corrected energies are reported relative to the lowest energy transition state in solution, C2-XX-ON.

Table 5.5 Free Energies in Solution (ΔG) of Hydride Transfer Transition States (Single Point PCM)

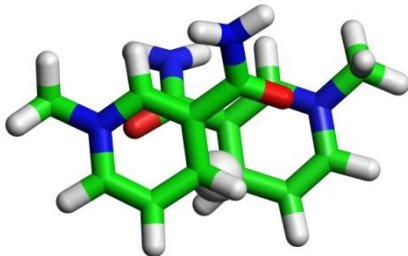
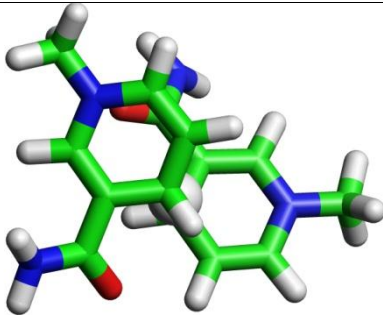
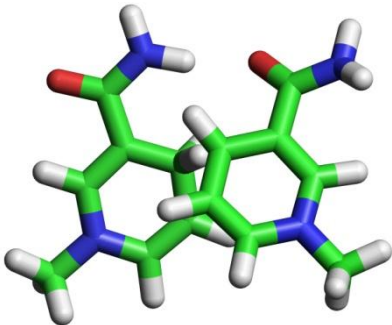
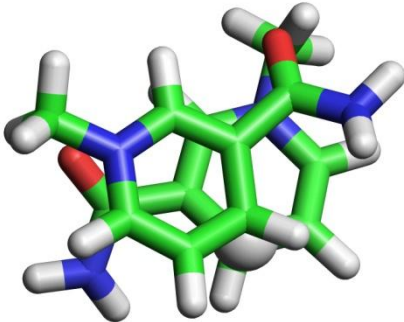
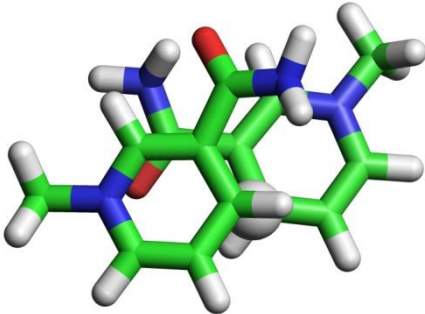
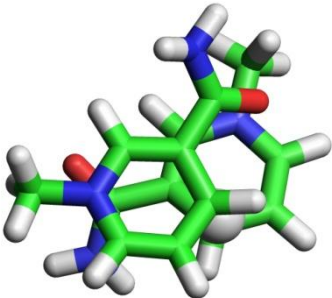
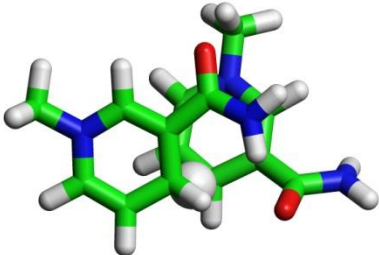
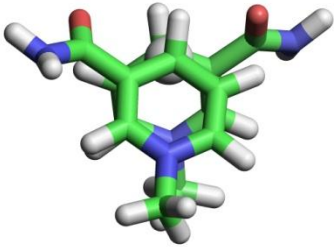
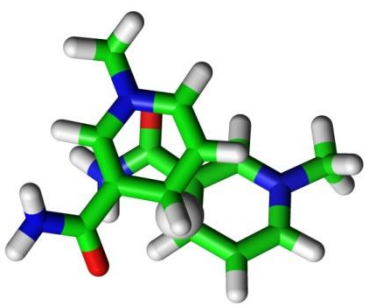
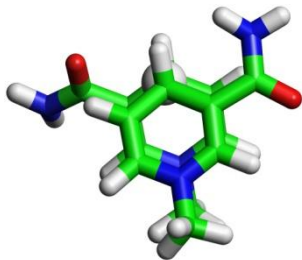
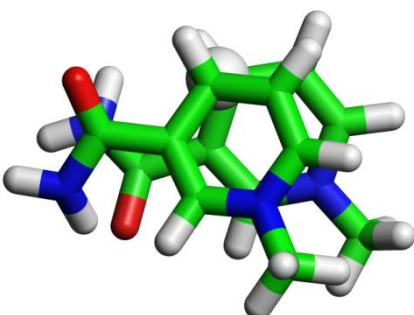
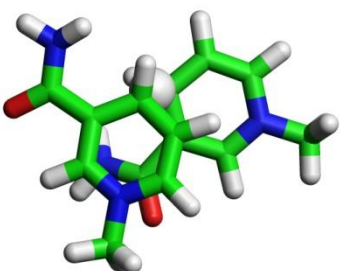
Designation	Transition State Structure	Relative Energy / kJ mol ⁻¹
C2-EE-OO (+)		+22.14
C2-XE-OO (+)		+23.50
C2-XX-ON (+)		0

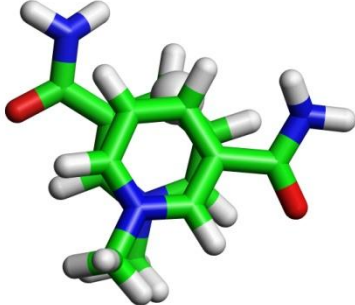
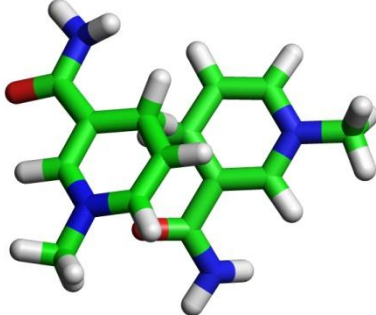
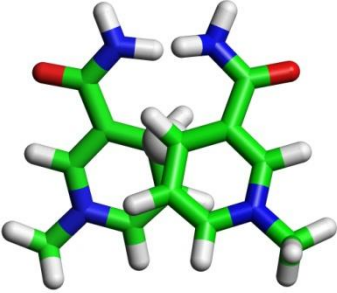
Table 5.5 Free Energies in Solution (ΔG) of Hydride Transfer Transition States (Single Point PCM)
[continued]

Designation	Transition State Structure	Relative Energy / kJ mol ⁻¹
C2-EE-NN (+)		+28.12
C2-EE-NO (+)		+22.51
C2-EE-ON (+)		+29.26
C2-EX-NO (+)		+26.29

**Table 5.5 Free Energies in Solution (ΔG) of Hydride Transfer Transition States (Single Point PCM)
[continued]**

Designation	Transition State Structure	Relative Energy / kJ mol ⁻¹
(E)-C2V-OO		+21.95
C2-XE-ON (+)		+26.49
(E)-C2V-NO		+19.47
(Z)-C2V-ON		+22.16
C2-XE-NN (-)		+26.96

**Table 5.5 Free Energies in Solution (ΔG) of Hydride Transfer Transition States (Single Point PCM)
[continued]**

Designation	Transition State Structure	Relative Energy / kJ mol ⁻¹
(E)-C2V-NN		+23.94
C2-XE-NO (-)		+25.32
C2-XX-NN (+)		+22.69

Perhaps the most striking thing about these results is that the total free energies in solution (ΔG) for each of the transition states are all remarkably similar, with all of them falling in the range of $\sim +20$ - 30 kJ mol^{-1} relative to the lowest energy structure in solution, C2-XX-ON, which is the most asymmetric transition state that was isolated and which was energetically ranked 3rd overall in the gas phase results reported in Table 5.1. Here this structure has a solvent corrected energy of -914.049443 Hartree.

Clearly, further future work is needed in order to overcome the geometry optimisation problems experienced with the PCM methodology to obtain more meaningful and accurate free energies that would be more representative of the true free energies of these TS structures fully optimised in solution as opposed to these partially optimised single point free energies which only give a partial idea of the effect of solvation on the energies of the isolated transition states, so nothing definite can be inferred from these results.

5.8 Implications for Hydride Transfer

The biological implications of all of the research discussed in this chapter for hydride transfer in the transhydrogenase enzyme primarily support and provide further evidence for the earlier findings of Bhakta *et al.* [6]. There is strong evidence for the asymmetry of the transition state as a driving force for hydride transfer in the enzyme provided by the fact that the results of this transition search coincide perfectly with the search conducted in 2007. The lowest energy gas-phase transition state isolated in both cases was the fully symmetric C2-EE-OO (Figure 5.7), which is not observed in the enzyme.

The nicotinamide ring conformation which resembles the one observed in the enzyme the most is a C_2 structure with one amide '*endo*', and the other '*exo*' with respect to the other

nicotinamide ring (Figure 5.8), and such a configuration is also observed to be asymmetric for hydride transfer and thus has led to the proposal of the asymmetric hypothesis which postulates that an asymmetric transition state gives rise to reactants and products with different energies in spite of being chemically identical if retained in a conformation that relates to the transition state, and that this sufficiently explains the change in reactant/product ratio brought about by transhydrogenase. A more reactant-like transition state should indicate an exothermic reaction.

Two transition states have been identified that could correspond to the conformation observed in the enzyme. Although the crystal structure (PDB: 2OO5) [6] discussed here offers no clues to the amide orientation of NAD(H) in *R. rubrum* transhydrogenase, owing to the fact protons do not appear and oxygen and nitrogen atoms are indistinguishable in the crystal structure (normally the positions of these atoms can be inferred from hydrogen bonding pattern, but the hydrogen bonding here is with an OH group and so this is not possible), there are a few theoretical studies [31, 93-95] with other transhydrogenases that suggest that this amide orientation may in fact be in the highly unusual N orientation. Singh *et al.* [93] provide supposition for this N orientation of NAD(H) in dI from the fact that in their work thio-NAD(H) is shown to restrict conversion from a distal to a proximal conformation in the enzyme [93]. They go on to state how the crystals of the lithium salt of NAD⁺ dihydrate (and related nucleotides) are shown to have a N orientation of the amide group. This amide orientation of NAD(H) is contrary to a majority of crystal structures with bound nucleotides which suppose that the amide orientation must be O in NAD(H) based primarily on the hydrogen bonding pattern of donors and acceptors [93]. It is even more striking an observation when you consider that the hydride transfer transition state calculations on the model systems discussed at length in this chapter demonstrate quite

clearly that the C2-XE-OO conformation with both of its amides in a O orientation is lower in energy than the alternative C2-XE-ON conformation with one of its amides in a N orientation. These seemingly contradictory suppositions has lead to the proposal by Singh *et al.* [93] that the C3-C7 bond of NAD(H) rotates to the preferred O orientation of the carbonyl oxygen during turnover, when the enzyme undergoes a conformational change whilst bringing the nicotinamide rings into a reacting conformation (proximal state) and therefore facilitating hydride transfer from a distal state where the two rings are held apart and hydride transfer is no longer possible [93]. This state corresponds to the enzyme configuration either prior to or just after hydride transfer (Figure 5.16).

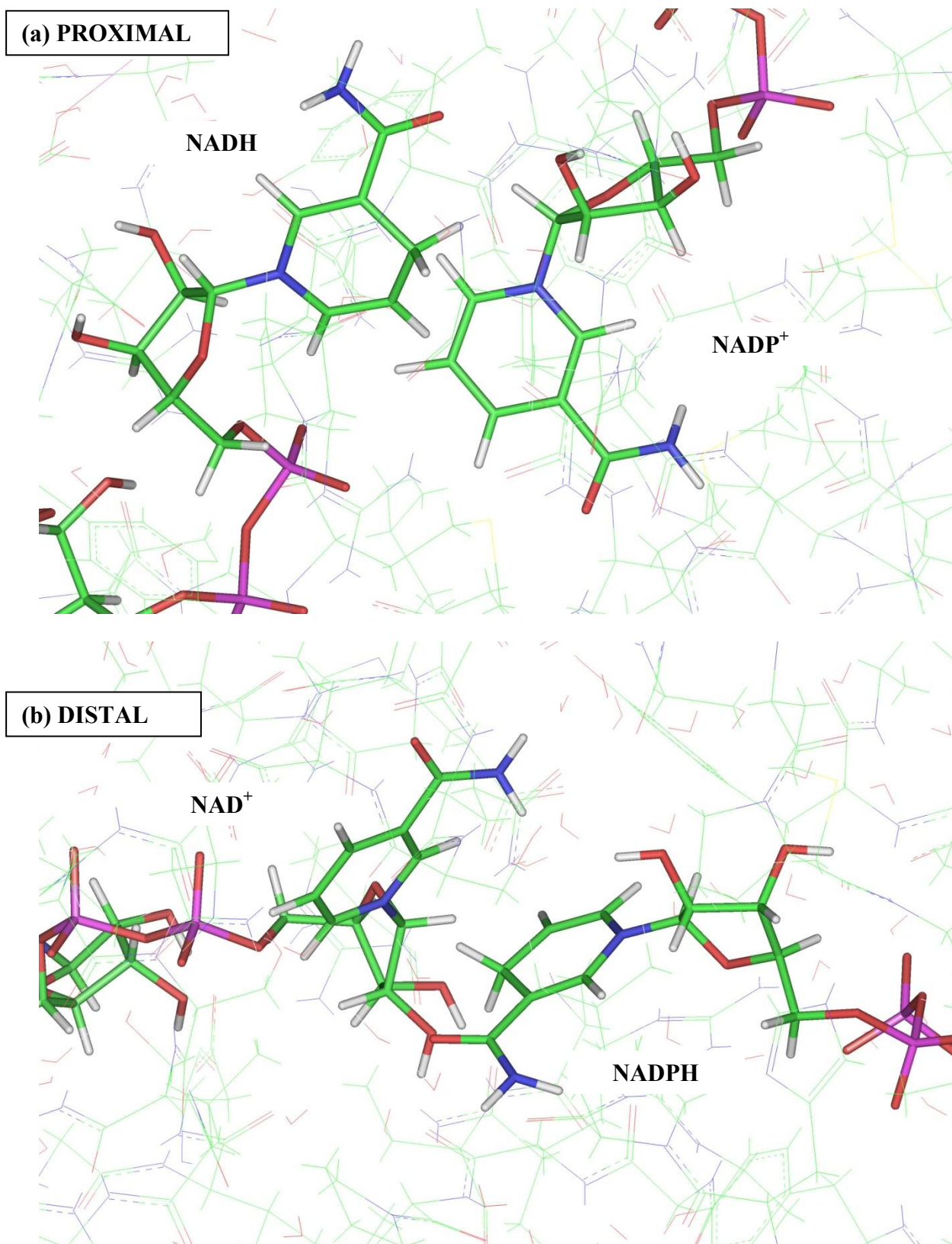


Fig. 5.16 The two conformational states in the enzyme, the nicotinamide rings of NAD(H) and NADP(H) are either (a) held close together in a proximal conformation, allowing hydride transfer, or (b) held apart in a distal conformation, blocking hydride transfer.

Previous studies [31, 95] have also shown that for nicotinamide and 1,4-dihydronicotinamide the amide rotation barrier is energetically only a few kcal mol⁻¹ and therefore it can be theorized that the rotational switch is an important additional preventative measure in blocking hydride transfer in the enzyme until the rings are properly orientated to facilitate hydride transfer, minimising slippage in proton translocation coupling. The amide bond rotation may also have a role in facilitating hydride transfer by being influential in the positioning of the (dihydro-) nicotinamide ring between the hydrogen bonding arrangements of the distal and proximal configurations of the enzyme active site [93].

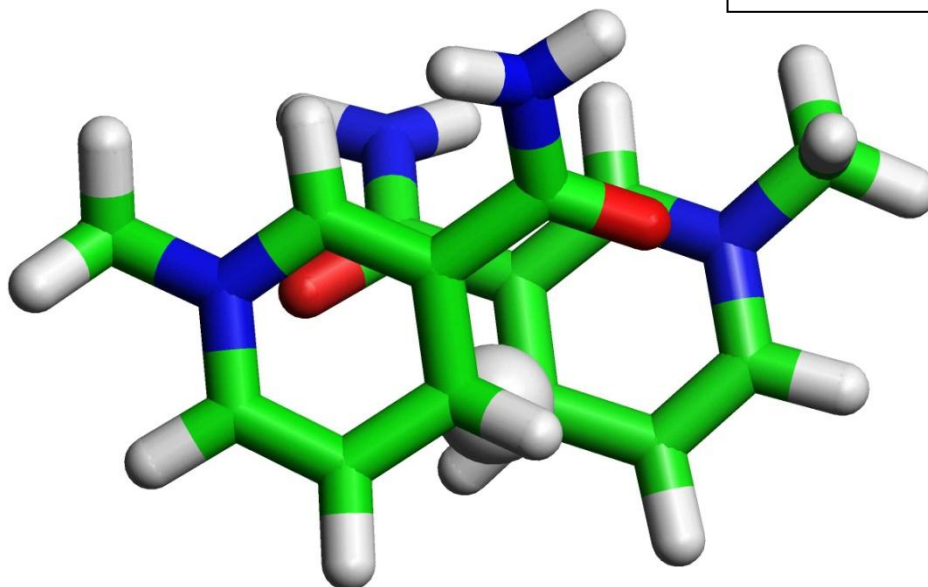
The work of Bhakta *et al.* [6] and Singh *et al.* [93] both support the observation that for NADPH, the amide is indeed in an O orientation, but Singh *et al.* proceed also to anticipate other conformational changes in the dIII binding site during enzyme turnover, these changes are not expected to be simply equivalent nucleotide amide rotations in NADP(H) to those hypothesised to occur in the dI binding site for NAD(H) [93].

The transition state structure, C2-XE-OO(+), shown in Figure 5.17 (b) most closely resembles the arrangement of the nicotinamide rings observed in the X-ray crystal structure of transhydrogenase. In this transition state and in the crystal structure, the ring planes are almost parallel and the torsion angle between the N1-C4 axes of both rings is 120.67° when seen through a normal to the rings plane as in Figure 5.17 (b). The C5-C6 bond of the top ring overlaps and stacks upon the C3-C7 bond of the amide group of the bottom ring, while the top amide group is orientated away from the bottom ring, thus defining the '*exo/endo*' configuration of the structure.

(a) C2-EE-OO(+) - 1st

Fully Symmetric Transition State (both C-H bond distances 1.34 Å)

Lowest energy and **not** observed in enzyme.



(b) C2-XE-OO(+) - 2nd

Asymmetric Transition State (both C-H bond distances 1.29 and 1.38 Å respectively).

Second lowest energy (+6.29 kJ mol⁻¹) and most closely resembles configuration observed in enzyme.

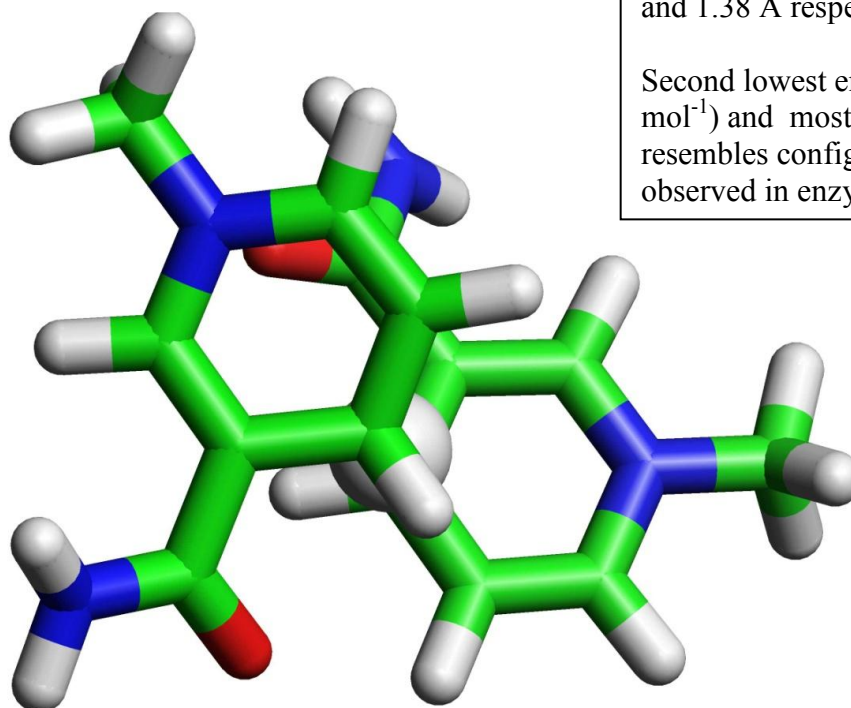
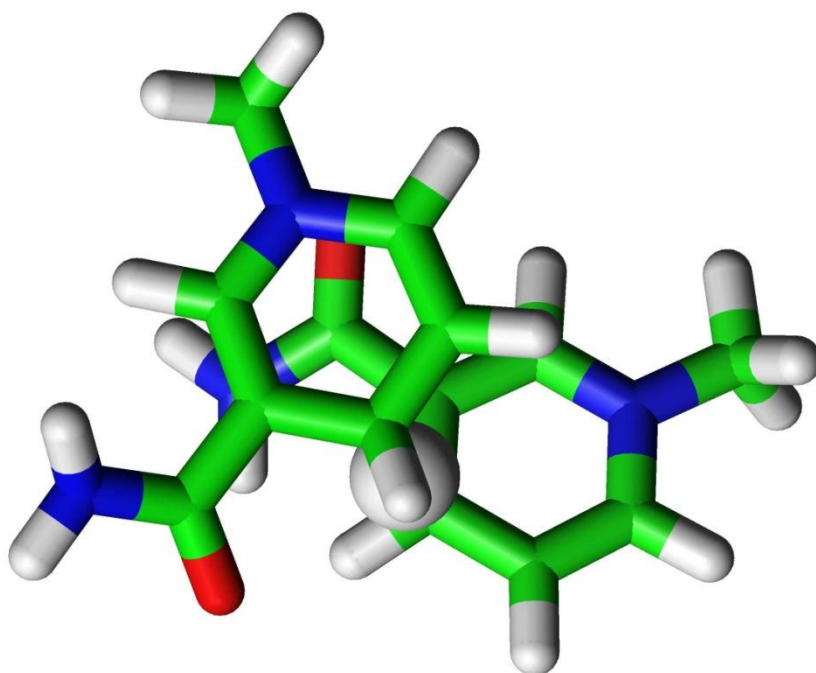


Fig. 5.17 The two lowest energy transition states, C2-EE-OO(+) and C2-XE-OO(+), (a) is fully symmetric and not observed in the X-ray crystal structure, whilst (b) is asymmetric and most closely resembles the arrangement of the rings observed in the crystal structure

In the C2-XE-OO transition state structure, the two C4 atoms are 2.67 Å apart, and both show partial sp^3 hybridization as expected for hydride transfer. The two amide groups in this TS are both orientated O in each nicotinamide ring. In comparison to the lowest energy TS structure, C2-EE-OO(+), shown in Figure 5.17 (a), which is fully symmetric with both C-H bond distances for hydride transfer being 1.34 Å, C2-XE-OO(+) is asymmetric, having C-H bond distances for hydride transfer being 1.29 Å and 1.38 Å respectively. The hydride transfer reaction is essentially a symmetrical reaction, so for it to proceed through an asymmetric transition state is highly unusual, and typically indicative of a two-step mechanism that proceeds through a symmetrical intermediate [6]. Owing to the fact that both the IRC and frequency calculations confirm the nature of this structure as a first-order transition state, there is proof that the reaction does not progress through a two-step mechanism and a symmetrical intermediate, and that hydride transfer instead proceeds via a single step.

(a) C2-XE-ON(+) - 9th



(b) C2-XE-NO(-) - 14th

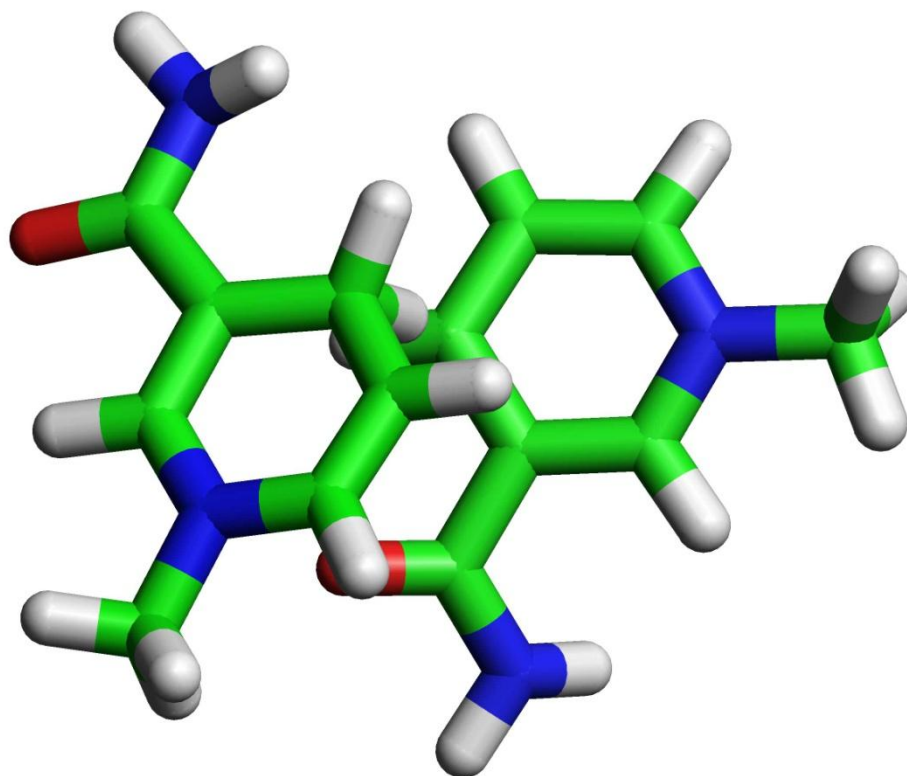
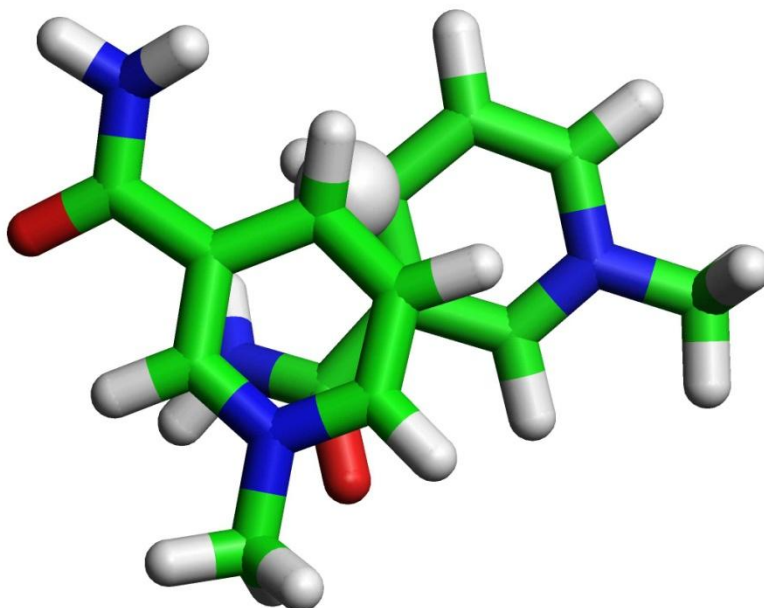


Fig. 5.18 Two transition states which are similar to C2-XE-OO:

(a) C2-XE-ON and (b) C2-XE-NO

Transition state structures C2-XE-ON (+11.33 kJ mol⁻¹) and C2-XE-NO (+26.02 kJ mol⁻¹), shown in Figure 5.18 are structurally similar to C2-XE-OO (+6.29 kJ mol⁻¹) except that one of the amide groups is rotated N in its nicotinamide ring, this small amide rotation leads to an increase in energy of approximately 5 kJ mol⁻¹ for C2-XE-ON and approximately 20 kJ mol⁻¹ for C2-XE-NO when compared to the energy of C2-XE-OO. This observation is in good *agreement* with previous results in [6]. Once again with these asymmetric C₂ transition states as seen with the two symmetric C_{2v} transition states discussed earlier a simple amide rotation results in an almost four fold increase in the energy of the transition state. Another structurally similar transition state, C2-XE-NN (+23.55 kJ mol⁻¹), shown in Figure 5.19 (a), has both amide groups rotated N and has an energy increase of approximately 17.3 kJ mol⁻¹ when compared to the energy of C2-XE-OO. The crystal structure resolution cannot directly identify the amide rotation of the C3-C7 bond in NAD(H), leading to some ambiguity owing to the fact that hydrogen atoms do not appear and both oxygen and nitrogen are indistinguishable in the crystal structure. The hydrogen bonding pattern suggests that the amide of NADP(H) has an O rotation [6], thus the ring arrangement in the crystal structure could match either transition states C2-XE-OO or C2-XE-ON (Figure 5.20) but not transition states C2-XE-NN or C2-XE-NO (Figure 5.19).

(a) C2-XE-NN(-) - 12th



(b) C2-XE-NO(-) - 14th

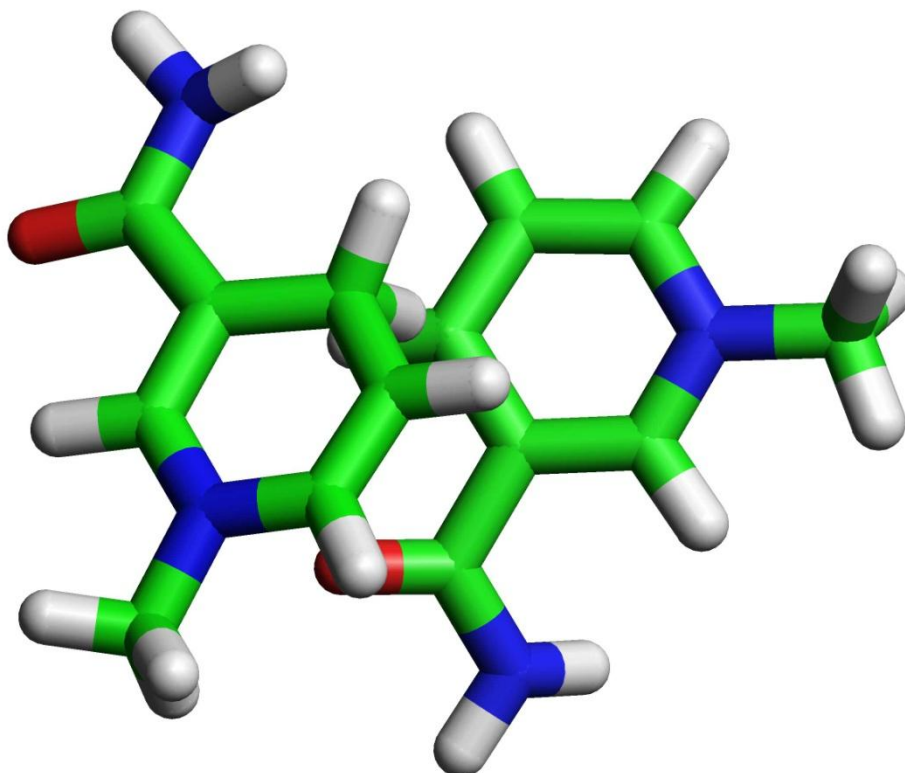
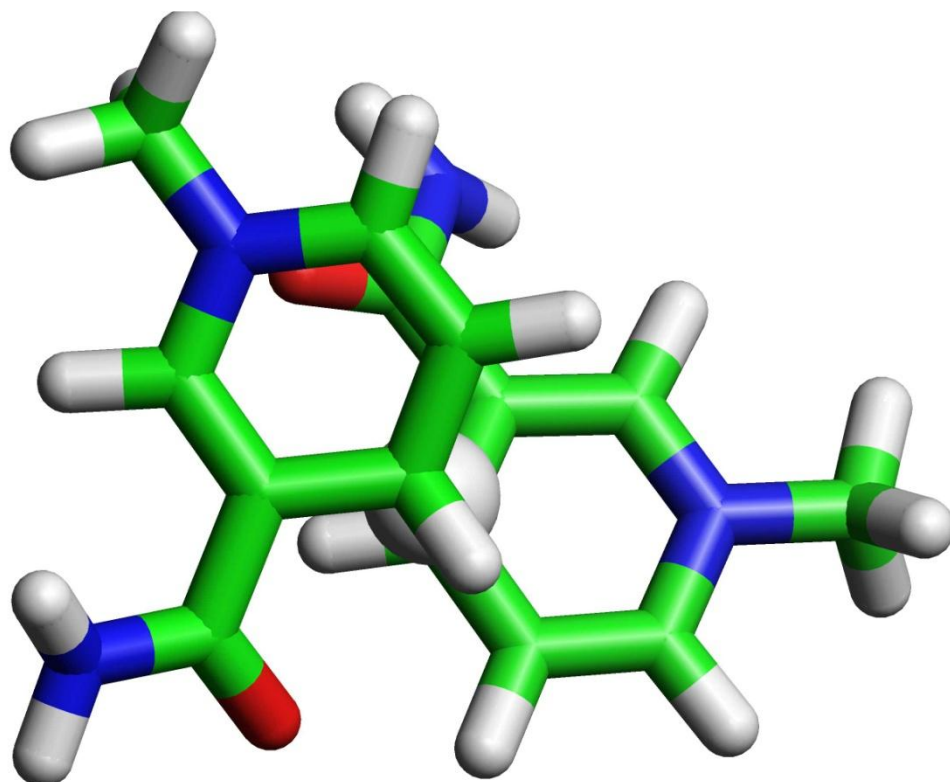


Fig. 5.19 Two high energy transition states which differ only in amide orientation:

(a) C2-XE-NN and (b) C2-XE-NO

(a) C2-XE-OO(+) - 2nd



(b) C2-XE-ON(+) - 9th

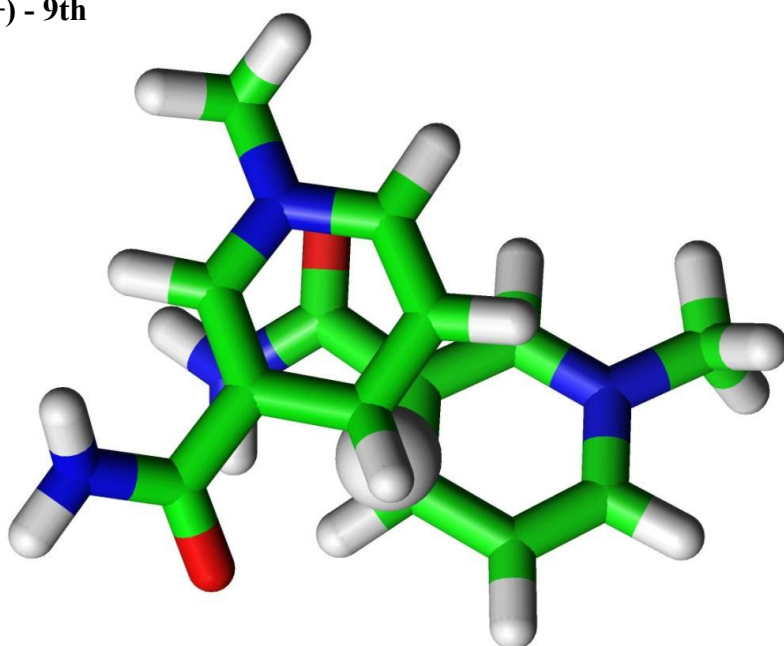


Fig. 5.20 Ambiguity in the amide orientation of one of the nucleotides in the crystal structure of transhydrogenase means that the corresponding ring arrangement could resemble either of these transition states: (a) C2-XE-OO or (b) C2-XE-ON

Figures 5.19 and 5.20 show all four possible amide orientations side by side for the hydride transfer transition state between 1,4-dihydronicotinamide and protonated nicotinamide when the ring arrangement is C₂ and the amide positions are *exo* and *endo* respectively. A number of interesting observations can be made in regards to these four structures that closely resemble the ring arrangement found in the X-ray crystal structure of transhydrogenase. C2-XE-NN and C2-XE-NO (Figure 5.19) are significantly higher in energy and each have a negative torsion angle compared to C2-XE-OO and C2-XE-ON (Figure 5.20) which each have a positive torsion angle. All four of these transition states are perhaps unsurprisingly asymmetric, but what is most striking is just how much of an impact simple changes in amide orientation seem to have on the energy of the transition state.

The next chapter concerns taking each of the nicotinamide ring configuration possibilities which correspond most closely to the ring configuration found in the crystal structure, C2-XE-OO and C2-XE-ON, and modelling each of them within the *R. rubrum* transhydrogenase enzyme as reactants and products and running ONIOM simulations on these models to probe the thermodynamics of the hydride transfer reaction by looking at how ring geometry and binding interactions of the surrounding protein environment affect the reaction, and whether the effects of geometry alone are sufficient to impact the reaction thermodynamics and explain the raised equilibrium constant. By accounting for both conformational possibilities, it is hoped that we can also resolve computationally something which has proven to be crystallographically ambiguous - the amide rotation of NAD(H) [6].

CHAPTER SIX

REACTION THERMODYNAMICS: GEOMETRIC AND ELECTROSTATIC EFFECTS UPON HYDRIDE TRANSFER

This chapter comprehensively covers the hybrid QM/MM (ONIOM) study of *R. rubrum* transhydrogenase to investigate the hydride transfer reaction and what effect the electrostatics of the surrounding protein environment may have on hydride transfer itself.

6. Reaction Thermodynamics Geometric and Electrostatic Effects Upon Hydride Transfer

This chapter concerns taking each of the nicotinamide ring configuration possibilities which correspond most closely to the ring configuration found in the crystal structure, C2-XE-OO and C2-XE-ON, and modelling each of them within the *R. rubrum* transhydrogenase enzyme as reactants and products and running ONIOM simulations on these models to probe the thermodynamics of the hydride transfer reaction by looking at how ring geometry and binding interactions of the surrounding protein environment affect the reaction, and whether the effects of geometry alone are sufficient to impact the reaction thermodynamics and explain the raised equilibrium constant. By accounting for both conformational possibilities, it is hoped that we can also resolve computationally something which has proven to be crystallographically ambiguous - the amide rotation of NAD(H) found in the enzyme [6].

As elaborated upon in Section 3.5 and 3.6, the ONIOM methodology is a hybrid QM/MM approach that allows different levels of theory to be applied to different layers in a molecular system. In the model systems investigated here, the nicotinamide rings of both cofactors are in the quantum mechanical (QM) high layer, whilst residues within 10 Å around them are in the molecular mechanical (MM) low layer (Figure 6.1). The QM region is flexible, whilst the MM region can be fully rigid or partially flexible (residues within 5 Å of the QM region) and also charged or uncharged, thus allowing the role of the protein as a simple steric block to be investigated. The influences of simple changes in geometry on the enthalpy of the hydride transfer reaction are compared to the influences of electrostatics on the protein environment in this study with a focus on the dIdIII dimeric model systems.

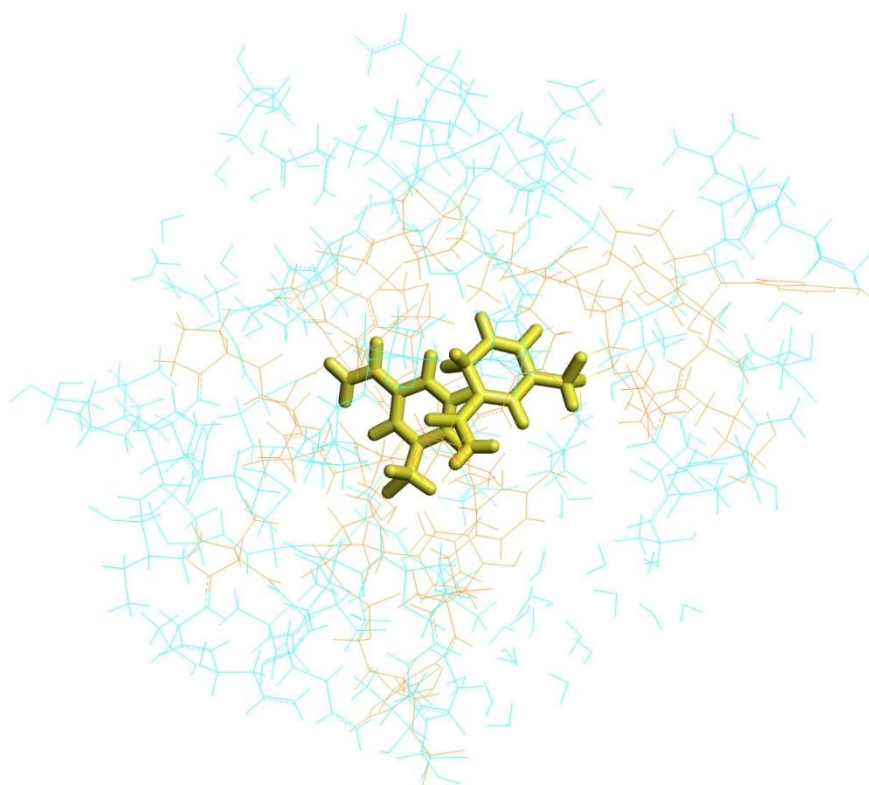


Fig. 6.1 A representation of ONIOM layers. The nicotinamide rings in the QM layer are shown in yellow, the flexible part of the MM (residues within 5 Å of the QM layer) is shown in orange, and the outer rigid part of the MM layer is shown in blue.

Before discussing the results of the ONIOM simulations, it is important to note here that the molecular dynamics simulations with the models of *R. rubrum* transhydrogenase discussed here implemented the protein and nucleic acid parameters from the AMBER ff03 set, and so it is consistent to use the AMBER force field in these hybrid QM/MM calculations also, with additional NAD(H) and NADP(H) cofactor parameters [68-70] specified as outlined in Section 3.5. More practical reasons for the use of the AMBER force field in the ONIOM study of the hydride transfer reaction in the enzyme is the ease with which the partial charges and AMBER atom types necessary in the input files can be incorporated. Once again, it is important to reiterate here that, as mentioned in Chapter 3,

all the necessary input files for these types of calculations were all generated through the implementation of entirely new programs written with Fortran 90. After the ONIOM calculations, Open Babel [97] was used to convert the Gaussian output file into PDB structure files for visualisation.

6.1 ONIOM Simulations on Model Protein Systems

As outlined above, protein models of *R. rubrum* can be fully rigid or partially flexible and the MM layer of each model can be with or without partial charges on the amino acid residues, allowing the effects of simple geometry and binding interactions in the surrounding protein environment on reaction thermodynamics to be investigated. Each of these model variations are presented and discussed below for the dIdIII dimers. The effect of amide rotation within NAD(H) on hydride transfer is also discussed later for both the dimeric model systems.

Each of the dimeric protein models discussed herein were set up for MD simulations using Sander within Amber 8. Three MD simulations were run for each enzyme complex, whether it is a reactant or product state and which NAD(H) orientation is under consideration in each complex. Each 500 ps snapshot was taken and minimised to generate a starting structure for ONIOM calculations. The three MD runs are plotted individually in each of the graphs discussed from this point forward, in black, green and red respectively. Multiple MD runs on each model system demonstrate good repeatability and assess the reliability of each run, as similar behaviour and trends are expected to be reflected across each of the three runs to validate the results.

6.2 ONIOM Study of dIdIII Dimers

The dI₂dIII trimer (PDB: 2OO5) contains one complete active site (dimer) which is proximal in nature. In order to model a distal dimer, the tetramer needs to be built as described in Section 3.3-4. From Chapter 3, it is worth recalling the structural configurations of the dimers. Table 6.1 below summarises the dimer configurations that were modelled in this work for each of the two structural possibilities arising from the ambiguity of the amide rotation of NAD(H) in *R. rubrum* transhydrogenase crystal structure.

Binding Site Conformation	Binding Domains	Bound Cofactors
Distal	dI(A), dIII	NADH, NADP ⁺
Distal	dI (A), dIII	NAD ⁺ , NADPH
Proximal	dI (B), dIII	NADH, NADP ⁺
Proximal	dI (B), dIII	NAD ⁺ , NADPH

Table 6.1 The modelled single dIdIII dimer configurations

In this ONIOM study, the energies of the snapshots from the minimisation of the initial dynamics steps of the molecular dynamics simulations are typically observed to be significantly large for the dimeric systems and as such have been discounted from the results for each of the conformations discussed below, but are included in Table S3 and Table S4 in the Appendix for completeness. It is reasonable to discount these ONIOM

energies because these snapshots are prior to equilibration and as such result in high unnatural variations in energy.

6.2-1 XE-OO Conformations

Since the C2-XE-OO transition state discussed in the previous chapter is lower in energy than C2-XE-ON and that the *syn* preference for hydride transfer as observed by Wu *et al.* [84] is reflected in this transition state, the XE-OO conformation is expected to be the one likely to be found in the enzyme and hence yield the lowest and more reasonable reaction enthalpies compared to the XE-ON conformation. Furthermore, if the initial hypothesis that underpins the motivation for this research is correct whereby it is postulated that the asymmetry of the transition state is responsible for the raised equilibrium constant, the non-electrostatic results which only account for changes in geometry and the surrounding protein environment acts simply as a steric block should show a significant impact on the thermodynamics of hydride transfer as the reaction progresses. If, as postulated in the asymmetric hypothesis, the modified transition state geometry is sufficient to impact the equilibrium constant, the model systems should show a similar modification in the reaction thermodynamics, both with and without electrostatics. If however, the binding-change mechanism (see Chapter 1, Section 1.6), there should be a substantial further modification to the thermodynamics upon inclusion of electrostatics.

Figure 6.2 below shows a couple of ONIOM snapshots of the two states of the hydride transfer site of *R. rubrum* transhydrogenase with the XE-OO conformations in the enzyme. Figure 6.2 (a) shows the proximal state of the enzyme where the two nicotinamide rings are held close enough together to facilitate hydride transfer, and Figure 6.2 (b) shows the

distal state of the enzyme, this conformation is one in which hydride transfer is blocked and is representative of the state prior to or following hydride transfer. The proximal snapshot has the reactant bound and therefore represents the conformation before hydride transfer, the C4-C4 atom distance is 3.81 Å. The distal snapshot has the product bound and therefore represents the conformation after hydride transfer, the C4-C4 atom distance is 5.58 Å. The C4-C4 atom distance is considerably larger in the reactant and product states of the enzyme, compared to the transition state where it is only 2.67 Å, due to the loss of a substantial 'bond' between the transferring hydride and the two C4 atoms, allowing van der Waals repulsion to push the unbound C4 atom away from the hydride.

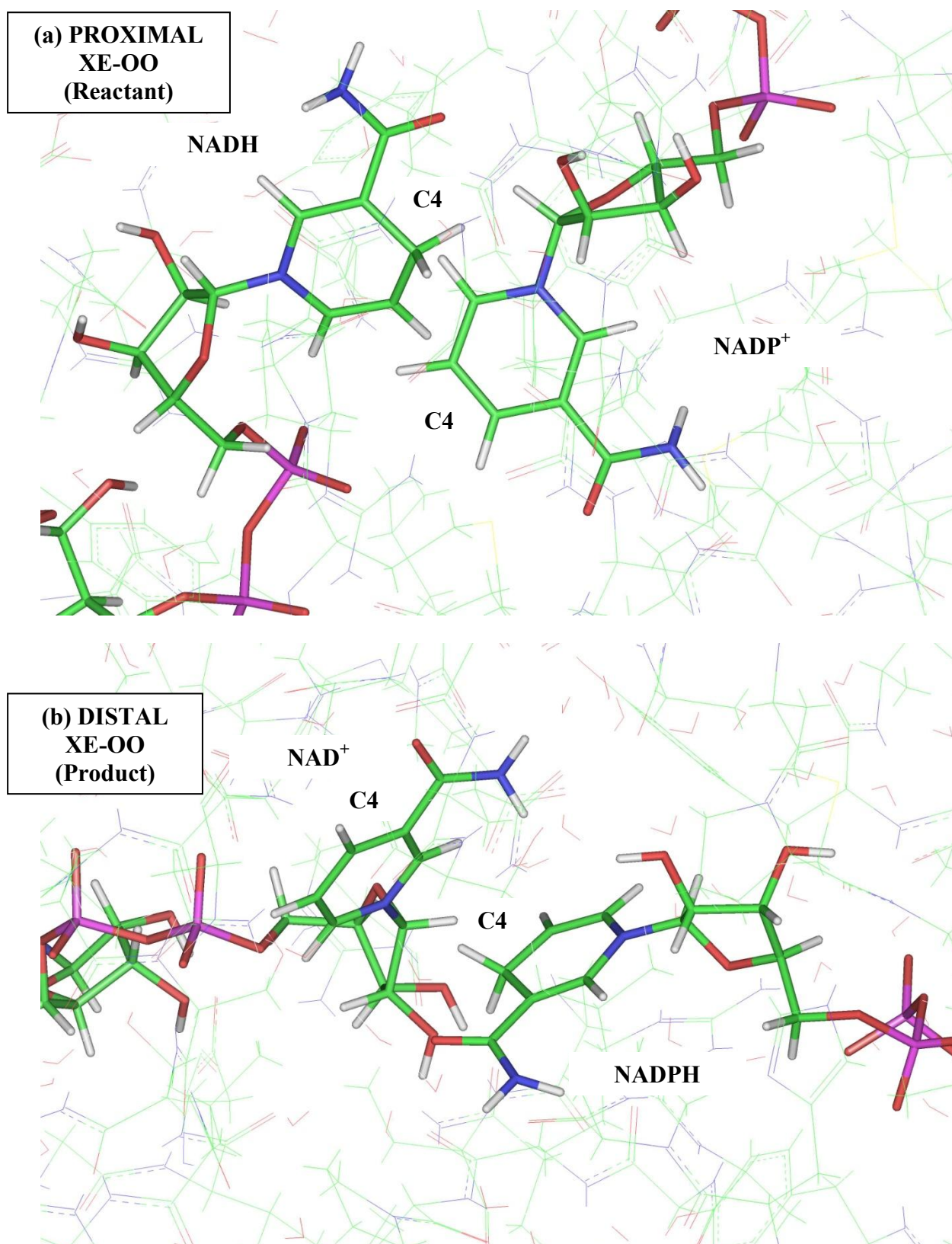


Fig. 6.2 The two XE-OO conformational states in the enzyme, the nicotinamide rings of NAD(H) and NADP(H) are either (a) held close together in a proximal reactant-bound conformation, prior to hydride transfer, or (b) held apart in a distal product-bound conformation, after hydride transfer.

In order to discuss the hydride transfer reaction in *R. rubrum* transhydrogenase, the results from the ONIOM simulations of the XE-OO protein models are presented below and discussed in the context of following the order of reaction in the enzyme - i.e. distal reactant bound, proximal reactant bound, proximal product bound and distal product bound states of the enzyme. The complete enzyme-catalysed reaction passes through all four of these states sequentially, whereas the chemical step of the hydride transfer reaction involves just the two proximal states.

Figures 6.3-6.6, below, show the energies of each of these four states across the three molecular dynamics runs plotted for the XE-OO conformations in the fully rigid protein models. Only the QM region (yellow in Figure 6.1) is included in the energy values given for these rigid systems.

These rigid model systems provide an indication of the thermodynamic changes that occur as the reaction progresses, but the partially flexible models, which include the QM region and residues within 5 Å of this (yellow and orange regions in Figure 6.1) as discussed later give a more accurate picture of these changes providing a more accurate simulation of the biological system as natural protein flexibility is accounted for in these models.

Figures 6.3 and 6.4 show the sole effect of geometry on the thermodynamics of the fully rigid model systems whereas Figures 6.5 and 6.6 incorporate the effects of binding interactions (electrostatics) into each of the rigid models to determine their influence on the thermodynamics of the reaction.

Rigid XE-OO Conformations		
Conformational State	Average Energy / Hartree	Relative Energy / kJ mol⁻¹
Reactant (Distal)	-913.195608686043 (\pm 0.10)	0
Reactant (Proximal)	-913.184229451752 (\pm 0.12)	+29.88
Product (Proximal)	-913.123779946750 (\pm 0.09)	+188.59
Product (Distal)	-913.085612757858 (\pm 0.10)	+288.79

Table 6.2 Relative energies in non electrostatic fully rigid XE-OO conformations. Standard deviation shown in parentheses.

All ONIOM energies in Table 6.2 are reported relative to the energy of the reactant bound in the distal state (this is the zero point energy) as this conformation represents the state at the start of the reaction. The non-electrostatic results for the rigid XE-OO model systems, show a preference for the reactant in the distal state. The proximal product state has a

substantially higher energy than the proximal reactant state, once again showing a preference for reactants over products. Overall, the reaction is shown to be strongly endothermic, and this may be due to protein rigidity in these models. The intersecting lines across each of the three plots across these fully rigid non-electrostatic XE-OO results demonstrate similar and consistent behaviour replicated across each molecular dynamics run.

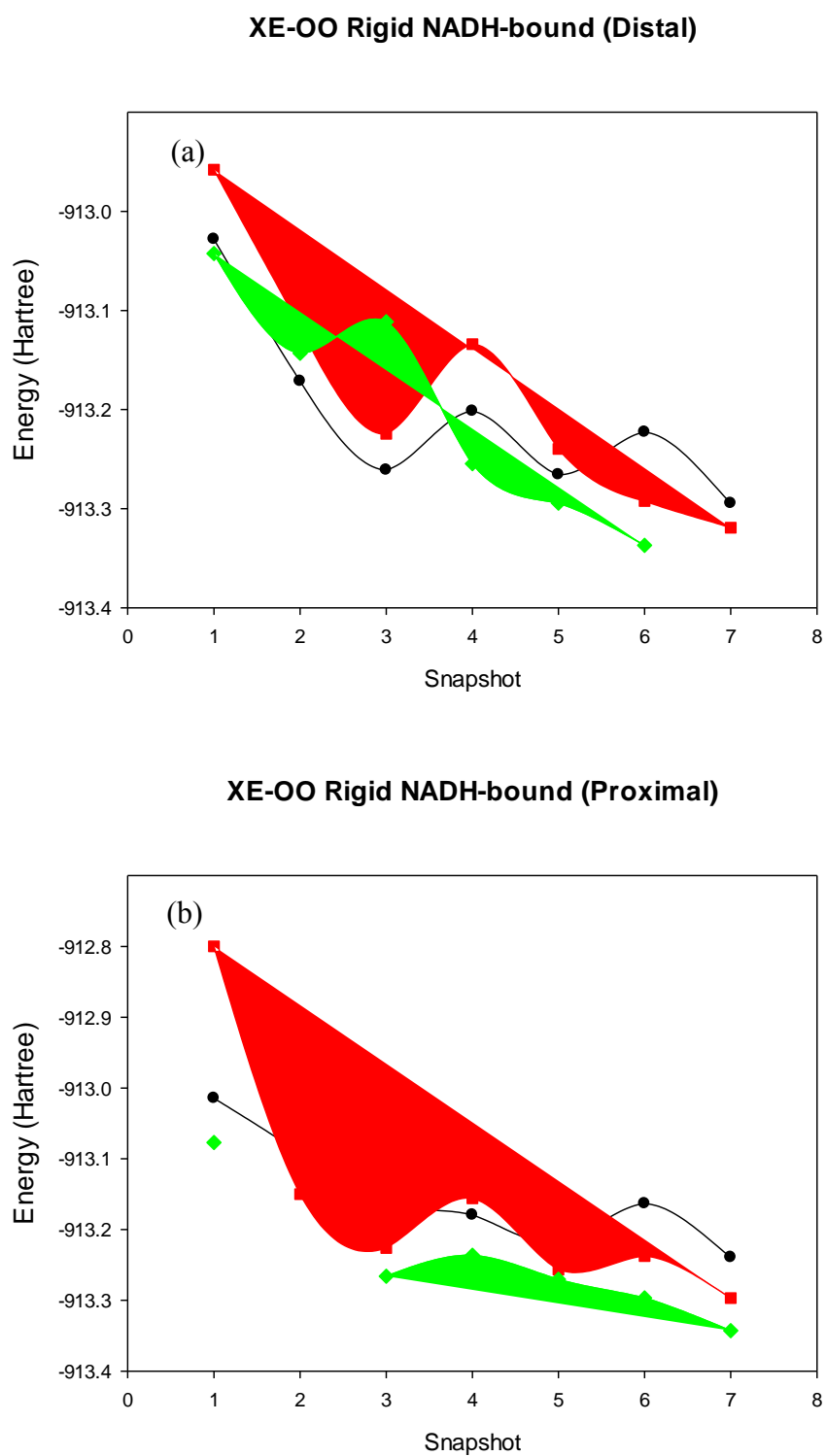
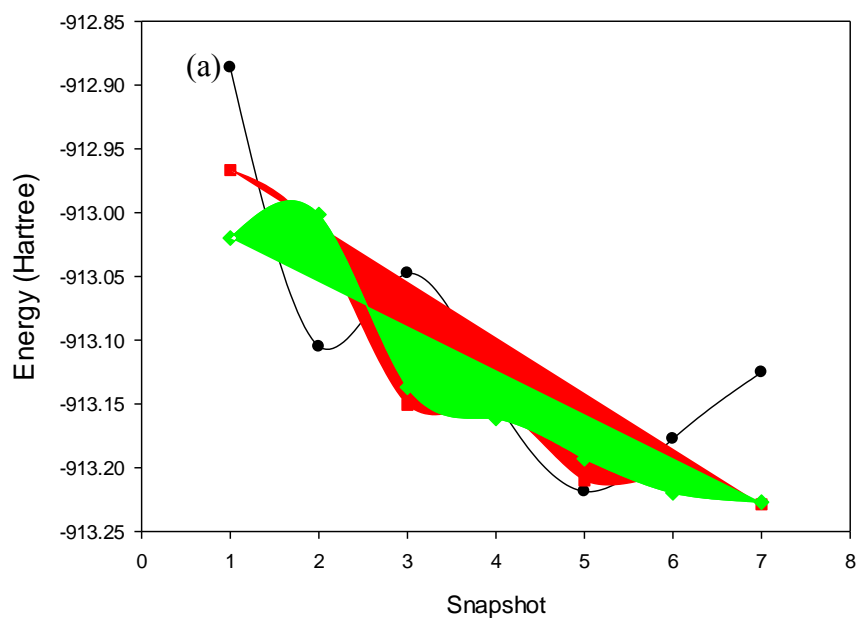


Fig. 6.3 Reactant energies in fully rigid XE-OO protein model systems:

(a) in a distal conformation and (b) in a proximal conformation. Plotted with Sigmaplot 12 [98].

XE-OO Rigid NADPH-bound (Proximal)



Rigid NADPH-bound (Distal)

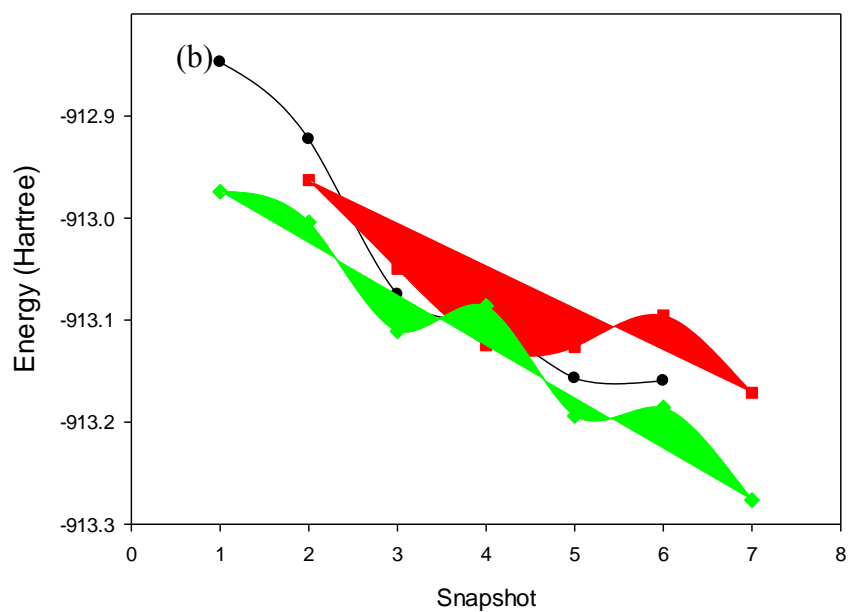


Fig. 6.4 Product energies in fully rigid XE-OO protein model systems:
(a) in a proximal conformation and (b) in a distal conformation

The importance of electrostatics upon these fully rigid XE-OO systems is seen in Figures 6.5 and 6.6 below which again present the results with respect to the order of reaction. All ONIOM energies in Table 6.3 are once again reported relative to the energy of the reactant bound in the distal state.

Rigid XE-OO Conformations (Electrostatics)		
Conformational State	Average Energy / Hartree	Relative Energy / kJ mol⁻¹
Reactant (Distal)	-918.270273725937 (± 0.51)	0
Reactant (Proximal)	-918.239626597495 (± 0.54)	+80.46
Product (Proximal)	-917.946427579946 (± 0.37)	+850.26
Product (Distal)	-918.191280367570 (± 0.51)	+207.40

Table 6.3 Electrostatic effects on the relative energies in fully rigid XE-OO conformations. Standard deviation shown in parentheses.

Although the reaction remains endothermic in both the distal and proximal configurations, greater differentiation in energy results in an energy increase as the reaction progresses through the first three states, before significantly decreasing and becoming less endothermic as the reaction cycle completes. The overall reaction is less endothermic as indicated by comparison of the distal states, than the hydride transfer reaction in the proximal states.

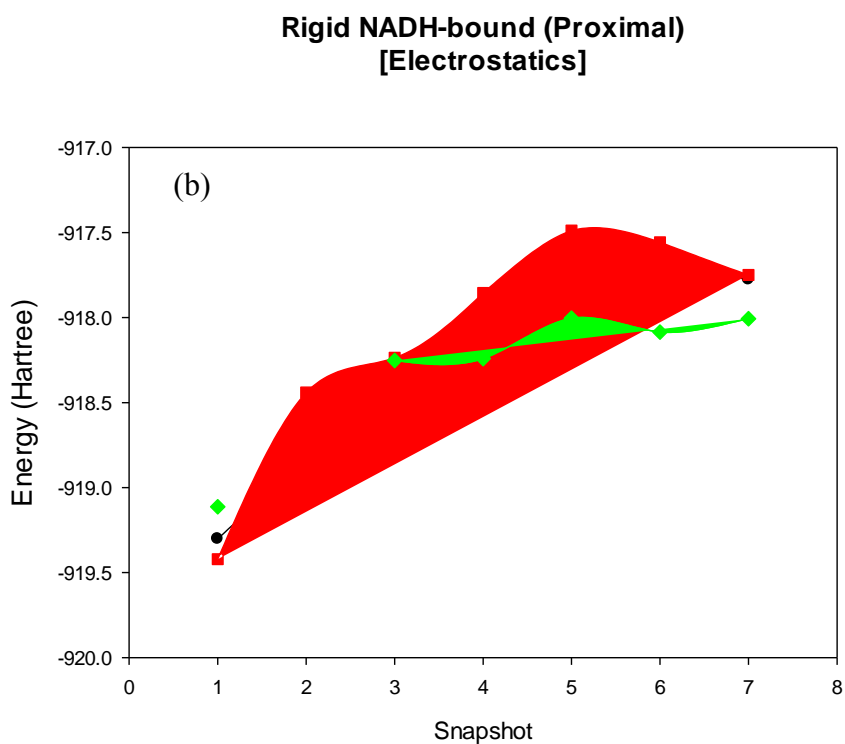
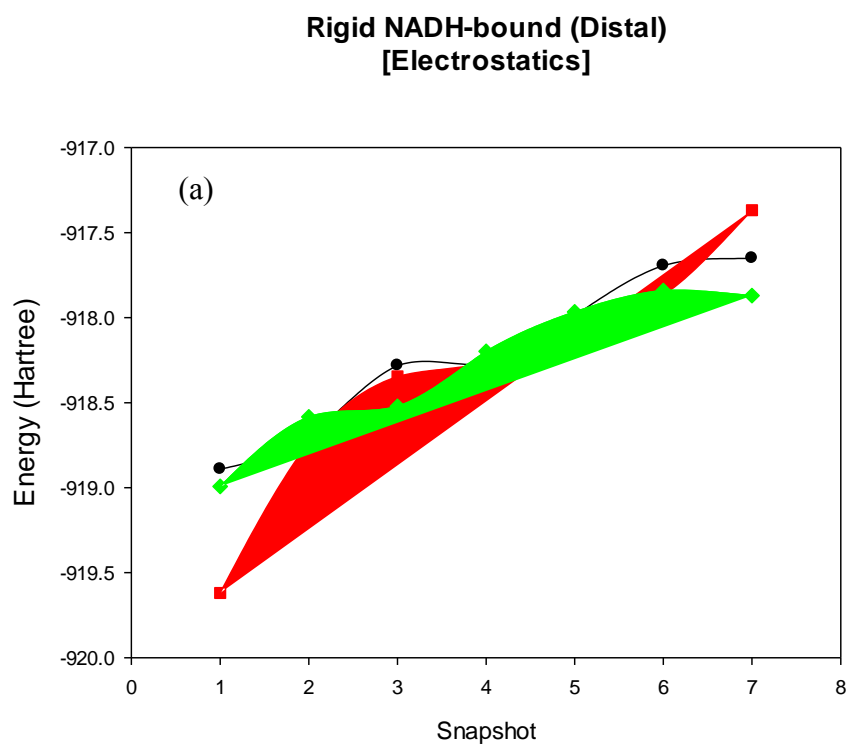


Fig. 6.5 Electrostatic effects on reactant energies in fully rigid XE-OO protein model systems:

(a) in a distal conformation and (b) in a proximal conformation

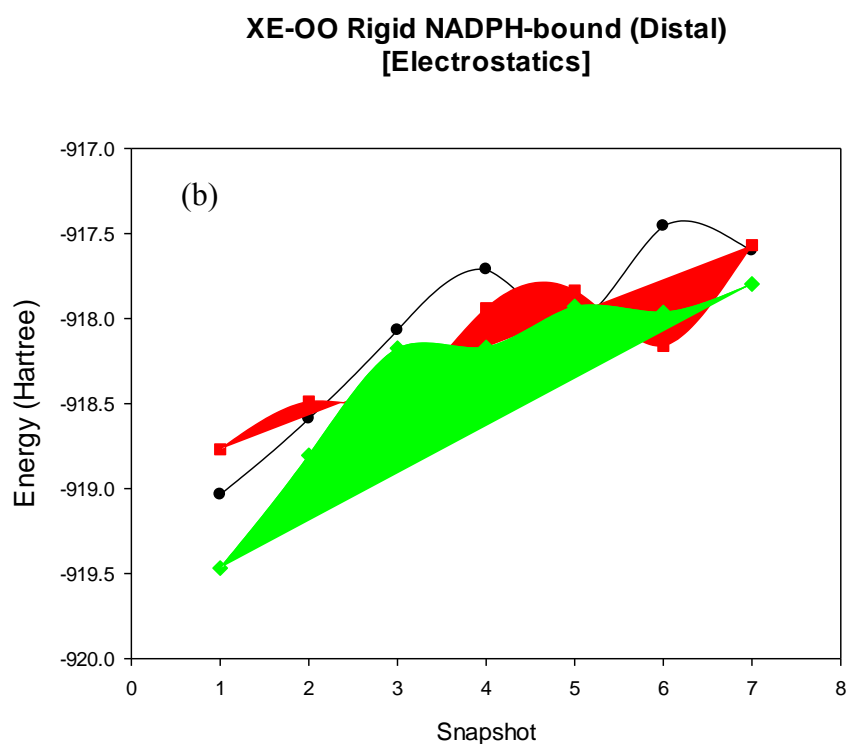
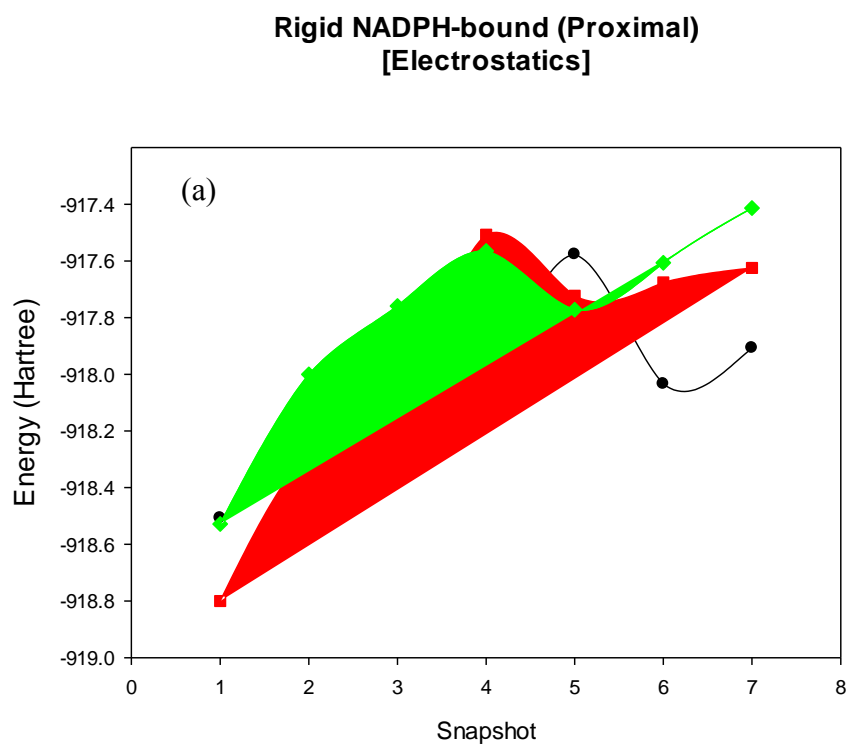


Fig. 6.6 Electrostatic effects on product energies in fully rigid XE-OO protein model systems:
(a) in a proximal conformation and (b) in a distal conformation

As rigidity prevents any relaxation of the protein structure to accommodate the specific bound species, it was felt that this rigidity may be responsible for the large positive energies observed as the reaction progresses, making the overall reaction strongly endothermic, even though it is known experimentally from the modified equilibrium constant that the reaction is exothermic and prefers product. Even though the individual snapshots are minimised first at the molecular mechanics level before being used to generate the ONIOM starting structures, introduction of a quantum mechanical portion may still have an impact on the nearby protein structure. With the aim of obtaining more reliable results, some flexibility was allowed in the ONIOM simulations by introducing a 5 Å flexible sphere around the reaction centre (see Figure 6.1), at the expense of a considerable increase in calculation time. Figures 6.7 and 6.8 show the sole effect of geometry on the energy of the partially flexible model systems whereas Figures 6.9 and 6.10 incorporate the effects of binding interactions (electrostatics) into each of the flexible models to determine their influence on the thermodynamics of the reaction. There is generally a reasonable intersection once again between the three lines in each plot suggesting that each molecular dynamics run behaves similarly, also the standard deviation values are observed to be considerably smaller.

Flexible XE-OO Conformations		
Conformational State	Average Energy / Hartree	Relative Energy / kJ mol ⁻¹
Reactant (Distal)	-913.669850693456 (± 0.079)	0
Reactant (Proximal)	-913.658914011467 (± 0.086)	+28.71
Product (Proximal)	-913.696891632973(± 0.068)	-71.00
Product (Distal)	-913.634728562038(± 0.087)	+92.21

Table 6.4 Relative energies of non electrostatic partially flexible XE-OO conformations. Standard deviation shown in parentheses.

These non-electrostatic results for these flexible XE-OO systems show that the overall reaction is now considerably lower in energy as it progresses through each of the four stages compared to the fully rigid results. The hydride transfer reaction is now also substantially exothermic in the proximal states, while the overall reaction is still endothermic, as evidenced by comparison of the distal results. The distal state preferentially binds the substrate, whilst the proximal state favours the product, as would

be expected. Hydride transfer occurs in the two proximal states and these results show the ΔH for this redox reaction is -99.71 mol^{-1} .

These proximal non-electrostatic results give a very large equilibrium constant of around 2.3×10^{17} when only the effects of enthalpy are considered and entropy is taken to be zero (a reasonable assumption based on that the proximal results reflect the hydride transfer reaction in the enzyme prior to any conformational changes) leading to the reasoning that the transition state asymmetry is influential in significantly changing the on-enzyme equilibrium for the hydride transfer reaction, shifting the equilibrium to favour the product. This equilibrium constant is considerably larger than the expected value of around 400. However this means that the hydride transfer reaction strongly favours the product at equilibrium as expected and that overall these results fit the observations made in the experimental system.

In an effort to avoid gross changes in protein structure (though there is still some relaxation due to protein flexibility in these models, but no slow changes due to dynamics), the relative energies of the proximal states using swapped oxidation states in the enzyme active site are also considered. These results may be less useful because the energies are likely to be biased due to the system used in the particular molecular dynamics calculation.

The relative energy of the product bound in the proximal reactant site is $+35.75 \text{ kJ mol}^{-1}$ and the relative energy of the reactant bound in the proximal product state is $+16.63 \text{ kJ mol}^{-1}$. The average of these swapped oxidation state relative energies is $+26.19 \text{ kJ mol}^{-1}$ and thus it can be seen that this energy is significantly endothermic when compared to the exothermic energy observed when the reaction progresses through its normal oxidation states.

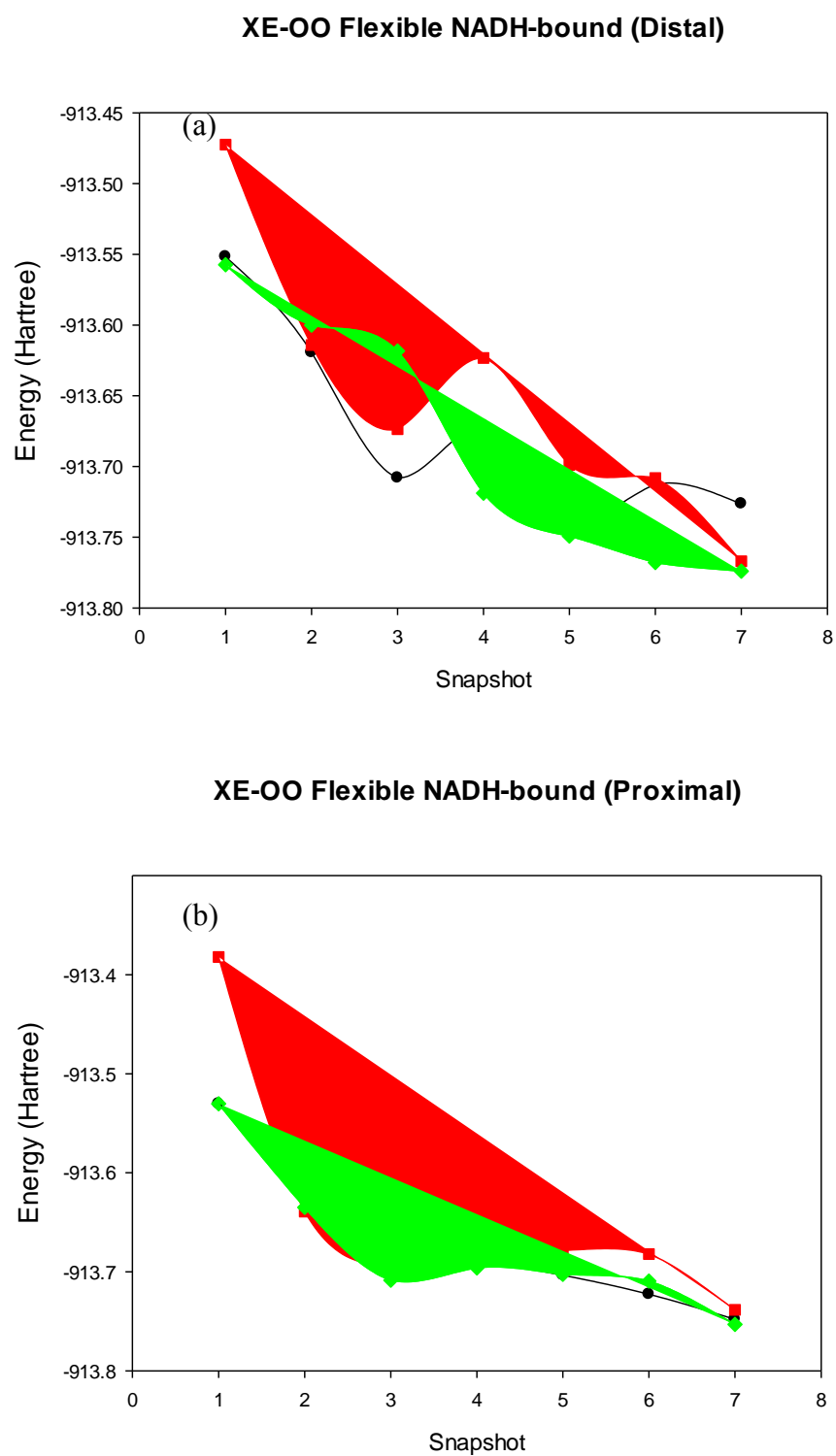


Fig. 6.7 Reactant energies in partially flexible XE-OO protein model systems:

(a) in a distal conformation and (b) in a proximal conformation

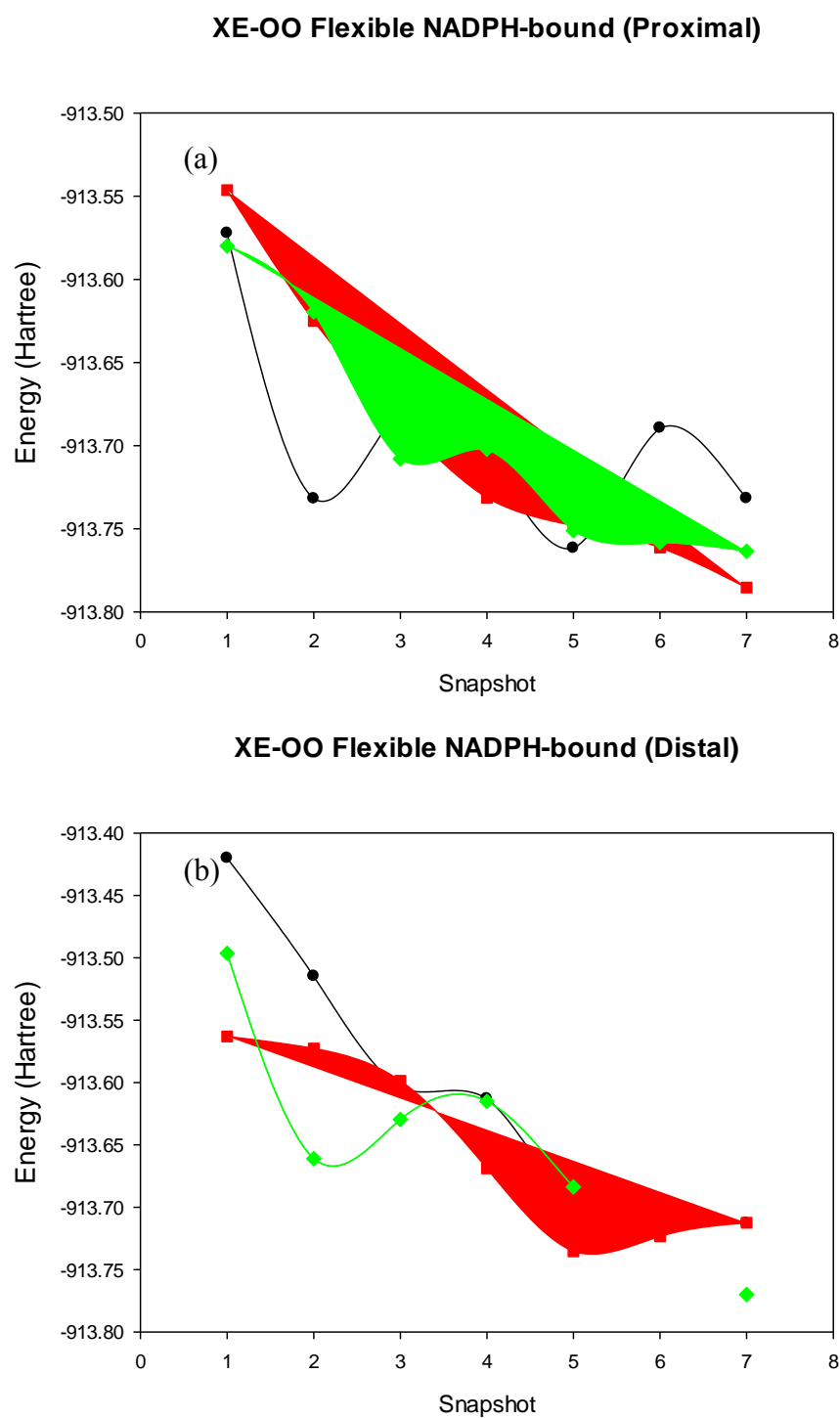


Fig. 6.8 Product energies in partially flexible XE-OO protein model systems:
(a) in a proximal conformation and (b) in a distal conformation

The influence of electrostatics upon these flexible XE-OO systems is seen in Figures 6.9 and 6.10 . Once again, all ONIOM energies are relative to the average energy of the distal reactant state at the start of the reaction, as reported in Table 6.5.

Flexible XE-OO Conformations (Electrostatics)		
Conformational State	Average Energy / Hartree	Relative Energy / kJ mol⁻¹
Reactant (Distal)	-918.635340198532 (\pm 0.47)	0
Reactant (Proximal)	- -918.624144010249 (\pm 0.50)	+29.40
Product (Proximal)	-918.375541898783 (\pm 0.36)	+682.10
Product (Distal)	-918.623662338595(\pm 0.48)	+30.66

Table 6.5 Electrostatic effects on the relative energies of partially flexible XE-OO conformations.

Standard deviation shown in parentheses.

These electrostatic results are surprising as they no longer coincide with the experimental system, but do show the same thermodynamic trend as observed for the fully rigid electrostatic results. Now the reaction remains endothermic in both the distal and proximal configurations, and the energy increases considerably as the reaction progresses through the first three states, before significantly decreasing and becoming less endothermic as the reaction cycle completes. The overall reaction is considerably less endothermic than the redox reaction of the two proximal states.

If the asymmetry of the transition state was solely responsible for the raised equilibrium constant, the reaction enthalpy would have been the same in both the non-electrostatic and electrostatic results, but since this is not the case and the inclusion of all the binding interactions on our model systems significantly alters the reaction profile and modifies the reaction thermodynamics. The ΔH of the redox reaction in the proximal states of the enzyme is now $+652.70 \text{ kJ mol}^{-1}$ compared to the non electrostatic result of $-99.71 \text{ kJ mol}^{-1}$.

The equilibrium constant in this case is approximately zero suggesting that with incorporating binding interactions into the model system, the reaction rapidly reaches equilibrium and cannot react further unless conditions are altered. It is known that hydride transfer is fast in the proximal state and these results support this observation. It is also important to remember that the entropy is taken to be zero here, and thus it is entirely possible that the inclusion of an entropic contribution would push the reaction away from equilibrium to favour products. Given the substantially large relative energy of the proximal product state ($+682.10 \text{ kJ mol}^{-1}$), which is in itself a surprising result, it is also possible that the reaction profile is modified by the binding interactions of the surrounding

protein environment such that hydride transfer does not progress through the proximal product state, instead being transferred directly from a proximal reactant state to give a distal product state. This is not an unreasonable possibility given the evidence presented by these electrostatic results, which show that the change in energy between these second and fourth state is much more reasonable than the energy change between the second and third state, going from +29.40 kJ mol⁻¹ to +30.66 kJ mol⁻¹.

If this is the case, and hydride transfer occurs concurrently as the active site switches between a proximal and a distal state, it would mean that the enthalpy of hydride transfer reaction in the enzyme would inherently be expected to have a significant entropic component due to the change in entropy upon the active site undergoing a conformational conversion from its proximal state to its distal state. As this proximal to distal conversion involves substantial changes to protein conformation and disruption of the solvent structure, a substantial entropic contribution to the reaction enthalpy would be expected in comparison to the case of hydride transfer occurring across the two proximal states, where only a small entropic component would be expected as there are no significant changes to the protein conformation or disruption of the solvent structure.

The relative energies for the swapped oxidation states on this electrostatic system are +191.90 kJ mol⁻¹ for the product bound in the proximal reactant state and +2.34 kJ mol⁻¹ for the reactant bound in the proximal product state. While these energies are wildly different, they cannot be averaged to give a single relative energy for hydride transfer in the proximal state of the enzyme. These separate energies are however, both considerably less endothermic than the energy observed when the reaction progresses through its normal states.

Any attempts to fully explain these findings further would be purely speculative without further research. As well as considering any entropic contribution on this electrostatic model system, it is also possible that the inclusion of binding interactions alters the electrostatic potential throughout transhydrogenase, just as reported in literature for hydride transfer in Dihydrofolate Reductase (DHFR) . Wong *et al.* [99] observed that although the secondary and tertiary structure of the DHFR enzyme remained intact during hydride transfer, the loop regions of the protein underwent important conformational changes, alongside these changes, a number of hydrogen bonding interactions were observed to form and break along the reaction coordinate. The conformational changes in the loop regions significantly changed the electrostatic potential throughout the enzyme and a charge deletion electrostatic analysis scheme showed key residues throughout DHFR that have a significant energetic impact on the hydride transfer reaction [99], therefore it may be reasonable to speculate that similar changes occur in *R. rubrum* transhydrogenase during hydride transfer upon inclusion of electrostatics.

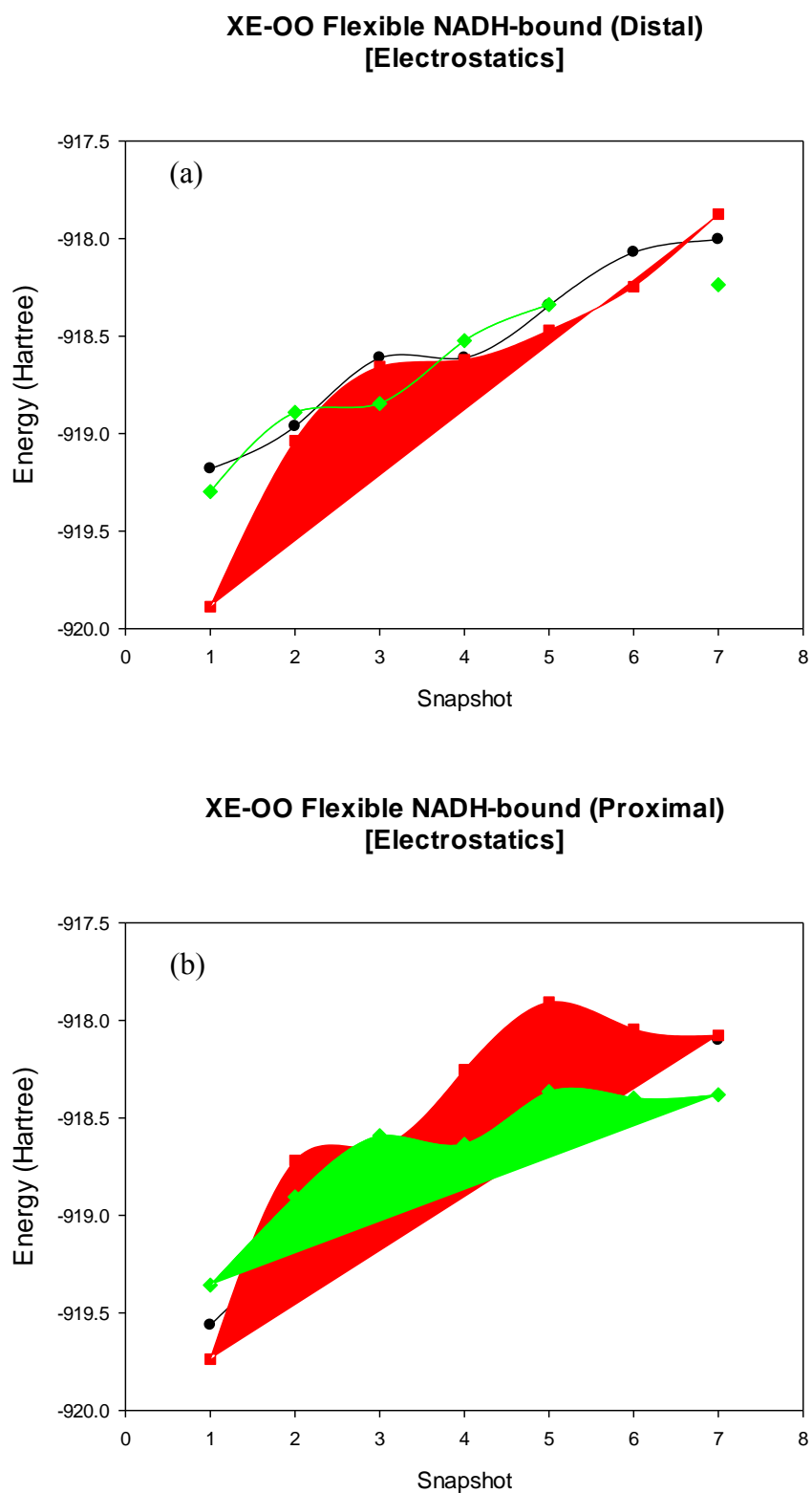
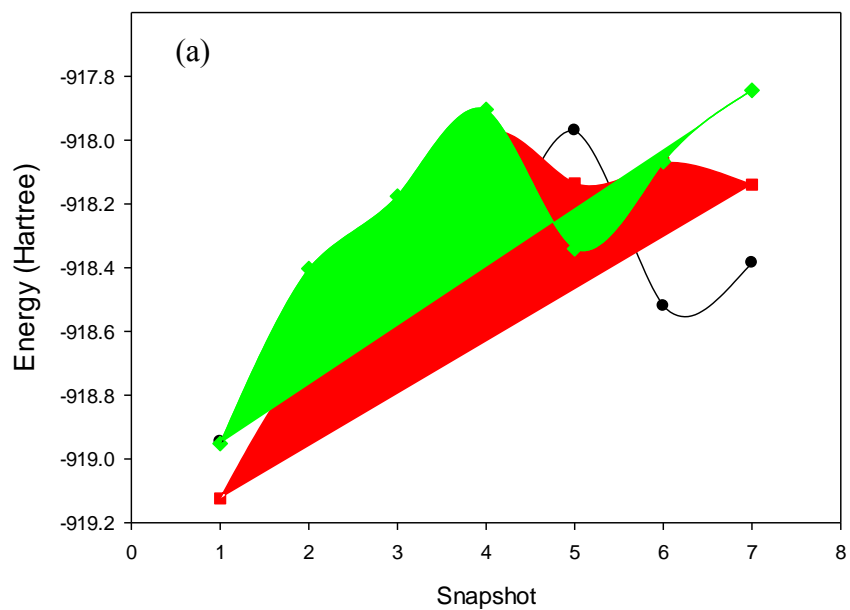


Fig. 6.9 Electrostatic effects on reactant enthalpies in partially flexible XE-OO protein model systems:
(a) in a distal conformation and (b) in a proximal conformation

**XE-OO Flexible NADPH-bound (Proximal)
[Electrostatics]**



**XE-OO Flexible NADPH-bound (Distal)
[Electrostatics]**

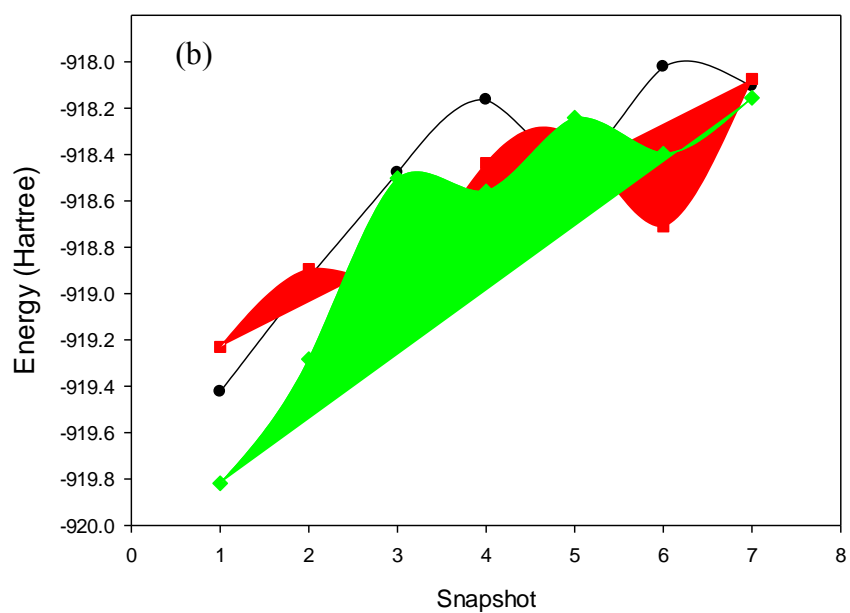


Fig. 6.10 Electrostatic effects on product energies in partially flexible XE-OO protein model systems:

(a) in a proximal conformation and (b) in a distal conformation

6.2-2 XE-ON Conformations

Figure 6.11 below shows a couple of ONIOM snapshots of the two states of the hydride transfer site of *R. rubrum* transhydrogenase with the XE-ON conformations in the enzyme. Figure 6.11 (a) shows the proximal state of the enzyme where the two nicotinamide rings are held close enough together to facilitate hydride transfer, and Figure 6.11 (b) shows the distal state of the enzyme, this conformation is one in which hydride transfer is blocked and is representative of the state prior to or following hydride transfer. The proximal snapshot has the reactant bound and therefore represents the conformation before hydride transfer, the C4-C4 atom distance is 3.95 Å. The distal snapshot has the product bound and therefore represents the conformation after hydride transfer, the C4-C4 atom distance is 5.15 Å. Once again, the C4-C4 atom distance is considerably larger in the reactant and product states of the enzyme, compared to the transition state where it is only 2.67 Å, due to the loss of a substantial 'bond' between the transferring hydride and the two C4 atoms, allowing van der Waals repulsion to push the unbound C4 atom away from the hydride, just as observed in the C2-XE-OO case discussed earlier.

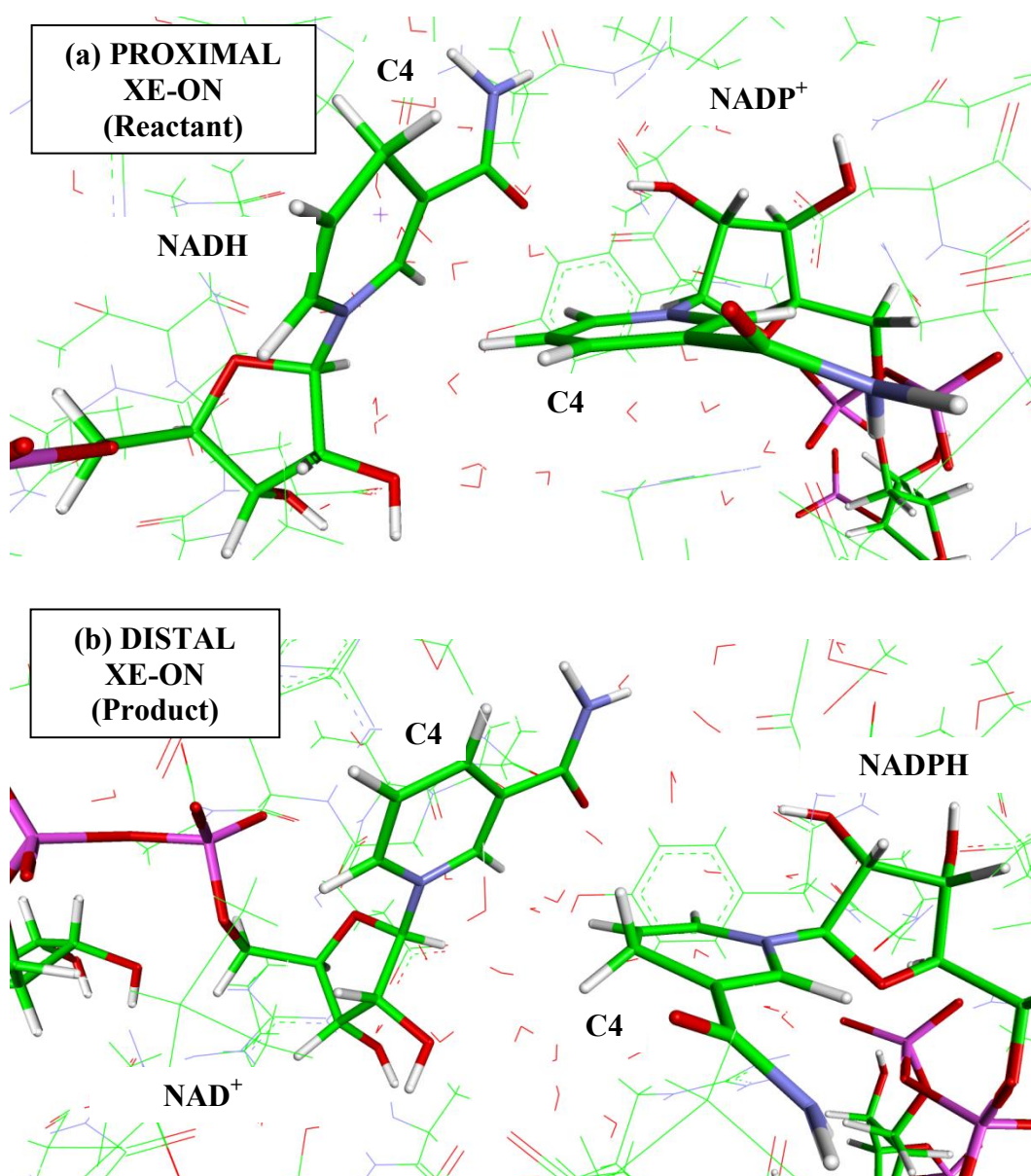


Fig. 6.11 The two XE-ON conformational states in the enzyme, the nicotinamide rings of NAD(H) and NADP(H) are either (a) held close together in a proximal reactant-bound conformation, prior to hydride transfer, or (b) held apart in a distal product-bound conformation, after hydride transfer.

Image generated by Discovery Studio [100].

The results from the ONIOM simulations of the XE-ON protein models are presented below in a way that is entirely analogous to the XE-OO results discussed in the previous section and they are once again presented in the context of following the order of reaction in the enzyme. All energies are once again relative to the energy of the distal reactant state.

Figures 6.12-6.15 below show the energies across the three molecular dynamics plotted for the XE-ON conformations in the fully rigid protein models.

Figures 6.12 and 6.13 show the sole effect of geometry on the thermodynamics of the fully rigid model systems whereas Figures 6.14 and 6.15 incorporate the effects of binding interactions (electrostatics) into each of the rigid models to determine their influence on the thermodynamics of the reaction. Just as seen for the XE-OO conformations, generally all three molecular dynamics runs seem to behave similarly in each case, and any missing data points can be attributed most commonly to convergence problems during optimisation whilst attempting to run ONIOM simulations on particular snapshots in the run.

Rigid XE-ON Conformations		
Conformational State	Average Energy / Hartree	Relative Energy / kJ mol⁻¹
Reactant (Distal)	-913.198442133556 (± 0.10)	0
Reactant (Proximal)	-913.230929633348 (± 0.09)	-85.30
Product (Proximal)	-913.078446166387 (± 0.10)	+315.05
Product (Distal)	-913.101035291177 (± 0.10)	+255.74

Table 6.6 Relative energies of fully rigid XE-ON conformations. Standard deviation shown in parentheses.

The non-electrostatic results for the rigid XE-ON model systems, summarised in Table 6.6 show that the overall reaction is endothermic as seen by comparison of the two distal results, the hydride transfer reaction is also endothermic, with a strong preference for the reactant over product in the proximal state.

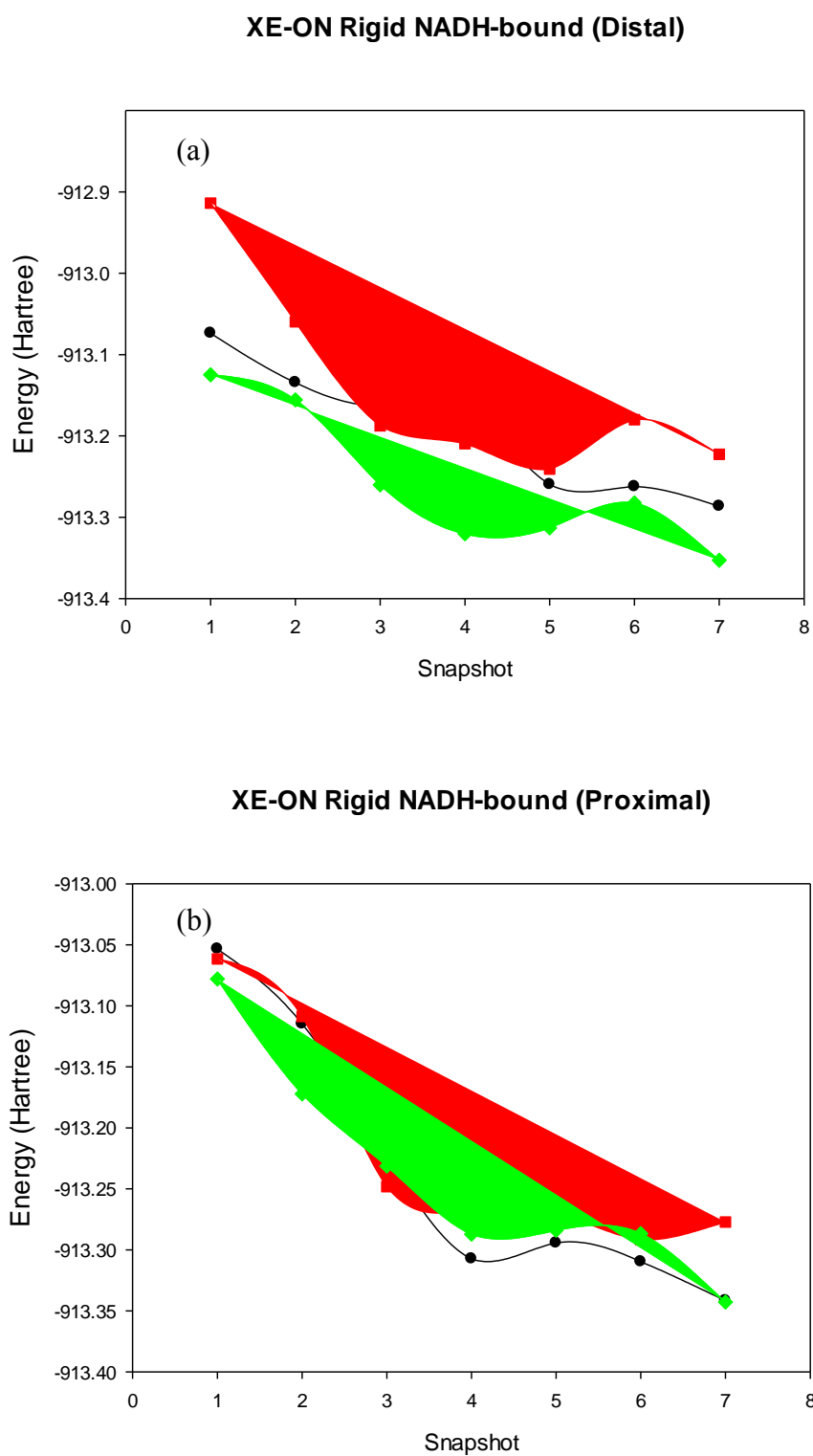
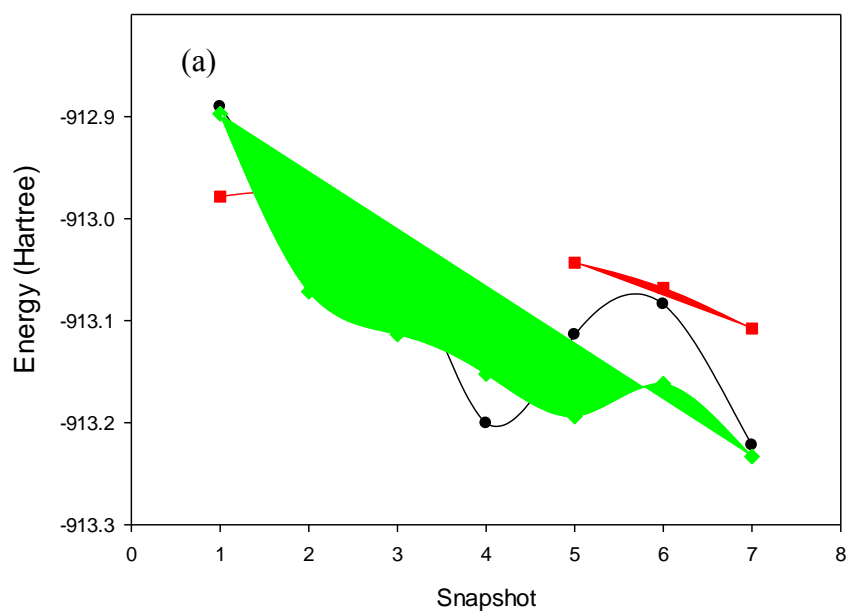


Fig. 6.12 Reactant energies for hydride transfer in fully rigid XE-ON protein model systems:
(a) reactant bound in a distal conformation and (b) reactant bound in a proximal conformation

XE-ON Rigid NADPH-bound (Proximal)



XE-ON Rigid NADPH-bound (Distal)

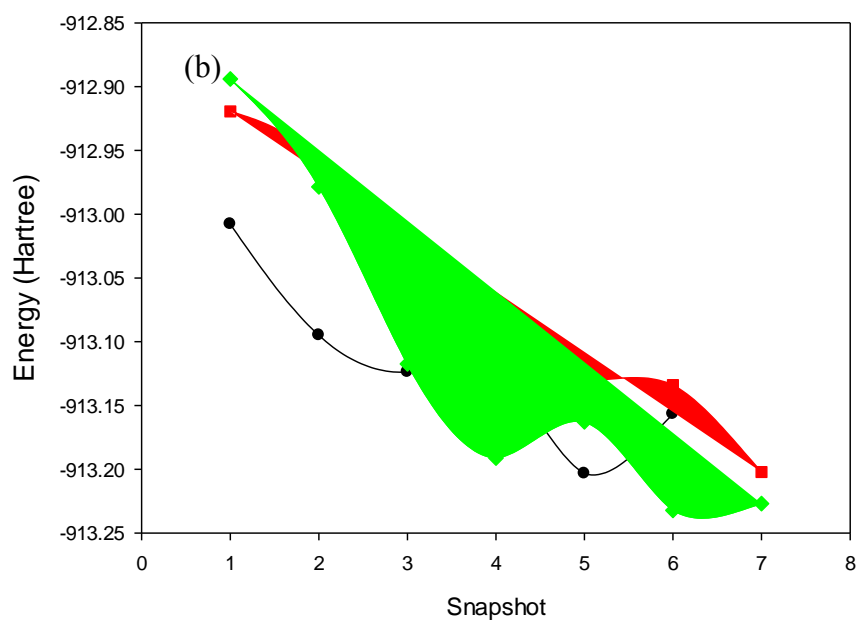


Fig. 6.13 Product energies in fully rigid XE-ON protein model systems:

(a) in a proximal conformation and (b) in a distal conformation

Generally the plots once again intersect rather well in each case for the electrostatic results in spite of missing data points for the fully rigid XE-ON systems, ensuring no wildly different MD run.

Rigid XE-ON Conformations (Electrostatics)		
Conformational State	Average Energy / Hartree	Relative Energy / kJ mol⁻¹
Reactant (Distal)	-918.381668700061 (± 0.43)	0
Reactant (Proximal)	-918.358676115175 (± 0.46)	+60.37
Product (Proximal)	-918.077499356162 (± 0.46)	+798.60
Product (Distal)	-918.210408836379 (± 0.50)	+449.64

Table 6.7 Electrostatic effects on the relative energies of fully rigid XE-ON conformations. Standard deviation shown in parentheses.

The reaction is endothermic in both the distal and proximal configurations, greater differentiation in energy results in an energy increase as the reaction progresses through the first three states, before significantly decreasing and becoming less endothermic as the reaction cycle completes. The overall reaction is less endothermic as indicated by comparison of the distal states, than the hydride transfer reaction in the proximal states.

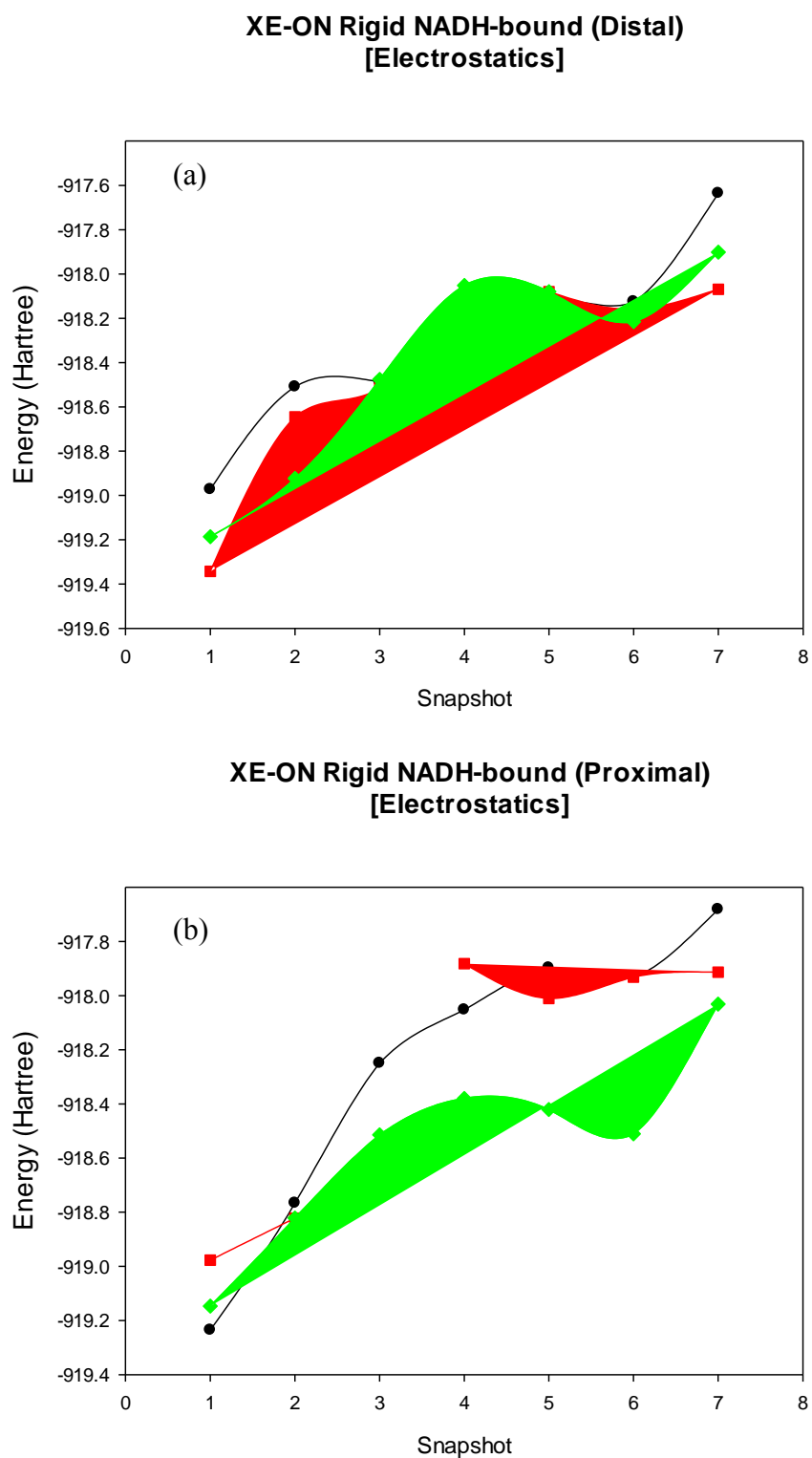


Fig. 6.14 Electrostatic effects on reactant energies in fully rigid XE-ON protein model systems:
(a) in a distal conformation and (b) in a proximal conformation

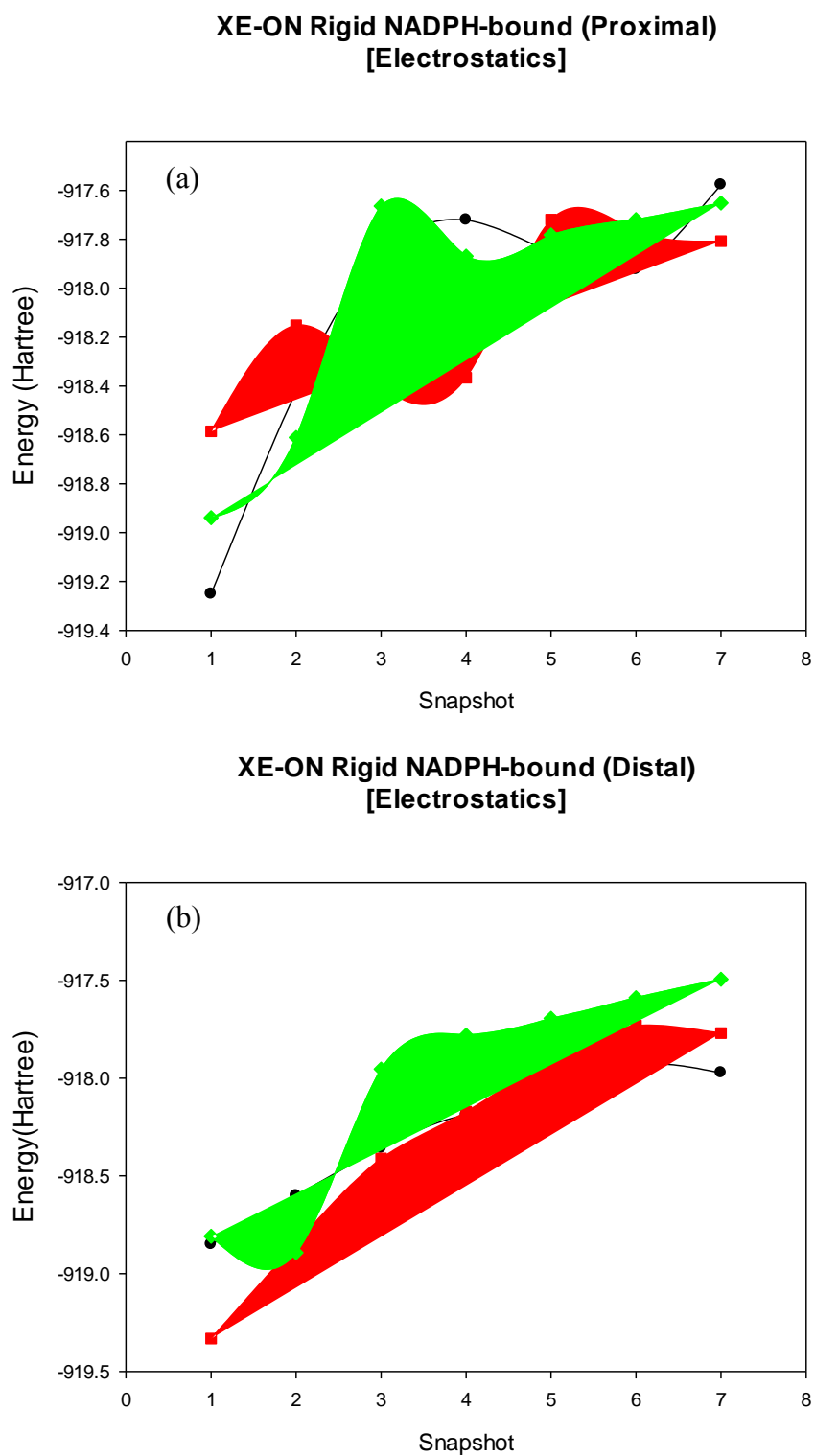


Fig. 6.15 Electrostatic effects on product energies in fully rigid XE-ON protein model systems:
(a) in a proximal conformation and (b) in a distal conformation

Figures 6.16 and 6.17 show the sole effect of geometry on the energies of the partially flexible (residues within 5 Å of the QM region - see Figure 6.1) XE-ON model systems whereas Figures 6.18 and 6.19 incorporate the effects of binding interactions (electrostatics) into each of the flexible models to determine their influence on the thermodynamics of the reaction.

Flexible XE-ON Conformations		
Conformational State	Average Energy / Hartree	Relative Energy / kJ mol ⁻¹
Reactant (Distal)	-913.684095073164 (± 0.065)	0
Reactant (Proximal)	-913.686875535587 (± 0.071)	-7.30
Product (Proximal)	-913.634679534794 (± 0.085)	+129.74
Product (Distal)	-913.673543836660 (± 0.078)	+27.70

Table 6.8 Relative energies of partially flexible XE-ON conformations. Standard deviation shown in parentheses.

The non-electrostatic results for the partially flexible XE-ON system, as summarised in Table 6.8 above, show that the overall reaction is endothermic as shown by the comparison of distal results, but the hydride transfer reaction in the proximal state is considerably more exothermic. There is a slight preference for binding the reactant in the proximal state, which is contrary to the results of both the non-electrostatic equivalent XE-OO model system and the experimental system, which both favour product.

The key result here is that these proximal non-electrostatic results show that the equilibrium is now strongly favouring the reactant rather than the product, the value of the ΔH for the hydride transfer is $+137.04 \text{ kJ mol}^{-1}$. This is supported by the very small equilibrium constant for this redox reaction of 1.38×10^{-24} when only the effects of enthalpy are considered and entropy is taken to be zero (recall that this is a reasonable assumption based on the fact that the proximal results reflect the hydride transfer reaction in the enzyme prior to any conformational changes).

These results support the conclusion that it is XE-OO and not XE-ON that is found in transhydrogenase as the XE-OO results for hydride transfer, which occurs in the proximal state both favour the product and hence fit the observations in the experimental system. It is also important to note that the overall reaction is kept endothermic and therefore binding the substrate is favoured in the distal states as expected.

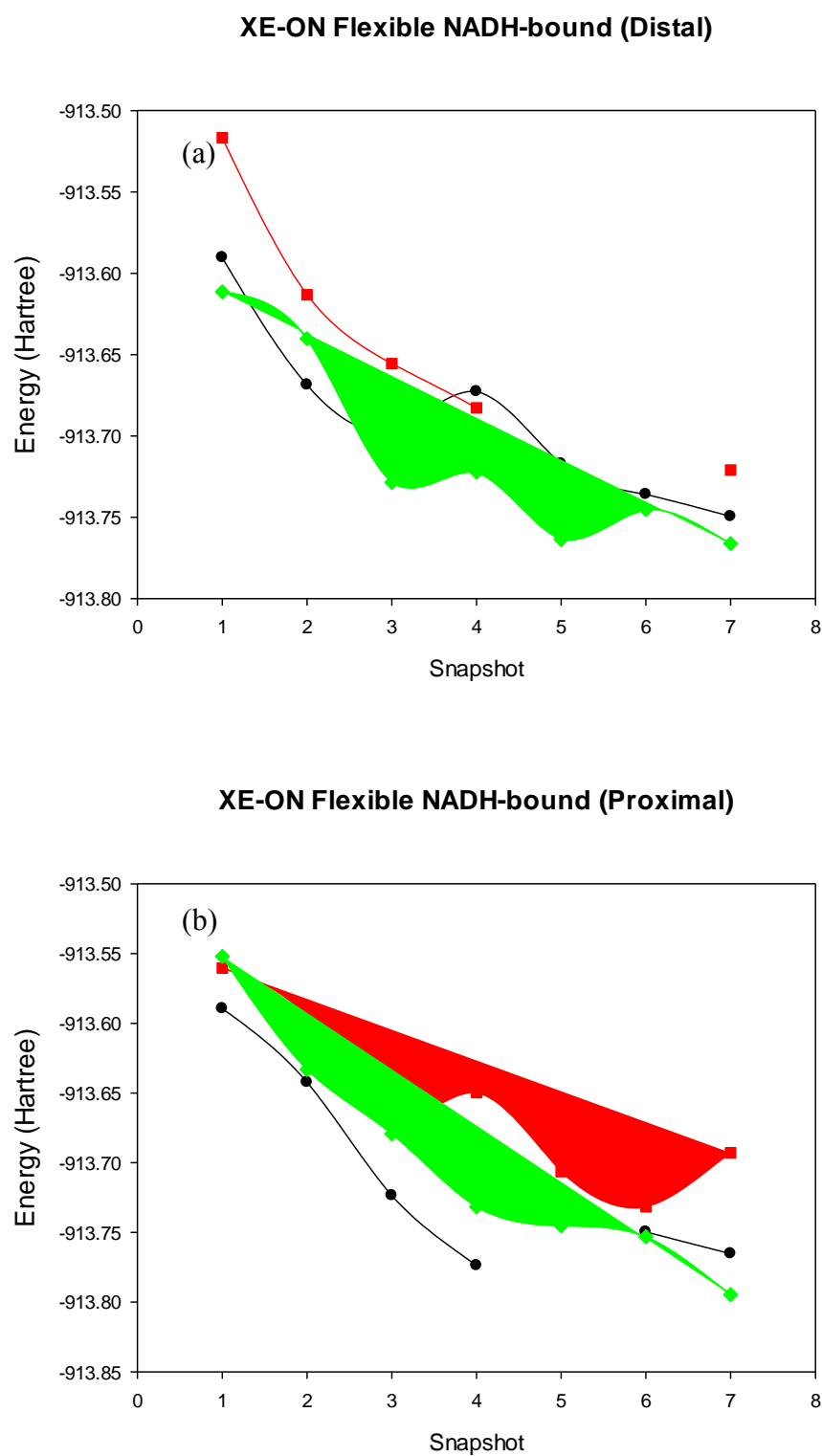
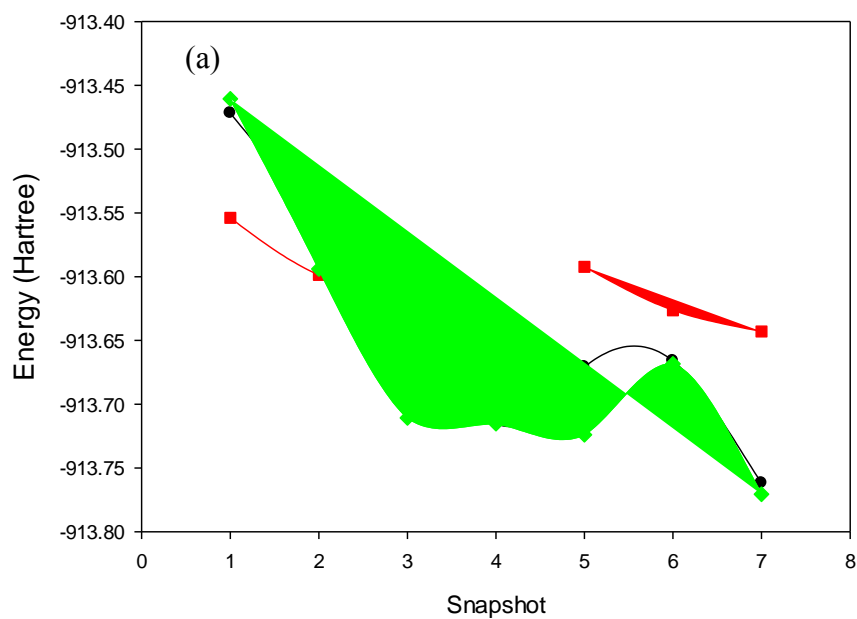


Fig. 6.16 Reactant energies in partially flexible XE-ON protein model systems:
(a) in a distal conformation and (b) in a proximal conformation

XE-ON Flexible NADPH-bound (Proximal)



XE-ON Flexible NADPH-bound (Distal)

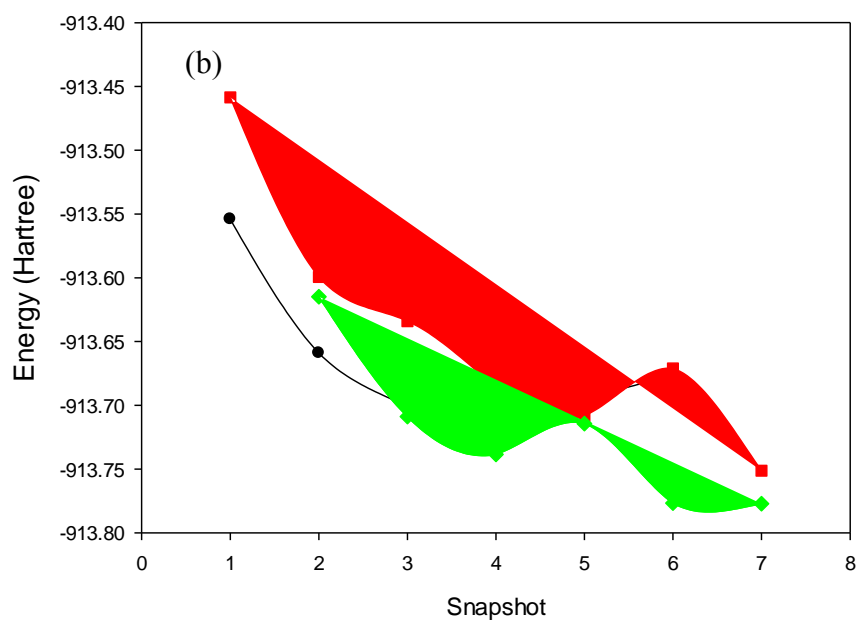


Fig. 6.17 Product energies in partially flexible XE-ON protein model systems:

(a) in a proximal conformation and (b) in a distal conformation

Electrostatic effects upon the energies of these partially flexible XE-ON systems are seen in Figures 6.18 and 6.19 below which again present the results with respect to the order of reaction.

Flexible XE-ON Conformations (Electrostatics)		
Conformational State	Average Energy / Hartree	Relative Energy / kJ mol⁻¹
Reactant (Distal)	-918.754069394152 (± 0.40)	0
Reactant (Proximal)	-918.656777152158 (± 0.41)	+255.44
Product (Proximal)	-918.485005492242 (± 0.42)	+706.43
Product (Distal)	-918.664308650965 (± 0.46)	+235.67

Table 6.9 Electrostatic effects on the relative energies of partially flexible XE-ON conformations.

Standard deviation shown in parentheses.

The reaction is endothermic in both the distal and proximal configurations, greater differentiation in energy results in an energy increase as the reaction progresses through the first three states, before significantly decreasing and becoming less endothermic as the reaction cycle completes. The overall reaction is less endothermic as indicated by comparison of the distal states, than the hydride transfer reaction in the proximal states.

These electrostatic results show that ΔH for hydride transfer in the proximal state is $+450.99 \text{ kJ mol}^{-1}$ and that equilibrium once again favours the reactant just as seen in the non-electrostatic XE-ON results, not fitting with the experimental system which is known to favour product. It is also apparent that the inclusion of the binding interactions into the model modifies the thermodynamics of the reaction such that the equilibrium constant observed for these electrostatic results is even smaller when compared to the non-electrostatic results. This electrostatic equilibrium constant is 2.97×10^{-79} compared to the non-electrostatic equilibrium constant which is 1.38×10^{-24} .

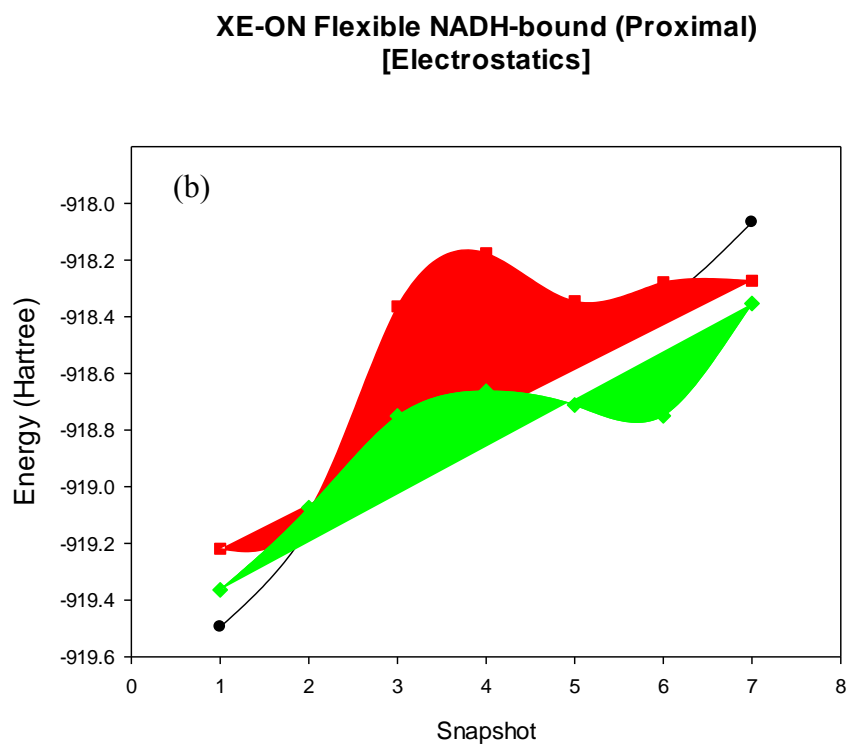
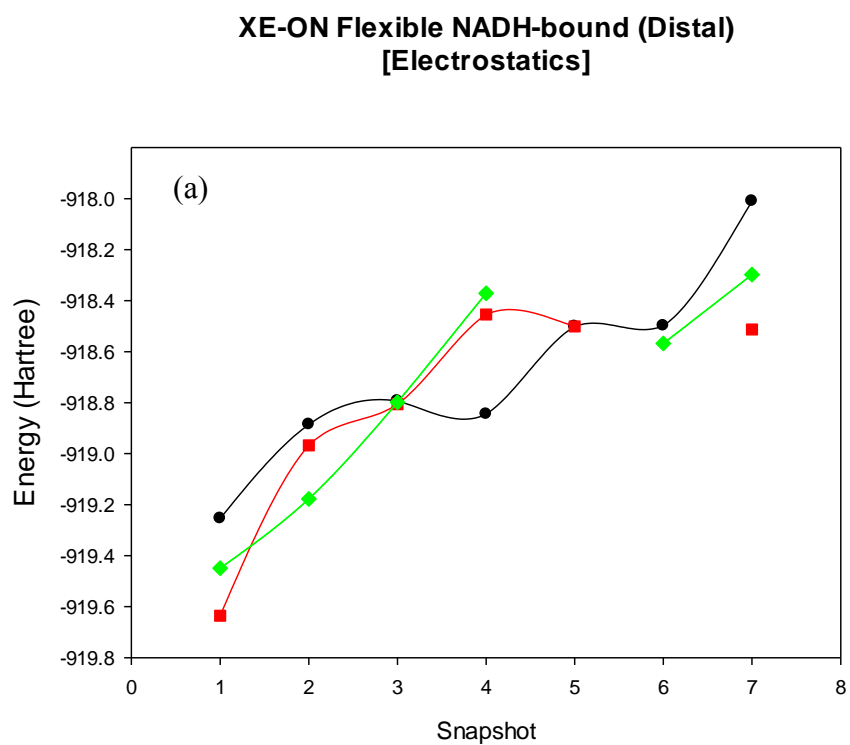
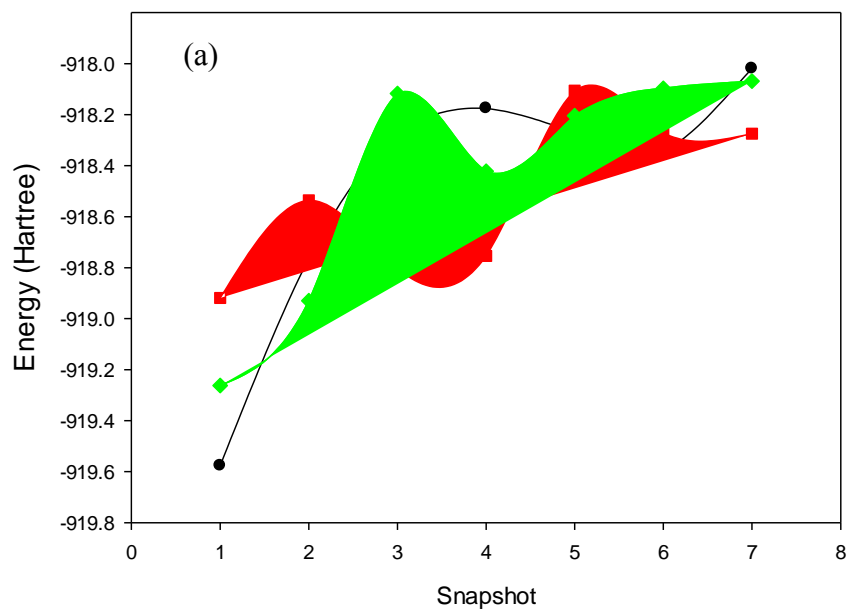


Fig. 6.18 Electrostatic effects on reactant energies in partially flexible XE-ON protein model systems:
(a) in a distal conformation and (b) in a proximal conformation

**XE-ON Flexible NADPH-bound (Proximal)
[Electrostatics]**



**XE-ON Flexible NADPH-bound (Distal)
[Electrostatics]**

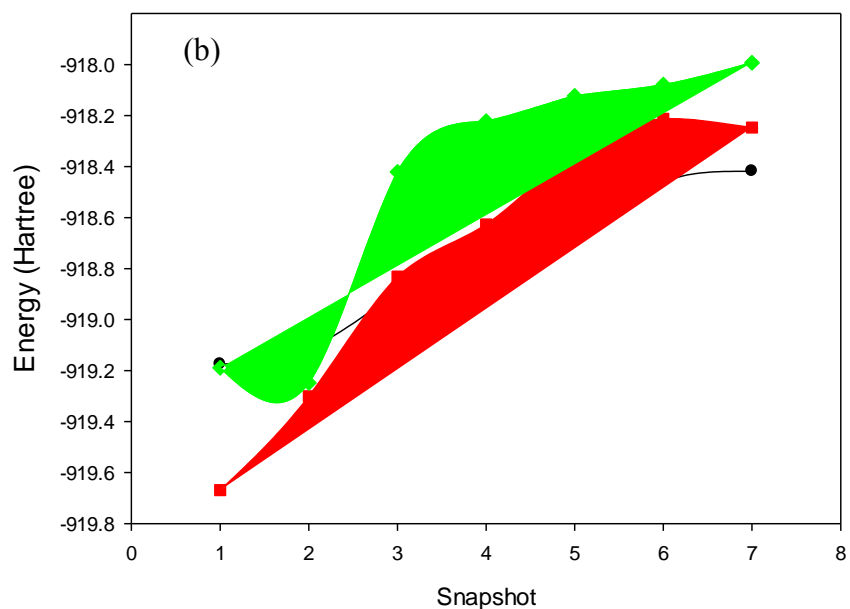


Fig. 6.19 Electrostatic effects on product energies in partially flexible XE-ON protein model systems:

(a) in a proximal conformation and (b) in a distal conformation

6.3 Preliminary ONIOM Study of dl_2dIII_2 Tetramers

In living organisms, transhydrogenase exists as a tetramer with two active sites and it is possible that the state of one active site may have an impact on the detailed conformation, and therefore reaction energetics of the other. As mentioned in Section 3.3-4, the presence of an additional dIII NADP(H) binding domain in a tetrameric structure results in many more possible reactant/product conformations compared to the simple dimers because the conformation of the second binding site, whether it is proximal or distal in nature, is not known for the crystal structure. Therefore all possible configuration combinations - whether the sites are distal or proximal, whether each site is occupied by reactants or products, and even the amide orientation - have to be considered when investigating the reaction within these tetrameric models. We can exploit symmetry for the tetramers where both sites are in the same conformation, i.e. both distal or both proximal, only one combination of sites where one site binds reactant and the other product has to be modelled. In the case where both sites have different conformations, one distal and one proximal, all combination of sites and reactants or products have to be considered. That is to say the distal site with NADH (and NADP^+) bound and the proximal site with NAD^+ (and NADPH) bound or vice versa.

Ultimately the results discussed here are very preliminary in nature as they are from the initial dynamics of an intact tetramer with the appropriate number of sites, reactants and products as described in Table 6.10. Each dynamics snapshot then generates two starting points for ONIOM calculations, one from each active site. Thus the ONIOM results represent the energetics of a single active site modified by the influence of a second site nearby in whatever specified conformation and reactant or product state. Each of the two amide orientations (XE-OO and XE-ON) have to be considered separately, though it was

assumed that each site would have the same amide orientation within any given model so the total number of combinations is twice those shown in Table 6.10.

Binding Site Conformations	Bound Cofactor (dI, dI)	Bound Cofactor (dIII, dIII)
Distal/Distal	NADH, NADH	NADP⁺, NADP⁺
Distal/Distal	NADH, NAD⁺	NADP⁺, NADPH
Distal/Distal	NAD⁺, NAD⁺	NADPH, NADPH
Distal/Proximal	NADH, NADH	NADP⁺, NADP⁺
Distal/Proximal	NADH, NAD⁺	NADP⁺, NADPH
Distal/Proximal	NAD⁺, NADH	NADPH, NADP⁺
Distal/Proximal	NAD⁺, NAD⁺	NADPH, NADPH
Proximal/Proximal	NADH, NADH	NADP⁺, NADP⁺
Proximal/Proximal	NADH, NAD⁺	NADP⁺, NADPH
Proximal/Proximal	NAD⁺, NAD⁺	NADPH, NADPH

Table 6.10 The modelled dI₂dIII₂ tetrameric configurations

It is obvious that a complete and comprehensive study of tetrameric model systems with multiple reactant and product snapshots equivalent to the one presented for the dimeric systems earlier in this chapter would be very involved and time consuming and as such is largely beyond the scope of this study. Here only a couple of snapshots from each possible tetramer conformation are considered, thus the ONIOM study of the tetramers can be considered to be preliminary in nature, offering clues to the possible influence of the second active site in the intact enzyme on hydride transfer. Another reason these preliminary findings for the tetrameric systems should be viewed with some caution is the

fact that each of the snapshots for ONIOM are from the initial minimization step rather than MD simulations.

6.3-1 XE-OO Tetramers

Though there are many possible tetrameric structures as they can have either both sites distal, both sites proximal or one site distal and the other proximal and the NAD(H) and NADP(H) cofactors in various oxidation states depending on whether a reactant or a product is bound into each site, (as listed in Table 6.2). Figure 6.20 shows the protein with each site in a different configuration, each cofactor in a different oxidation state and the nicotinamide rings in an XE-OO conformation within each site.

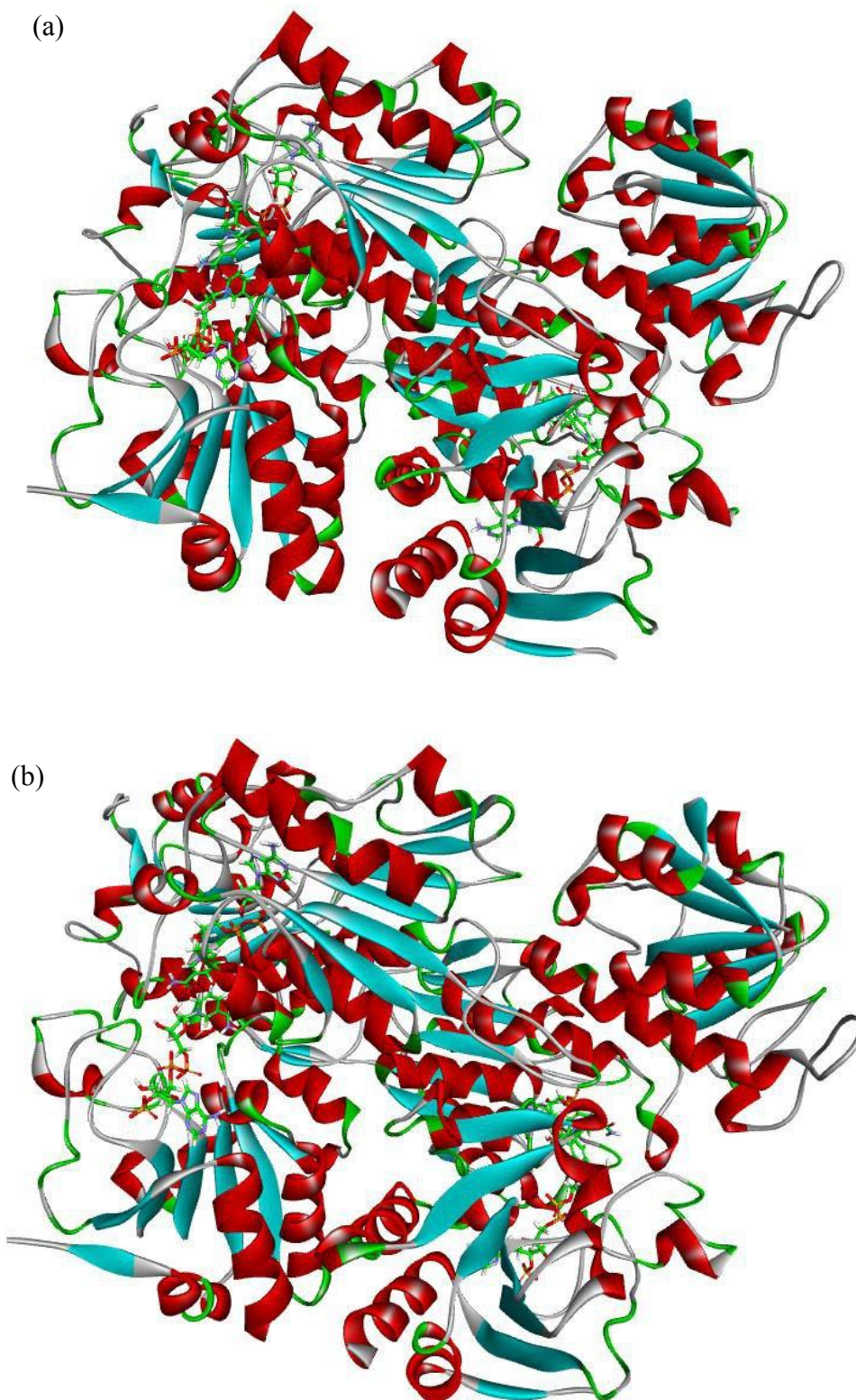


Figure 6.20 A tetrameric XE-OO dI₂dIII₂ configuration of *R. rubrum* transhydrogenase, with one site proximal (left) and one site distal (right). The bound cofactors in each active site (see text) are shown in stick format, the protein is shown as a ribbon, with secondary structure features, α -helices (red) and β -sheets shown. Images generated by Discovery Studio [100].

Figure 6.20 shows an XE-OO dI₂dIII₂ configuration for *R. rubrum* transhydrogenase with each site in a different conformation. In Figure 6.20 (a) the proximal site (nicotinamide C4-C4 ring atom distance of 3.47 Å) on the left has NAD⁺ and NADPH bound (product), conversely in the distal site (nicotinamide C4-C4 ring atom distance of 5.58 Å) on the right, NADH and NADP⁺ are bound (reactant). In Figure 6.20 (b) the proximal site (nicotinamide C4-C4 ring atom distance of 3.53 Å) on the left has NADH and NADP⁺ bound (reactant), and the distal site (nicotinamide C4-C4 ring atom distance of 5.68 Å) has NAD⁺ and NADPH bound (product).

XE-OO Tetrameric Site Configurations		XE-OO Tetrameric Site Energies / Hartree	
Conformations	Bound States	Flexible	Flexible [Electrostatics]
Distal	Reactant	--	-918.15
Distal	Reactant	-913.55	-918.83
Distal	Product	-913.63	-918.09
Distal	Product	--	-918.05
Distal	Reactant	-913.71	-918.14
Distal	Product	-913.61	-918.03
Distal	Reactant	-913.61	-918.79
Proximal	Reactant	--	-918.46
Distal	Product	-913.68	-917.99
Proximal	Product	-913.68	-918.01
Distal	Reactant	--	-918.34
Proximal	Product	-913.63	-918.01
Distal	Product	-913.64	-918.31
Proximal	Reactant	-913.64	-918.22
Proximal	Reactant	-913.62	-918.39
Proximal	Reactant	-913.68	-918.15
Proximal	Product	-913.61	-918.13
Proximal	Product	-913.67	-917.61
Proximal	Reactant	-913.64	-918.39
Proximal	Product	-913.69	-917.56

Table 6.11 Energies of the XE-OO dI₂dIII₂ tetramers in different conformations

Table 6.11 shows the energies of the individual active site snapshots from the partially flexible dI₂dIII₂ XE-OO tetramers, both with and without electrostatics. There is high variation in some of the energies, possibly owing to the fact that these snapshots were from the initial dynamics where the structures had not yet equilibrated. These large variations in energy, for sites that are notionally equivalent, and the fact that some snapshots are missing because they failed to optimise demonstrate the need for further snapshots in order to reduce this variation. The lack of data means that reliable conclusions about the impact of one active site on the energetics of the other cannot be made. However, generally these initial results seem to support the results of the XE-OO dimer study and show that binding the reactant is favoured in the distal site and the product is favoured in the proximal site, and that these preferences are strongly influenced by the binding interactions of the surrounding protein environment.

6.3-2 XE-ON Tetramers

Figure 6.21 shows the protein with each site in a different configuration, each cofactor in a different oxidation state and the nicotinamide rings in an XE-ON conformation within each site. In Figure 6.21 (a) the proximal site (nicotinamide C4-C4 ring atom distance of 3.45 Å) on the left has NAD⁺ and NADPH bound (product), conversely in the distal site (nicotinamide C4-C4 ring atom distance of 5.69 Å) on the right, NADH and NADP⁺ are bound (reactant). In Figure 6.21 (b) the proximal site (nicotinamide C4-C4 ring atom distance of 3.45 Å) on the left has NADH and NADP⁺ bound (reactant), and the distal site (nicotinamide C4-C4 ring atom distance of 5.57 Å) has both NAD⁺ and NADPH bound (product).

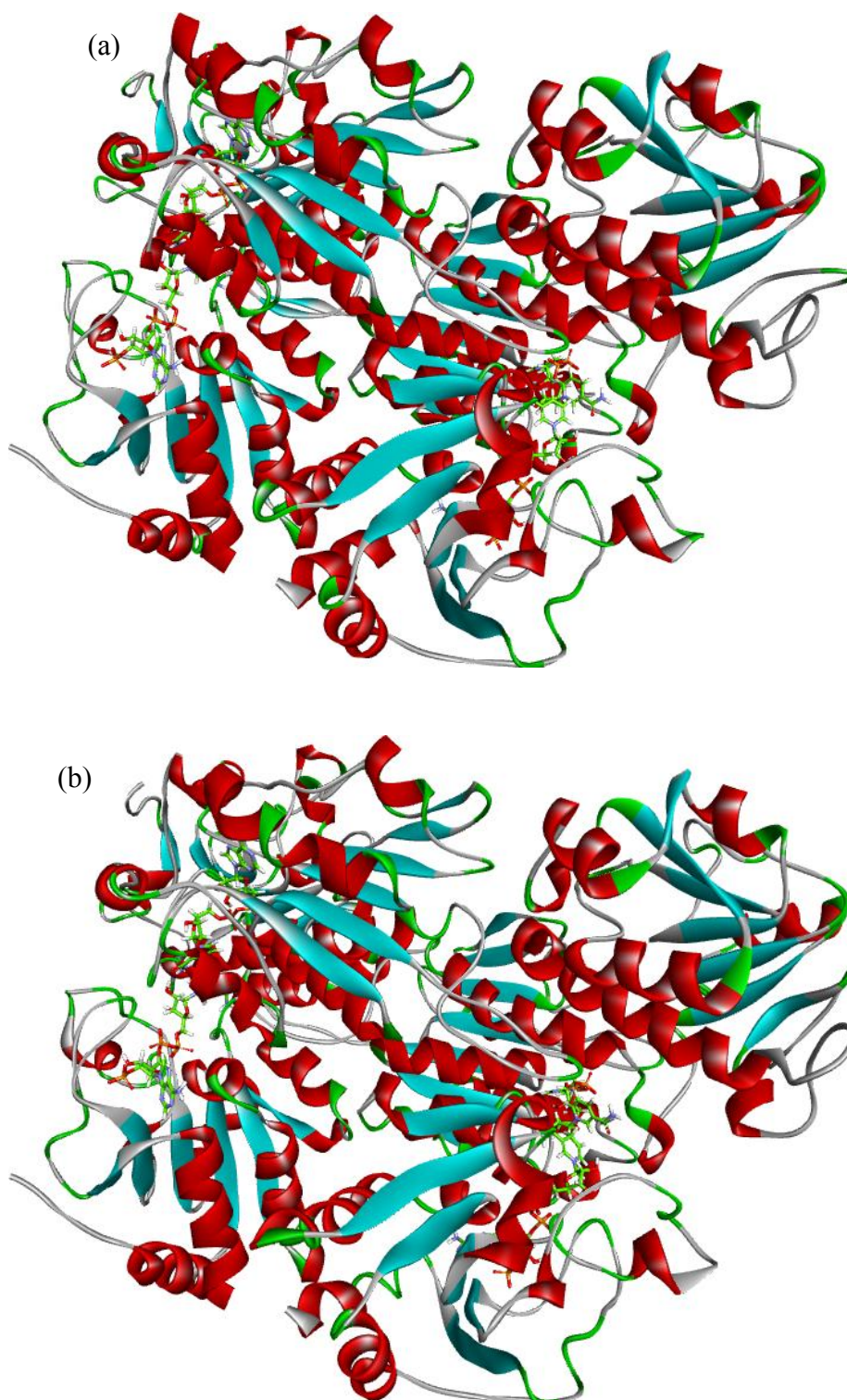


Figure 6.21 A tetrameric XE-ON dI₂dIII₂ configuration of *R. rubrum* transhydrogenase, with one site proximal (left) and one site distal (right). The bound cofactors in each active site (see text) are shown in stick format, the protein is shown as a ribbon, with secondary structure features, α -helices (red) and β -sheets shown.

XE-ON Tetrameric Site Configurations		XE-ON Tetrameric Site Energies / Hartree	
Conformations	Bound States	Flexible	Flexible [Electrostatics]
Distal	Reactant	-913.74	-918.18
Distal	Reactant	-913.67	-918.66
Distal	Product	-913.77	-917.90
Distal	Product	-913.72	-918.01
Distal	Reactant	-913.71	--
Distal	Product	-913.69	-917.94
Distal	Reactant	--	-918.33
Proximal	Reactant	-913.72	-918.62
Distal	Product	-913.69	-917.90
Proximal	Product	-913.72	-918.91
Distal	Reactant	-913.70	-918.17
Proximal	Product	-913.69	-917.91
Distal	Product	-913.65	-917.94
Proximal	Reactant	-913.73	-918.44
Proximal	Reactant	-913.64	-918.48
Proximal	Reactant	-913.65	-918.41
Proximal	Product	-913.66	-917.95
Proximal	Product	-913.58	-917.87
Proximal	Reactant	-913.64	-918.48
Proximal	Product	-913.59	-917.93

Table 6.12 Energies of the XE-ON dI₂dIII₂ tetramers in different conformations

Table 6.12 shows the energies of the individual active site snapshots from the partially flexible dI₂dIII₂ XE-ON tetramers, both with and without electrostatics. There is once again high variation in some of the energies. These large variations in energy again demonstrate the need for further snapshots in order to reduce this variation. Without more data it is once again impossible to draw reliable conclusions about the impact of one active site on the other.

6.4 Key Conclusions for the dIdIII Dimeric Systems

It has been shown that protein rigidity in the dimeric model systems considerably influences the reliability of the results, and that for a more useful set of results, some flexibility needs to be introduced into the system. This is not surprising as proteins are naturally flexible. For the partially flexible model systems, both the XE-OO and XE-ON results show that geometry can and does significantly impact the enthalpy of the hydride transfer reaction in the proximal state of *R. rubrum* transhydrogenase where the redox reaction occurs, but also crucially that it is not the only contributing factor to modifying the thermodynamics of the reaction, shifting the equilibrium constant to values far greater than 400. However, the ΔH of the hydride transfer reaction is not the same in both the non-electrostatic and electrostatic results. If it were solely the asymmetry of the transition state that was responsible for the raised equilibrium constant, the ΔH for hydride transfer would be the same in both the electrostatic and non-electrostatic model systems, and our results show that this is not the case.

The inclusion of the binding interactions (electrostatics) of the surrounding protein environment into our model systems seems to suggest that these effects can and do further

modify the reaction thermodynamics, with particular impact on modifying the on-enzyme equilibrium, taking the reaction rapidly to equilibrium in the XE-OO systems, supporting the notion that hydride transfer is rapid in the proximal state of the enzyme. It is possible that inclusion of an entropic contribution would move the reaction away from equilibrium to favour the products.

It has also been shown that the overall reaction stays endothermic in all of our models which would be expected as in the distal states of the enzyme, the substrate should bind and the product should be released. The XE-OO models fit with the experimental observation from the modified equilibrium constant that the hydride transfer reaction favours product, whereas the XE-ON models are shown to shift the equilibrium in the wrong direction and favour the reactants, thus it has been shown that XE-OO and not XE-ON is likely to be the ring arrangement found in the enzyme, providing strong evidence to resolve the ambiguity of the amide orientation of NAD(H) in the crystal structure.

CHAPTER

SEVEN

CONCLUSIONS

AND

FUTURE WORK

This chapter summarises the key findings of this research as well as suggesting directions for future research which may enhance and provide further evidence for the conclusions reached.

7. Conclusions and Future Work

This chapter summarises the key conclusions reached as a result of this research as well as suggesting directions for future research on the hydride transfer reaction in the *R. rubrum* transhydrogenase enzyme which could provide further evidence to support the findings reported here and enhance understanding of how the enzyme acts as a molecular machine by altering the equilibrium constant.

It has been shown that protein rigidity in the dimeric model systems considerably influences the reliability of the results, and that for a more useful set of results, some flexibility needs to be introduced into the system. This is not surprising as proteins are naturally flexible. For the partially flexible model systems, both the XE-OO and XE-ON results show that geometry can and does significantly impact the enthalpy of the hydride transfer reaction in the proximal state of *R. rubrum* transhydrogenase where the redox reaction occurs, but also crucially that it is not the only contributing factor to modifying the thermodynamics of the reaction, an observation supported by the preliminary results of the XE-OO tetramers, which indicate the importance of electrostatics in modifying the energies of bound reactants and products in various conformational states.

The ΔH of the hydride transfer reaction is not the same in both the non-electrostatic and electrostatic results. If it were solely the asymmetry of the transition state that was responsible for the raised equilibrium constant, the ΔH for hydride transfer would be the same in both the electrostatic and non-electrostatic model systems, and our results show that this is not the case.

The inclusion of the binding interactions (electrostatics) of the surrounding protein environment into our model systems seems to suggest that these effects can and do further modify the reaction thermodynamics, with particular impact on modifying the on-enzyme equilibrium, taking the reaction rapidly to equilibrium in the XE-OO systems, supporting the notion that hydride transfer is rapid in the proximal state of the enzyme. It is possible that inclusion of an entropic contribution would move the reaction away from equilibrium to favour the products.

The work presented here did not take into account any changes in entropy partly because these changes were expected to be small, but also because entropies are typically quite complex and time-consuming to calculate. Our interest was primarily in both the overall hydride transfer reaction, $\text{NADH}/\text{NADP}^+$ transferring a hydride to give $\text{NAD}^+/\text{NADPH}$, and in the on-enzyme equilibrium in the proximal state. Both of these involve little in the way of conformational changes in the protein or changes in solvation so the largest contribution to entropy would lie in the changes of the relative concentrations of the species. Switching from distal to proximal conformations however, does involve changes in protein conformation and substantial disruption to the solvent structure, and so these changes may be expected to be accompanied by substantial changes in entropy. Entropic change could be tackled by free-energy perturbation calculations, but these are potentially time-consuming and also we have no experience in this area.

It has also been shown that the overall reaction stays endothermic in all of our models which would be expected as in the distal states of the enzyme, the substrate should bind and the product should be released. The XE-OO models fit with the experimental observation from the modified equilibrium constant that the hydride transfer reaction

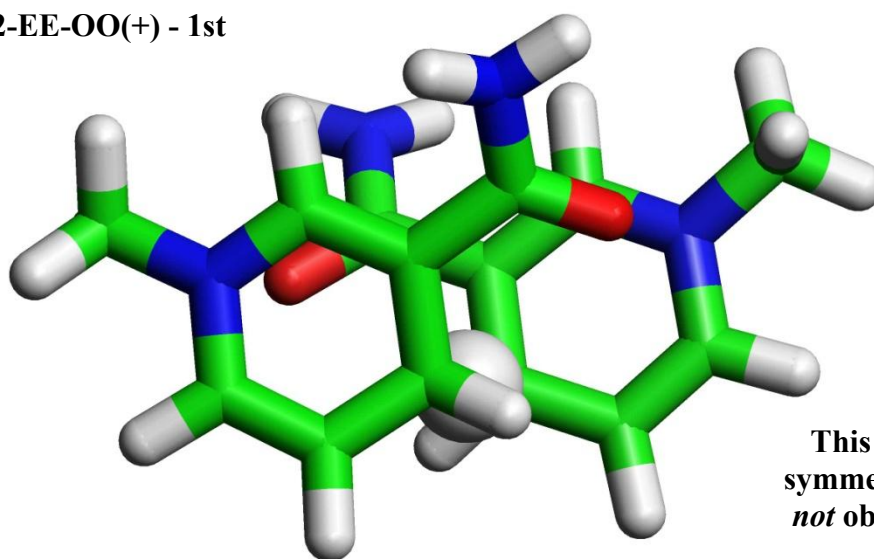
favours product, whereas the XE-ON models are shown to shift the equilibrium in the wrong direction and favour the reactants, thus it has been shown that XE-OO and not XE-ON is likely to be the ring arrangement found in the enzyme, providing strong evidence to resolve the ambiguity of the amide orientation of NAD(H) in the crystal structure.

The findings of the precursory work of this study that the main conclusion of other studies of the sequential electron/proton/electron transfer mechanism in solution [33] also holds in the gas phase – specifically that such a mechanism yields high energy intermediates which are not formed. Disproving this mechanistic possibility for hydride transfer in transhydrogenase was key before proceeding to investigate the validity of the asymmetric transition state hypothesis as an alternative mechanism that was proposed as the main aim of this research.

A thorough search for all the possible first order TS for hydride transfer between nicotinamide and dihydronicotinamide rings yields 11 C_2 and 4 C_{2v} structures and no first order C_{2h} structures (although a handful of second order transition state structures were isolated). Of these fifteen unique transition state possibilities, only four were not previously reported by Bhakta *et al.* [6], these included two new C_2 and two new C_{2v} structures. The reason for the lack of first order C_{2h} transition states is likely due to the fact that extra attractive interactions introduced by the amides (principally amide-amide and amide-pyridine nitrogen) modify the potential energy surface. The N-methyl pyridine analogues show that a C_{2h} ring orientation is the highest in energy as shown by the results of the energies of N-methyl pyridine analogues for hydride transfer. These N-methyl pyridine analogues also highlight the danger of oversimplifying a system as the N-methyl pyridine system gives a C_{2v} ring orientation as the lowest energy transition state for

hydride transfer, but the nicotinamide system favours a C₂ ring geometry for the lowest energy transition state – with the fully symmetrical C2-EE-OO TS structure being lowest energy. Importantly this is not the structural organisation which resembles the crystal structure of transhydrogenase which shows that one of the amide groups is *exo* as opposed to both being *endo* as they are in C2-EE-OO (Figure 7.1).

C2-EE-OO(+) - 1st



This lowest energy, fully symmetric transition state is *not* observed in the enzyme

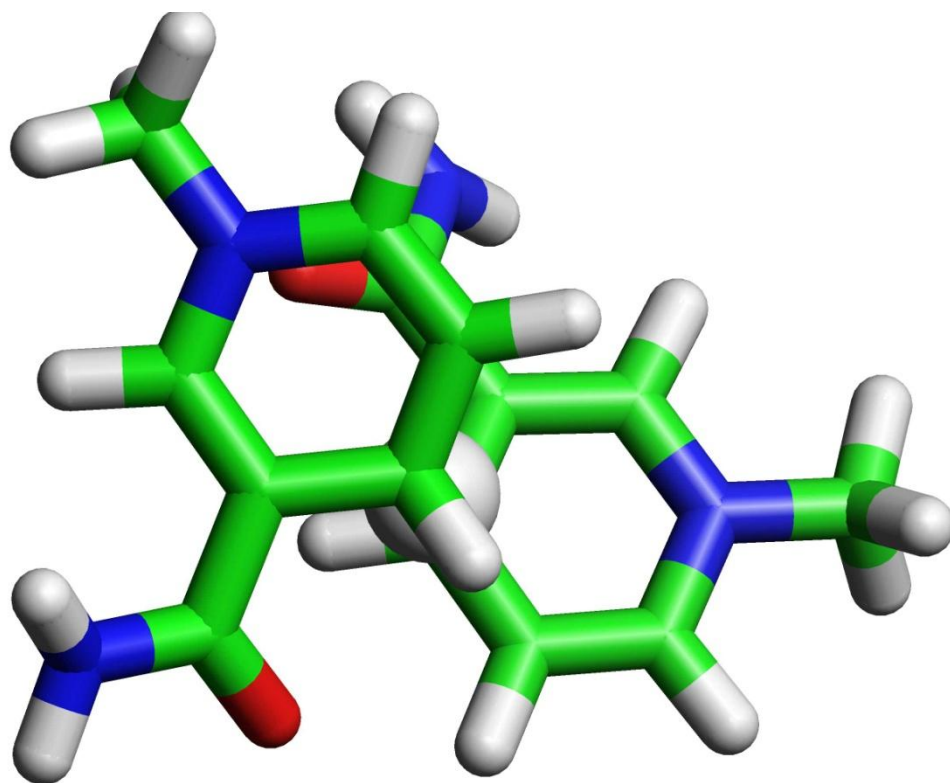
Fig 7.1 The lowest energy TS, C2-EE-OO(+), with a total energy of -913.971715621 Hartree (1 Hartree = 2625.498 kJ mol⁻¹). This structure is fully symmetric, with both C-H bond distances for hydride transfer being 1.34 Å. This TS structure is *not* observed in the enzyme.

The basis for the asymmetric hypothesis arose from the fact that the TS structure which most closely resembles the arrangement found in the crystal structure - C2-XE-OO - was observed to be asymmetric in nature and is found to be the second lowest in energy overall. Whilst the hydrogen bonding pattern in the crystal structure confirm that the amide of NADP(H) is in an O orientation, there was ambiguity about the orientation of the amide

of NAD(H) arising from the fact that protons do not appear and oxygen and nitrogen atoms are indistinguishable in the crystal structure. This lead to two asymmetric transition state possibilities, either of which could fit the crystal structure - C2-XE-OO or C2-XE-ON (Figure 7.2).

Whilst the transition state work was rather comprehensive in nature, clearly further future work is needed in order to overcome the geometry optimisation problems experienced with the PCM methodology to obtain more meaningful and accurate free energies that would be more representative of the true free energies of these TS structures fully optimised in solution as opposed to these partially optimised single point free energies which only give a partial idea of the effect of solvation on the energies of the isolated transition states, so nothing definite can be inferred from these PCM results.

(a) C2-XE-OO(+)- 2nd



(b) C2-XE-ON(+)- 9th

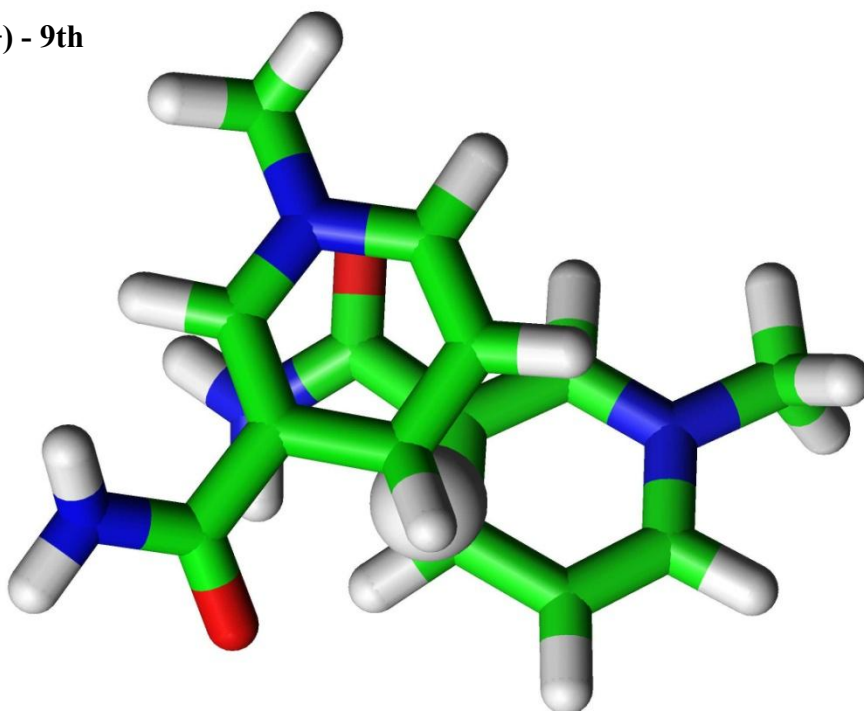


Fig. 7.3 Ambiguity in the amide orientation of one of the nucleotides in the crystal structure of transhydrogenase means that the corresponding ring arrangement could resemble either of these transition states: (a) C2-XE-OO or (b) C2-XE-ON

This thesis has described a range of transition states and identified which of those were capable of fitting the active site of the proximal state of transhydrogenase, and so correspond to the transition state for the enzyme-catalysed reaction. Though the effects of the transition states on relative reactant and product energies can be shown and a preferred transition state identified from comparison with experiment, it would be preferable to calculate the structure of the transition state in situ and follow the reaction pathway from that point by IRC or another similar method. Whilst an effort was made to try and isolate transition state structures within the enzyme to give an indication of the kind of reaction barrier for hydride transfer in transhydrogenase, this proved difficult in all cases. Rigid systems often failed or lead to isolating local minima as opposed to transition states within the enzyme. For the flexible systems, none of the calculations ran to completion despite a running time of several months. Subsequent work should focus on alternative methods for finding these enzyme-bound transition states as they would lead to further understanding of exactly how transhydrogenase works as a molecular machine.

References

- [1] Jackson, J.B. (2003) FEBS Lett. 545, 18-
- [2] Jackson, J.B. (1991) J. Bioenerg. Biomembr. 23, 715-741.
- [3] Sazanov, L.A. and Jackson, J.B. (1994) FEBS Lett. 344, 109-116.
- [4] Bizouarn, T. and Jackson, J.B. (1993) Eur. J. Biochem. 217, 7633-7770.
- [5] Rydstrom, J. and Hoek, J.B. (1988) Biochem. J. 254, 1-10.
- [6] Bhakta, T, Whitehead, S.J., Snaith, J.S., Dafforn, T.R., Wilkie, J., Rajesh, S., White, S.A., Jackson, J.B. (2007) Biochemistry 46, 3304-3318
- [7] Buckley, P.A., Jackson, J.B., Schneider, T., White, S.A., Rice, D.W. and Baker, P.J. (2000) Structure 8, 809-815.
- [8] Cotton, N.P.J., White, S.A., Peake, S.J., McSweeney, S. and Jackson, J.B. (2001) Structure 9, 165-176.
- [9] Prasad, G.S., Wahlberg, M., Sridhar, V., Sundaresan, V., Yamaguchi, M., Hatefi, Y. and Stout, C.D. (2002) Biochemistry 41, 12745-12754.
- [10] Insight II (2000) Accelrys Inc. San Diego, CA, USA

- [11] Wierenga, R.K., De Maeyer, M.C.H. and Hol, W.G.J. (1985) *Biochemistry* 24, 4346-1357.
- [12] Prasad, G.S., Sridhar, V., Yamaguchi, M., Hatefi, Y. and Stout, C.D. (1999) *Nat. Struct. Biol.* 6, 1126-1131.
- [13] Jackson, J.B., Peake, S.J. and White, S.A. (1999) *FEBS Lett.* 464, 1-8.
- [14] White, S.A., Peake, S.J., McSweeney, S., Leonard, G., Cotton, N.N.J. and Jackson, J.B. (2000) *Structure* 8, 1-12.
- [15] Venning, J.D., Rodrigues, D.J., Weston, C.J., Cotton, N.P.J., Quirk, P.G., Errington, N., Finet, S., White, S.A. and Jackson, J.B. (2001) *J. Biol. Chem.* 276, 30678-30685.
- [16] Jackson, J.B., White, S.A., Quirk, P.G. and Venning, J.D. (2002) *Biochemistry* 41, 4173-4185.
- [17] Bizouarn, T., Fjellstrom, O., Meuller, J., Axelsson, M., Bergkvist, A., Johansson, C., Karlsson, G. and Rydstrom, J. (2000) *Biochim. Biophys. Acta* 1457, 211-218.
- [18] Hatefi, Y. and Yamaguchi, M. (1996) *FASEB J.* 10, 444-452.

- [19] Meuller, J. and Rydstrom, J. (1999) *J. Biol. Chem.* 274, 19072-19080.
- [20] Bragg, P.D. and Hou, C. (2000) *Biochem. Biophys. Res. Commun.* 273, 955-959
- [21] Studley, W.K., Yamaguchi, M., Hatefi, Y. and Saier, M.H. (1999) *Microb. Comp. Genomics* 4, 173-186.
- [22] Diggle, C., Hutton, M., Jones, G.R., Thomas, C.M. and Jackson, J.B. (1995) *Eur. J. Biochem.* 228, 719-726.
- [23] Venning, J.D., Grimley, R.L., Bizouarn, T., Cotton, N.P.J. and Jackson, J.B. (1997) *J. Biol. Chem.* 272, 27535-27538.
- [24] Venning, J.D., Bizouarn, T., Cotton, N.P.J., Quirk, P.G. and Jackson, J.B. (1998) *Eur. J. Biochem.* 257, 202-209.
- [25] Venning, J.D. and Jackson, J.B. (1999) *Biochem. J.* 341, 329-337.
- [26] Pinheiro, T.J.T., Venning, J.D. and Jackson, J.B. (2001) *J. Biol. Chem.* 276, 44757-44761.
- [27] Spiegel, M.J. and Drysdale, G.R. (1960) *J. Biol. Chem.* 235, 2498-2501.

- [28] van Eikeren, P. and Grier, D.L. (1977) *J. Am. Chem. Soc.* 99, 8057-8060.
- [29] Jenks, W.P (1983) *Curr. Top. Membrane Transp.* 19, 1-18
- [30] Dewar, M.J.S., Zebisch, E.G., Healy, E.F., and Stewart, J.J.P. (1985) *J. Am. Chem. Soc.* 107, 3902
- [31] Almarrson, O. and Bruice, T.C. (1993) *J. Am. Chem. Soc.* 115, 2125-2138.
- [32] Kraulis, P.J. (1991) *J. Appl. Crystallogr.* 24, 946-950
- [33] Lee, I.-S.H., Jeoung, E.H. and Kreevoy, M.M. (1997) *J. Am. Chem. Soc.* 119, 2722-2728.
- [34] Page, C.C., Moser, C.C., Chen, X. and Dutton, P.L. (1999) *Nature* 402, 47-52.
- [35] Bruno, W.J. and Bialek, B. (1992) *Biophys. J.* 63, 689-699.
- [36] Sutcliffe, M.J. and Scrutton, N.S. (2002) *Eur. J. Biochem.* 269, 3096-3102.
- [37] BlueBEAR (2008) University of Birmingham, Birmingham, England, UK
- [38] Silicon Graphics International. Fermont, CA, USA

- [39] Tatham, S., Dunn, O., Harris, B., and Nevins J. (2007) Cambridge, England, UK
- [40] Cramer, C.J. (2006) Essentials of Computational Chemistry: Theories and Models
(Various Chapters therein)
- [41] Petersson, G.A., Bennett, A., Tensfeldt, T.G., Al-Laham, M.A., Shirley, W.A., and
Mantzaris, J., J. Chem. Phys., 89 (1988) 2193-2218.
- [42] Petersson, G.A. and Al-Laham, M.A. J. Chem. Phys., 94 (1991) 6081-6090.
- [43] Ditchfield, R., Hehre, W.J., and Pople, J.A. J. Chem. Phys., 54 (1971) 724.
- [44] Hehre, W.J., Ditchfield, R., and Pople, J.A. J. Chem. Phys., 56 (1972) 2257.
- [45] Hariharan , P.C., and Pople, J.A., Theor. Chem. Acc., 28 (1973) 213-222.
- [46] Hariharan, P.C. and Pople, J.A., Mol. Phys., 27 (1974) 209-214.
- [47] Gordon, M.S. Chem. Phys. Lett., 76 (1980) 163-168.
- [48] Francel, M.M., Pietro, W.J., Hehre, W.J., Binkley, J.S., DeFrees, D.J., Pople, J.A.,
and Gordon, M.S., *J. Chem. Phys.*, **77** (1982) 3654-3665.
- [49] Binning Jr., R.C, and Curtiss, L.A., J. Comp. Chem., 11 (1990) 1206-1216.

- [50] Blaudeau, J.P., McGrath, M.P., Curtiss, L.A., and Radom, L. J. Chem. Phys., 107 (1997) 5016-5021.
- [51] Rassolov, V.A., Ratner, M.A., Pople, J.A., Redfern, P.C., and Curtiss, L.A., J. Comp. Chem., 22 (2001) 976-984.
- [52] Rassolov, V.A., Pople, J.A., Ratner, M.A., and Windus, T.L., J. Chem. Phys., 109 (1998) 1223-1229.
- [53] Ademo, C., and Barone, V. (1998) J. Chem. Phys. 108, 664.
- [54] Perdew, J.P. (1991) Electronic Structure of Solids '91
- [55] Krishnan, R., Frisch, M. J., and Pople, J. A. (1980) J. Chem. Phys. 72, 4244.
- [56] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C.

Gonzalez, and J. A. Pople (2004) Gaussian 03, Revision E.01, Gaussian, Inc., Wallingford CT, USA.

[57] http://wanglab.bu.edu/g03guide/G03Guide/www.gaussian.com/g_ur/keywords.htm

Gaussian 03 Online Manual

[58] C. Peng, P. Y. Ayala, H. B. Schlegel, and M. J. Frisch, *J. Comp. Chem.* 17, 49 (1996).

[59] C. Gonzalez and H. B. Schlegel, *J. Chem. Phys.* 90, 2154 (1989).

[60] C. Gonzalez and H. B. Schlegel, *J. Phys. Chem.* 94, 5523 (1990).

[61] Jensen, F. (2006) *Introduction to Computational Chemistry* (Various Chapters therein)

[62] F. Maseras and K. Morokuma, *J. Comp. Chem.* 16, 1170 (1995)

[63] S. Humbel, S. Sieber, and K. Morokuma, *J. Chem. Phys.* 105, 1959 (1996).

[64] T. Matsubara, S. Sieber, and K. Morokuma, *Int. J. Quant. Chem.* 60, 1101 (1996).

[65] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, *J. Phys. Chem.* 100, 19357 (1996)

[66] M. Svensson, S. Humbel, and K. Morokuma, *J. Chem. Phys.* 105, 3654 (1996).

[67] S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma, and M. J. Frisch, *J. Mol. Struct. (Theochem)* 462, 1 (1999).

[68] T. Vreven and K. Morokuma, *J. Comp. Chem.* 21, 1419 (2000).

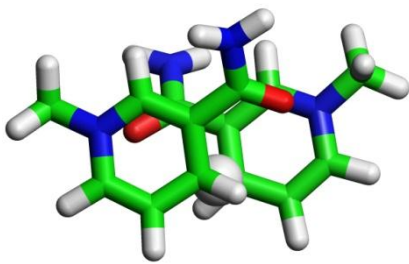
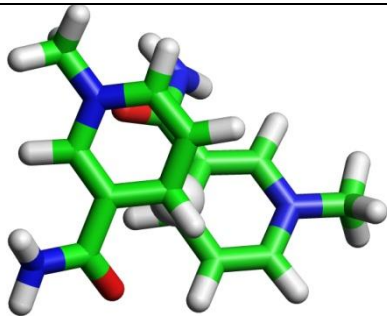
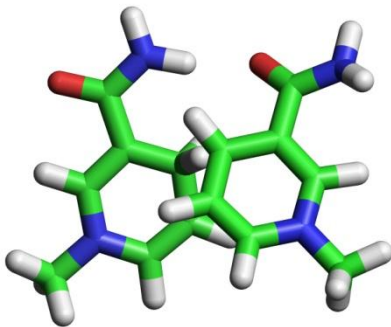
[69] Rappé, A.K., Casewit, C.J., Colwell, K.S., Goddard III, W.A., and Skiff, W.M. (1992) *J. Am. Chem. Soc.* 114, 10024-10035.

- [70] Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M. Jr, Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A. (1995) *J. Am. Chem. Soc.* 117: 5179–5197.
- [71] M. T. Cancès, B. Mennucci, and J. Tomasi, *J. Chem. Phys.* 107, 3032 (1997).
- [72] M. Cossi, V. Barone, B. Mennucci, and J. Tomasi, *Chem. Phys. Lett.* 286, 253 (1998).
- [73] B. Mennucci and J. Tomasi, *J. Chem. Phys.* 106, 5151 (1997).
- [74] M. Cossi, G. Scalmani, N. Rega, and V. Barone, *J. Chem. Phys.* 117, 43 (2002).
- [75] B. Mennucci, E. Cancès, and J. Tomasi, *J. Phys. Chem. B* 101, 10506 (1997).
- [76] J. Tomasi, B. Mennucci, and E. Cancès, *J. Mol. Struct. (Theochem)* 464, 211 (1999).
- [77] D. M. Chipman, *J. Chem. Phys.* 112, 5558 (2000).
- [78] D.A. Case, T.E. Cheatham, III, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang and R. Woods. *J. Computat. Chem.* 26, 1668-1688 (2005).
- [79] J.W. Ponder and D.A. Case *Adv. Prot. Chem.* 66, 27-85 (2003)
- [80] U. Ryde, *Proteins, Struct. Funct. Genet.* 21 40-50 (1995)
- [81] U. Ryde, *Prot. Sci.* 4 1124-1132 (1995)

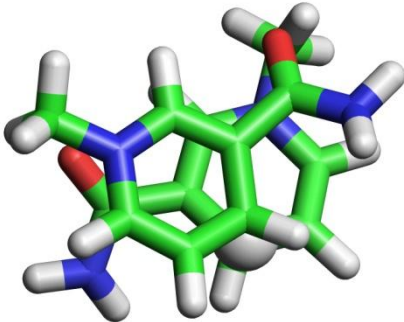
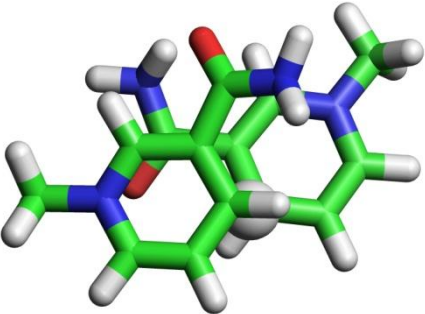
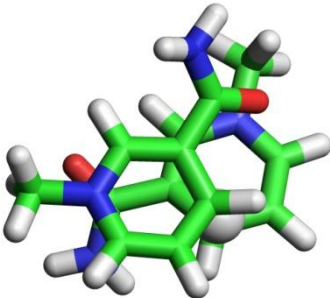
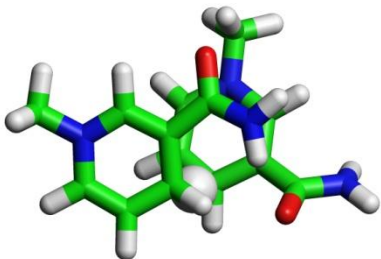
- [82] N. Holmberg, U. Ryde, and L. Bulow, *Prot. Engin.*, 12 851-856 (1999).
- [83] Christofferson, A., and Wilkie. J. (2009) *Biochem. Soc. Trans.* 37, 413-418.
- [84] Wu, Y.-D., Lai, D. K. W., and Houk, K. N. (1995) *J. Am. Chem. Soc.* 117, 4100-4108.
- [85] Møller, C. and Plesset, M.S., *Phys. Rev.*, 46 (1934) 618-622.
- [86] Head-Gordon, M., Pople, J.A. and Frisch, M.J. *Chem. Phys. Lett.*, 153 (1988) 503-506.
- [87] Saebø, S. and Almlöf, J., *Chem. Phys. Lett.*, **154** (1989) 83-89.
- [88] Frisch, M.J., Head-Gordon, M. and Pople, J.A., *Chem. Phys. Lett.*, 166 (1990) 275-280.
- [89] Frisch, M.J., Head-Gordon, M. and Pople, J.A. *Chem. Phys. Lett.*, 166 (1990) 281-289.
- [90] Head-Gordon, M. and Head-Gordon, T., *Chem. Phys. Lett.*, 220 (1994) 122-128
- [91] Mather, O. M., Singh, A., van Boxel, G. I., White, S. A., and Jackson, J. B. (2004), *Biochemistry*, 43, 10952-10964

- [92] Donkersloot, M. C. A.; Buck, H. M. J. *Am. Chem. Soc.* (1981), 103, 6554.
- [93] Singh, A., Venning, J. D., Quirk, P., van Boxel, G. I., Rodrigues, D. J., White, S. A., and Jackson, J. B. (2003) *J. Biol. Chem.* 278, 33208-33216.
- [94] Cummins, P. L., and Gready, J. E. (1989) *J. Mol. Struct. (Theochem.)* 183, 161–174
- [95] Wu, Y.D., and Houk, K. N. (1993) *J. Org. Chem.* 58, 2043–2045
- [96] Brondijk, T. H. C., van Boxel, G. I., Singh, A., Mather, O. M., White, H. A., Quirk, P. G., White, S. A., and Jackson, J. B. (2006) *J. Biol. Chem.* 281, 13345-13354.
- [97] Boyle, N.M.O., Banck, M., James, C.A., Morley, C., Vandermeersch, T. and Hutchison, G. R. (2011) *J. Cheminf.* 3, 33.
- [98] *Systat Software Inc., (2012) Sigmaplot, Version 12.0, San Jose CA, USA*
- [99] Wong, K.F., Watney, J.B. and Hammes-Schiffer, S. (2004) *J. Phys. Chem.* 108, 12231-12241 .
- [100] *Accelrys Software Inc., (2012) Discovery Studio Modelling Environment, Release 3.5, San Diego, CA, USA.*

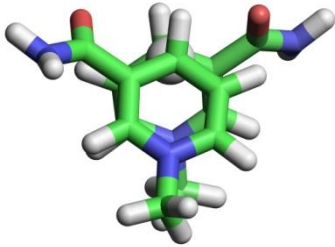
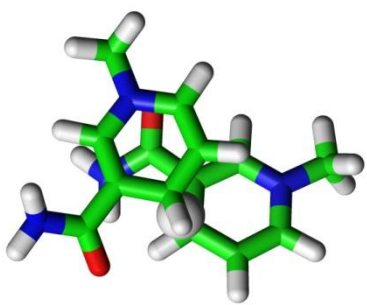
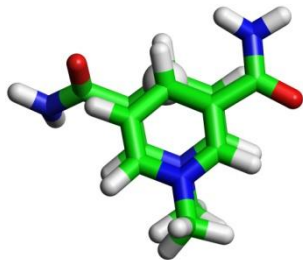
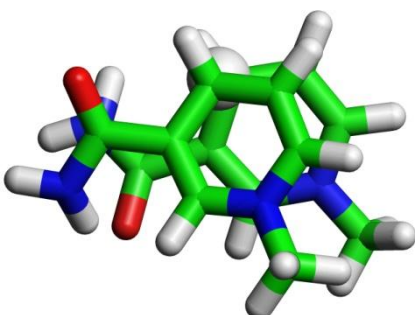
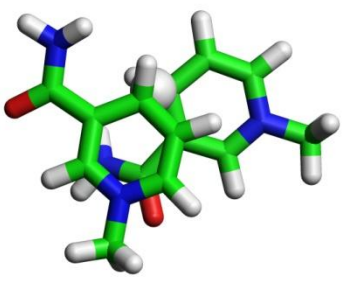
Supplementary Information: Transition State Structures

Rank	Designation	Structure	Energy / Hartree	Relative Energy / kJ mol ⁻¹
1	C2-EE-OO (+)		-913.971715621	0
2	C2-XE-OO (+)		-913.969320553	+6.29
3	C2-XX-ON (+)		-913.968457004	+8.56

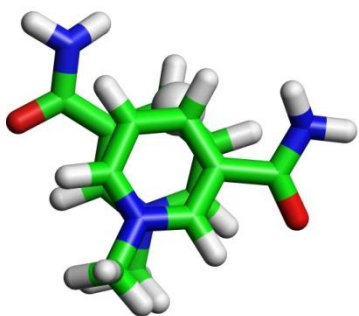
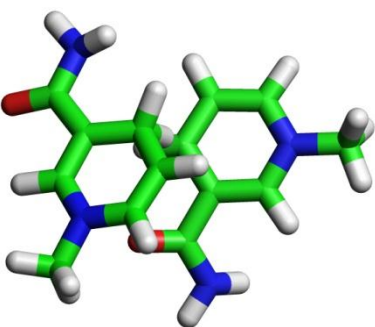
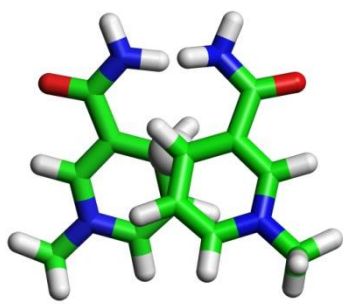
TRANSITION STATE STRUCTURES

Rank	Designation	Structure	Energy / Hartree	Relative Energy / kJ mol ⁻¹
4	C2-EE-NN (+)		-913.968014502	+9.72
5	C2-EE-NO (+)		-913.967815194	+10.24
6	C2-EE-ON (+)		-913.967728808	+10.47
7	C2-EX-NO (+)		-913.967567226	+10.89

TRANSITION STATE STRUCTURES

Rank	Designation	Structure	Energy / Hartree	Relative Energy / kJ mol ⁻¹
8	(E)-C2V-OO		-913.967508069	+11.05
9	C2-XE-ON (+)		-913.967401869	+11.33
10	(E)-C2V-NO		-913.965113129	+17.33
11	(Z)-C2V-ON		-913.962885776	+23.18
12	C2-XE-NN (-)		-913.962746403	+23.55

TRANSITION STATE STRUCTURES

Rank	Designation	Structure	Energy / Hartree	Relative Energy / kJ mol ⁻¹
13	(E)-C2V-NN		-913.962377037	+24.52
14	C2-XE-NO (-)		-913.961803878	+26.02
15	C2-XX-NN (+)		-913.959844943	+31.17

APPENDIX: TABLE S1 - NICOTINAMIDE HYDRIDE TRANSFER TRANSITION STATES

Name	C2-H / Å	C1-H / Å	Energy / Hartree	TV frequency	θ / Degrees (Initial)	θ / Degrees (Final)
c2-ee-oo (+)	1.34	1.34	-913.971715621	-767.3075	-179.98	112.19
c2-ee-oo (+)	1.34	1.34	-913.971715621	-767.3940	97.18	112.19
c2-ee-oo (-)	1.34	1.34	-913.971715620	-767.3261	-179.98	-112.17
c2-ee-oo (-)	1.34	1.34	-913.971715575	-767.3632	177.03	-112.15
c2-xe-oo (+)	1.38	1.29	-913.969320553	-694.3529	-1.14	120.67
c2-xe-oo (-)	1.38	1.29	-913.969320551	-694.2191	-179.98	-120.67
c2-xe-oo (-)	1.38	1.29	-913.969320544	-694.1887	-179.89	-120.68
c2-xe-oo (+)	1.38	1.29	-913.969320523	-694.2203	179.05	120.67
c2-xe-oo (-)	1.38	1.29	-913.969320517	-694.2193	-179.98	-120.67
c2-xe-oo (-)	1.38	1.29	-913.969320468	-694.2255	-120.63	-120.67
c2-xx-on (+)	1.24	1.48	-913.968457004	-528.1576	73.50	86.80
c2-xx-on (+)	1.24	1.49	-913.968456968	-528.0060	73.50	86.84
c2-xx-on (-)	1.24	1.48	-913.968456912	-528.2627	-174.84	-86.85
c2-ee-nn (+)	1.36	1.36	-913.968014502	-998.1557	77.49	77.48
c2-ee-no	1.34	1.36	-913.967815194	-889.1855	-174.86	121.46

APPENDIX: TABLE S1 - NICOTINAMIDE HYDRIDE TRANSFER TRANSITION STATES

Name	C2-H / Å	C1-H / Å	Energy / Hartree	TV frequency	θ / Degrees (Initial)	θ / Degrees (Final)
(+)						
c2-ee-no (-)	1.34	1.36	-913.967815185	-889.2305	-179.98	-121.46
c2-ee-no	1.34	1.36	-913.967815169	-889.0667	97.15	121.48
(+)						
c2-ee-on	1.33	1.38	-913.967728808	-891.8612	97.17	97.16
(+)						
c2-ex-no	1.38	1.30	-913.967567226	-807.4717	-179.98	63.48
(+)						
c2-ex-no (-)	1.39	1.30	-913.967567225	-807.4272	178.11	-63.50
(E)-c2v-oo	1.33	1.33	-913.967508069	-716.0715	-1.14	-15.86
(E)-c2v-oo	1.33	1.33	-913.967508064	-715.9531	73.50	15.83
c2-xe-on	1.35	1.34	-913.967401869	-837.2079	-179.98	79.16
(+)						
c2-xe-on	1.35	1.34	-913.967400300	-837.6670	177.72	79.10
(+)						
c2-xe-on (-)	1.35	1.34	-913.967400023	-837.7504	-79.03	-79.07
c2-xe-on	1.35	1.34	-913.967399928	-837.7338	-170.84	79.07
(+)						
c2-xe-on	1.35	1.34	-913.967399330	-837.8578	178.66	79.03
(+)						
(E)-c2v-no	1.37	1.31	-913.965113129	-745.4013	2.54	1.13
(E)-c2v-no	1.37	1.31	-913.965113102	-745.5551	-1.76	-1.11

APPENDIX: TABLE S1 - NICOTINAMIDE HYDRIDE TRANSFER TRANSITION STATES

Name	C2-H / Å	C1-H / Å	Energy / Hartree	TV frequency	θ / Degrees (Initial)	θ / Degrees (Final)
(E)-c2v-no	1.37	1.31	-913.965112906	-745.5411	-1.13	-1.13
(E)-c2v-no	1.37	1.31	-913.965050036	-739.4985	2.56	2.11
(Z)-c2v-on	1.39	1.31	-913.962885776	-824.7683	-14.82	-14.82
c2-xe-nn (-)	1.32	1.39	-913.962746403	-878.8370	-14.82	-78.74
c2-xe-nn (-)	1.32	1.39	-913.962745980	-878.6874	-14.82	-78.74
c2-xe-nn (+)	1.32	1.40	-913.962745957	-879.0808	178.64	78.72
c2-xe-nn (-)	1.32	1.39	-913.962745946	-878.7813	-78.74	-78.76
c2-xe-nn (+)	1.32	1.39	-913.962745857	-878.6524	137.25	78.76
(E)-c2v-nn	1.35	1.35	-913.962377037	-874.4048	2.56	-18.16
(E)-c2v-nn	1.35	1.35	-913.962376537	-874.2798	18.15	18.17
c2-xe-no (-)	1.47	1.25	-913.961803878	-588.8092	-179.98	-117.03
c2-xe-no (-)	1.47	1.25	-913.961803862	-589.1314	-78.77	-117.04
c2-xe-no (+)	1.47	1.25	-913.961803791	-589.4933	78.76	117.04
c2-xx-nn (+)	1.32	1.37	-913.959844943	-820.1746	100.80	89.10

APPENDIX: TABLE S2 - TRANSITION STATE IRC

Name	Energy /Hartree	TV frequency	Energy (IRC opt) / Hartree	TV freq (IRC opt)	Energy (opt min) / Hartree	TV freq (opt min)
c2-ee- oo (+)	-913.971715621	-767.3940	- 913.9717156 75	-767.2859	- 914.00281 0489	MINIMUM
c2-ee- oo (+)	-913.971715621	-767.3075			- 914.00280 8112	MINIMUM
c2-ee- oo (-)	-913.971715620	-767.3261	- 913.9717156 22	-767.3150	failed	unopt
c2-ee- oo (-)	-913.971715575	-767.3632	- 913.0028111 24		- 914.00281 1124	MINIMUM
c2-xe- oo (+)	-913.969320553	-694.3529	- 913.9693205 58	-694.3503	- 913.98653 9402	MINIMUM
c2-xe- oo (-)	-913.969320551	-694.2191	- 913.9693205 33	-693.6833	- 913.98653 9350	MINIMUM
c2-xe- oo (-)	-913.969320544	-694.1887			- 913.98653 9360	MINIMUM
c2-xe- oo (-)	-913.969320523	-694.2203	- 913.9693205 76	-694.3696	- 913.98653 9467	MINIMUM
c2-xe-	-913.969320517	-694.2193			-	MINIMUM

APPENDIX: TABLE S2 - TRANSITION STATE IRC

Name	Energy /Hartree	TV frequency	Energy (IRC opt) / Hartree	TV freq (IRC opt)	Energy (opt min) / Hartree	TV freq (opt min)
oo (-)					913.98653	
					9334	
			-		-	
c2-xe- oo (-)	-913.969320468	-694.2255	913.9693205	-694.2668	913.98653	MINIMUM
			36		9426	
			-			
c2-xx- on (+)	-913.968457004	-528.1576	913.9684570	-527.0391	failed	unopt
			15			
			-		-	
c2-xx- on (+)	-913.968456968	-528.0060	913.9684570	-526.4679	914.00180	-7.3211
			77		8652	
			-			
c2-xx- on (-)	-913.968456912	-528.2627	913.9684569	-528.0263	failed	unopt
			07			
			-		-	
c2-ee- nn (+)	-913.968014502	-998.1557			914.00249	MINIMUM
					6364	
			-		-	
c2-ee- no (+)	-913.967815194	-889.1855	913.9678155	-888.8972	914.00639	MINIMUM
			71		9739	
			-			
c2-ee- no (-)	-913.967815185	-889.2305	913.9678151	-891.6841		
			71			
			-			
c2-ee- no (+)	-913.967815169	-889.0667				

APPENDIX: TABLE S2 - TRANSITION STATE IRC

Name	Energy /Hartree	TV frequency	Energy (IRC opt) / Hartree	TV freq (IRC opt)	Energy (opt min) / Hartree	TV freq (opt min)
c2-ee- on (+)	-913.967728808	-891.8612				
c2-ex- no (+)	-913.967567226	-807.4717	913.9675665 65	-810.3741	914.00265 4944	MINIMUM
c2-ex- no (-)	-913.967567225	-807.4272			914.00257 3974	MINIMUM
(E)-c2v- oo	-913.967508069	-716.0715	913.9675081 14	-716.1182	913.98654 5911	MINIMUM
(E)-c2v- oo	-913.967508064	-715.9531	913.9675080 99	-715.9225	913.98654 5892	MINIMUM
c2-xe- on (+)	-913.967401869	-837.2079			913.99801 1832	MINIMUM
c2-xe- on (+)	-913.967400300	-837.6670	913.9673969 41	-838.4035		
c2-xe- on (-)	-913.967400023	-837.7504	913.9674011 69	-836.5529	913.99801 2374	MINIMUM
c2-xe- on (+)	-913.967399928	-837.7338	913.9674051	-835.4431		

APPENDIX: TABLE S2 - TRANSITION STATE IRC

Name	Energy /Hartree	TV frequency	Energy (IRC opt) / Hartree	TV freq (IRC opt)	Energy (opt min) / Hartree	TV freq (opt min)
			69			
c2-xe- on (+)	-913.967399330	-837.8578	- 913.9674039	-836.1514	- 913.99792	-7.3818
			98		9254	
(E)-c2v- no	-913.965113129	-745.4013	- 913.9651166	-743.9389	- 913.99815	MINIMUM
			43		0999	
(E)-c2v- no	-913.965113102	-745.5551			- 914.00184	MINIMUM
					2485	
(E)-c2v- no	-913.965112906	-745.5411	- 913.9651154	-742.8092	- 913.98952	MINIMUM
			14		0947	
(E)-c2v- no	-913.965050036	-739.4985	- 913.9650497	-740.1232	- 914.00257	MINIMUM
			09		4103	
(Z)-c2v- on	-913.962885776	-824.7683	- 913.9628846	-832.4052	- 914.00257	MINIMUM
			09		3904	
c2-xe- nn (-)	-913.962746403	-878.8370			failed	unopt
					-	
c2-xe- nn (-)	-913.962745980	-878.6874			914.00049	-8.2331
					4109	
c2-xe-	-913.962745957	-879.0808	-	-879.0645	failed	unopt

APPENDIX: TABLE S2 - TRANSITION STATE IRC

Name	Energy /Hartree	TV frequency	Energy (IRC opt) / Hartree	TV freq (IRC opt)	Energy (opt min) / Hartree	TV freq (opt min)
nn (+)			913.9627453			
			32			
			-		-	
c2-xe- nn (-)	-913.962745946	-878.7813	913.9627469	882.5594	914.00048	-7.5249
			54		8446	
			-		-	
c2-xe- nn (+)	-913.962745857	-878.6524	913.9627466	-878.6774	914.00048	-7.0699
			70		8911	
			-		-	
(E)-c2v- nn	-913.962377037	-874.4048	913.9623784	-875.2467	914.00072	MINIMUM
			18		2919	
			-		-	
(E)-c2v- nn	-913.962376537	-874.2798			913.99443	MINIMUM
					7826	
			-		-	
c2-xe- no (-)	-913.961803878	-588.8092	913.9618038		914.00639	MINIMUM
			47		3326	
			-		-	
c2-xe- no (-)	-913.961803862	-589.1314	913.9618038	-590.4458	914.00639	MINIMUM
			47		3326	
			-		-	
c2-xe- no (+)	-913.961803791	-589.4933	913.9618037	-588.4115	914.00639	MINIMUM
			79		0696	
			-		-	
c2-xx- nn (+)	-913.959844943	-820.1746	913.9598449	-819.5152	913.98643	MINIMUM

APPENDIX: TABLE S2 - TRANSITION STATE IRC

Name	Energy /Hartree	TV frequency	Energy (IRC opt) / Hartree	TV freq (IRC opt)	Energy (opt min) / Hartree	TV freq (opt min)
			77		3973	

APPENDIX: TABLE S3

Ring Conformation	MM Layer	Reactant Energy / Hartree
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.201044009387
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.028120828809
Dist-XE-OO (Dimer)	Rigid/Uncharged	-912.957917519463
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.042405908163
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.171246591594
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.116486264418
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.143083615628
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.260997710427
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.224565034300
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.111268149977
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.201676991448
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.133976666631
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.254586407993
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.265692724680
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.239811451430
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.294342727576
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.222782732808
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.292313801916
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.337051009918
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.294697934452
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.319149649227
Dist-XE-OO (Dimer)	Rigid/Uncharged	

APPENDIX: TABLE S3

Ring Conformation	MM Layer	Reactant Energy / Hartree
Dist-XE-OO (Dimer)	Rigid/Charged	-918.026683920824
Dist-XE-OO (Dimer)	Rigid/Charged	-918.892006405879
Dist-XE-OO (Dimer)	Rigid/Charged	-919.619173112224
Dist-XE-OO (Dimer)	Rigid/Charged	-918.992947626980
Dist-XE-OO (Dimer)	Rigid/Charged	-918.695942970206
Dist-XE-OO (Dimer)	Rigid/Charged	-918.735497635258
Dist-XE-OO (Dimer)	Rigid/Charged	-918.583386646799
Dist-XE-OO (Dimer)	Rigid/Charged	-918.281594527859
Dist-XE-OO (Dimer)	Rigid/Charged	-918.345441649796
Dist-XE-OO (Dimer)	Rigid/Charged	-918.521451372157
Dist-XE-OO (Dimer)	Rigid/Charged	-918.271294814664
Dist-XE-OO (Dimer)	Rigid/Charged	-918.259616383769
Dist-XE-OO (Dimer)	Rigid/Charged	-918.197786438828
Dist-XE-OO (Dimer)	Rigid/Charged	-917.988478969641
Dist-XE-OO (Dimer)	Rigid/Charged	-918.034982876737
Dist-XE-OO (Dimer)	Rigid/Charged	-917.966117123794
Dist-XE-OO (Dimer)	Rigid/Charged	-917.693952709644
Dist-XE-OO (Dimer)	Rigid/Charged	-917.866714687769
Dist-XE-OO (Dimer)	Rigid/Charged	-917.840028607352
Dist-XE-OO (Dimer)	Rigid/Charged	-917.651134295648
Dist-XE-OO (Dimer)	Rigid/Charged	-917.368152025600
Dist-XE-OO (Dimer)	Rigid/Charged	-917.870047364081

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Dist-XE-OO (Dimer)	Flex/Charged	-918.320497927263
Dist-XE-OO (Dimer)	Flex/Charged	-919.180442755355
Dist-XE-OO (Dimer)	Flex/Charged	-919.887134431103
Dist-XE-OO (Dimer)	Flex/Charged	-919.297299452919
Dist-XE-OO (Dimer)	Flex/Charged	-918.964038934373
Dist-XE-OO (Dimer)	Flex/Charged	-919.036140596506
Dist-XE-OO (Dimer)	Flex/Charged	-918.890850160415
Dist-XE-OO (Dimer)	Flex/Charged	-918.611144595555
Dist-XE-OO (Dimer)	Flex/Charged	-918.655787514076
Dist-XE-OO (Dimer)	Flex/Charged	-918.845254044467
Dist-XE-OO (Dimer)	Flex/Charged	-918.611827970701
Dist-XE-OO (Dimer)	Flex/Charged	-918.620801960215
Dist-XE-OO (Dimer)	Flex/Charged	-918.522791516970
Dist-XE-OO (Dimer)	Flex/Charged	-918.343627211577
Dist-XE-OO (Dimer)	Flex/Charged	-918.469945429182
Dist-XE-OO (Dimer)	Flex/Charged	-918.337949408821
Dist-XE-OO (Dimer)	Flex/Charged	-918.070669212829
Dist-XE-OO (Dimer)	Flex/Charged	-918.245853234442
Dist-XE-OO (Dimer)	Flex/Charged	
Dist-XE-OO (Dimer)	Flex/Charged	-918.004888162268
Dist-XE-OO (Dimer)	Flex/Charged	-917.874325411861
Dist-XE-OO (Dimer)	Flex/Charged	-918.236031967010

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.702626905246
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.590329690615
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.516695278005
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.611429565433
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.668647275885
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.613175096259
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.640146595720
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.696723311099
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.655515680567
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.728580874769
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.672773243648
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.682764398293
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.722247500465
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.717161874839
Dist-XE-ON (Dimer)	Flex/Uncharged	UNOPTIMISED
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.763747661447
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.735949684064
Dist-XE-ON (Dimer)	Flex/Uncharged	UNOPTIMISED
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.745229783067
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.749542266193
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.720943762024
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.766202847728

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.190331885893
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.014371024121
Prox-XE-OO (Dimer)	Rigid/Uncharged	-912.799954194665
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.076804373551
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.091989423297
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.149876178596
Prox-XE-OO (Dimer)	Rigid/Uncharged	
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.165794693502
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.226015453063
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.265904492677
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.179341001710
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.155895478669
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.236326813585
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.219375978540
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.256570408793
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.270285579990
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.163390052530
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.237413656471
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.296580803988
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.239298784115
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.296690539612
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.342710103559

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Prox-XE-OO (Dimer)	Rigid/Charged	-918.071785281804
Prox-XE-OO (Dimer)	Rigid/Charged	-919.302704082634
Prox-XE-OO (Dimer)	Rigid/Charged	-919.420926038399
Prox-XE-OO (Dimer)	Rigid/Charged	-919.112903863679
Prox-XE-OO (Dimer)	Rigid/Charged	-918.854400644362
Prox-XE-OO (Dimer)	Rigid/Charged	-918.439629744595
Prox-XE-OO (Dimer)	Rigid/Charged	
Prox-XE-OO (Dimer)	Rigid/Charged	-918.507944928646
Prox-XE-OO (Dimer)	Rigid/Charged	-918.236114371816
Prox-XE-OO (Dimer)	Rigid/Charged	-918.251817287938
Prox-XE-OO (Dimer)	Rigid/Charged	-917.943520552949
Prox-XE-OO (Dimer)	Rigid/Charged	-917.855066864558
Prox-XE-OO (Dimer)	Rigid/Charged	-918.241063413761
Prox-XE-OO (Dimer)	Rigid/Charged	-918.186487734086
Prox-XE-OO (Dimer)	Rigid/Charged	-917.487955862297
Prox-XE-OO (Dimer)	Rigid/Charged	-918.002719196101
Prox-XE-OO (Dimer)	Rigid/Charged	-917.775260467878
Prox-XE-OO (Dimer)	Rigid/Charged	-917.556749345899
Prox-XE-OO (Dimer)	Rigid/Charged	-918.085170683628
Prox-XE-OO (Dimer)	Rigid/Charged	-917.777065630445
Prox-XE-OO (Dimer)	Rigid/Charged	-917.749099587798
Prox-XE-OO (Dimer)	Rigid/Charged	-918.005931648430

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.614300653617
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.530321008276
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.382049438527
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.530307085626
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.607331860777
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.639132781298
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.634981612401
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.622479099372
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.692165822691
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.708484879911
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.677881909229
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.678132763979
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.696000913136
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.703431914188
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.678333653432
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.702515263081
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.722701121695
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.682137968249
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.709357119421
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.747973547939
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.738396085153
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.753078392430

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Prox-XE-OO (Dimer)	Flex/Charged	-918.311445072656
Prox-XE-OO (Dimer)	Flex/Charged	-919.562553697239
Prox-XE-OO (Dimer)	Flex/Charged	-919.736833903615
Prox-XE-OO (Dimer)	Flex/Charged	-919.357732074594
Prox-XE-OO (Dimer)	Flex/Charged	-919.161702608264
Prox-XE-OO (Dimer)	Flex/Charged	-918.717618030239
Prox-XE-OO (Dimer)	Flex/Charged	-918.905238718006
Prox-XE-OO (Dimer)	Flex/Charged	-918.799042998338
Prox-XE-OO (Dimer)	Flex/Charged	-918.648243986582
Prox-XE-OO (Dimer)	Flex/Charged	-918.590458262092
Prox-XE-OO (Dimer)	Flex/Charged	-918.312887473255
Prox-XE-OO (Dimer)	Flex/Charged	-918.252205669007
Prox-XE-OO (Dimer)	Flex/Charged	-918.635064416681
Prox-XE-OO (Dimer)	Flex/Charged	-918.530491239763
Prox-XE-OO (Dimer)	Flex/Charged	-917.905877953034
Prox-XE-OO (Dimer)	Flex/Charged	-918.364740815524
Prox-XE-OO (Dimer)	Flex/Charged	
Prox-XE-OO (Dimer)	Flex/Charged	-918.044771094459
Prox-XE-OO (Dimer)	Flex/Charged	-918.397961198253
Prox-XE-OO (Dimer)	Flex/Charged	-918.101399253656
Prox-XE-OO (Dimer)	Flex/Charged	-918.076867729624
Prox-XE-OO (Dimer)	Flex/Charged	-918.381189082743

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.127966002769
Prox-XE-OO (Dimer)	Rigid/Uncharged	-912.886249684303
Prox-XE-OO (Dimer)	Rigid/Uncharged	-912.966514429459
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.019823834074
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.105198253289
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.019813716199
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.001448075488
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.047525580960
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.150712395012
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.136927401089
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.146374455011
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.151549012019
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.161293898275
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.218716570888
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.210009511877
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.192903081583
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.177432083023
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.206045716402
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.219739701265
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.125292181888
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.228758302093
Prox-XE-OO (Dimer)	Rigid/Uncharged	-913.227050997562

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Prox-XE-OO (Dimer)	Rigid/Charged	-918.066811007613
Prox-XE-OO (Dimer)	Rigid/Charged	-918.507744783824
Prox-XE-OO (Dimer)	Rigid/Charged	-918.800120129392
Prox-XE-OO (Dimer)	Rigid/Charged	-918.528077753034
Prox-XE-OO (Dimer)	Rigid/Charged	-918.289571462582
Prox-XE-OO (Dimer)	Rigid/Charged	-918.274355221136
Prox-XE-OO (Dimer)	Rigid/Charged	-918.000063972439
Prox-XE-OO (Dimer)	Rigid/Charged	-918.188552336859
Prox-XE-OO (Dimer)	Rigid/Charged	-918.082683378810
Prox-XE-OO (Dimer)	Rigid/Charged	-917.758650897658
Prox-XE-OO (Dimer)	Rigid/Charged	-918.042223821728
Prox-XE-OO (Dimer)	Rigid/Charged	-917.506950413229
Prox-XE-OO (Dimer)	Rigid/Charged	-917.564550540564
Prox-XE-OO (Dimer)	Rigid/Charged	-917.577124686186
Prox-XE-OO (Dimer)	Rigid/Charged	-917.722482657067
Prox-XE-OO (Dimer)	Rigid/Charged	-917.771617840521
Prox-XE-OO (Dimer)	Rigid/Charged	-918.033710626167
Prox-XE-OO (Dimer)	Rigid/Charged	-917.675510150244
Prox-XE-OO (Dimer)	Rigid/Charged	-917.606257954606
Prox-XE-OO (Dimer)	Rigid/Charged	-917.907016829468
Prox-XE-OO (Dimer)	Rigid/Charged	-917.624495966702
Prox-XE-OO (Dimer)	Rigid/Charged	-917.413217756648

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.631528568343
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.572239595516
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.546369275325
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.579920780305
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.731998552117
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.625002202184
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.619543166000
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.678779135922
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.685192544722
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.708004950693
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.703721763425
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.731331938216
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.702189832906
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.761707420705
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.748095755328
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.751083530307
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.689477633076
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.761301939938
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.758185145453
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.731807914476
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.785333377921
Prox-XE-OO (Dimer)	Flex/Uncharged	-913.763437837893

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Prox-XE-OO (Dimer)	Flex/Charged	-918.397945296061
Prox-XE-OO (Dimer)	Flex/Charged	-918.946167073290
Prox-XE-OO (Dimer)	Flex/Charged	-919.123912457664
Prox-XE-OO (Dimer)	Flex/Charged	-918.951836842242
Prox-XE-OO (Dimer)	Flex/Charged	-918.663907112751
Prox-XE-OO (Dimer)	Flex/Charged	-918.682582206870
Prox-XE-OO (Dimer)	Flex/Charged	-918.402659914665
Prox-XE-OO (Dimer)	Flex/Charged	-918.683271240345
Prox-XE-OO (Dimer)	Flex/Charged	-918.452147501867
Prox-XE-OO (Dimer)	Flex/Charged	-918.175625359416
Prox-XE-OO (Dimer)	Flex/Charged	-918.451586671425
Prox-XE-OO (Dimer)	Flex/Charged	-917.977408171784
Prox-XE-OO (Dimer)	Flex/Charged	-917.903484420840
Prox-XE-OO (Dimer)	Flex/Charged	-917.969522312523
Prox-XE-OO (Dimer)	Flex/Charged	-918.135468116981
Prox-XE-OO (Dimer)	Flex/Charged	-918.340604855853
Prox-XE-OO (Dimer)	Flex/Charged	-918.520505152714
Prox-XE-OO (Dimer)	Flex/Charged	-918.070372974121
Prox-XE-OO (Dimer)	Flex/Charged	-918.067119582623
Prox-XE-OO (Dimer)	Flex/Charged	-918.385101770531
Prox-XE-OO (Dimer)	Flex/Charged	-918.139577821282
Prox-XE-OO (Dimer)	Flex/Charged	-917.843518314652

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.183147472646
Dist-XE-OO (Dimer)	Rigid/Uncharged	-912.847216138666
Dist-XE-OO (Dimer)	Rigid/Uncharged	
Dist-XE-OO (Dimer)	Rigid/Uncharged	-912.974041349979
Dist-XE-OO (Dimer)	Rigid/Uncharged	-912.922729699146
Dist-XE-OO (Dimer)	Rigid/Uncharged	-912.962646883371
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.004039259771
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.074898834316
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.049514665147
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.110986543801
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.104450461595
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.124636999832
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.086110189374
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.157044552877
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.126342041339
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.193875289792
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.159690535230
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.095374019022
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.185553388087
Dist-XE-OO (Dimer)	Rigid/Uncharged	
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.171191372183
Dist-XE-OO (Dimer)	Rigid/Uncharged	-913.276300175776

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Dist-XE-OO (Dimer)	Rigid/Charged	-917.746425777247
Dist-XE-OO (Dimer)	Rigid/Charged	-919.035867502926
Dist-XE-OO (Dimer)	Rigid/Charged	-918.768897786492
Dist-XE-OO (Dimer)	Rigid/Charged	-919.467905391684
Dist-XE-OO (Dimer)	Rigid/Charged	-918.588052747287
Dist-XE-OO (Dimer)	Rigid/Charged	-918.487210082257
Dist-XE-OO (Dimer)	Rigid/Charged	-918.804842947670
Dist-XE-OO (Dimer)	Rigid/Charged	-918.069265656470
Dist-XE-OO (Dimer)	Rigid/Charged	-918.458833180916
Dist-XE-OO (Dimer)	Rigid/Charged	-918.174366570705
Dist-XE-OO (Dimer)	Rigid/Charged	-917.711338380449
Dist-XE-OO (Dimer)	Rigid/Charged	-917.938776892923
Dist-XE-OO (Dimer)	Rigid/Charged	-918.169473116056
Dist-XE-OO (Dimer)	Rigid/Charged	-918.032471755677
Dist-XE-OO (Dimer)	Rigid/Charged	-917.834870600417
Dist-XE-OO (Dimer)	Rigid/Charged	-917.928372491529
Dist-XE-OO (Dimer)	Rigid/Charged	-917.456210925024
Dist-XE-OO (Dimer)	Rigid/Charged	-918.160588704506
Dist-XE-OO (Dimer)	Rigid/Charged	-917.963856122703
Dist-XE-OO (Dimer)	Rigid/Charged	-917.600937877731
Dist-XE-OO (Dimer)	Rigid/Charged	-917.568069134830
Dist-XE-OO (Dimer)	Rigid/Charged	-917.796679850706

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.651128482866
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.420096210094
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.562975533310
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.496375331108
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.514727066911
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.572652025100
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.661046282300
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.604095842703
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.598510579719
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.629657486363
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.612953372461
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.668842977909
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.615059431427
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.698792925830
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.735353406812
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.683755608189
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.700910254007
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.723363595481
Dist-XE-OO (Dimer)	Flex/Uncharged	
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.713027077768
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.712499076608
Dist-XE-OO (Dimer)	Flex/Uncharged	-913.769877156664

APPENDIX: TABLE S3 - XE-OO ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Dist-XE-OO (Dimer)	Flex/Charged	-918.101169333913
Dist-XE-OO (Dimer)	Flex/Charged	-919.424289426157
Dist-XE-OO (Dimer)	Flex/Charged	-919.230032367236
Dist-XE-OO (Dimer)	Flex/Charged	-919.818487096312
Dist-XE-OO (Dimer)	Flex/Charged	-918.930158910828
Dist-XE-OO (Dimer)	Flex/Charged	-918.893736758159
Dist-XE-OO (Dimer)	Flex/Charged	-919.283236859947
Dist-XE-OO (Dimer)	Flex/Charged	-918.478576483029
Dist-XE-OO (Dimer)	Flex/Charged	-918.903590306385
Dist-XE-OO (Dimer)	Flex/Charged	-918.502180474540
Dist-XE-OO (Dimer)	Flex/Charged	-918.164216585976
Dist-XE-OO (Dimer)	Flex/Charged	-918.436991569303
Dist-XE-OO (Dimer)	Flex/Charged	-918.559225111007
Dist-XE-OO (Dimer)	Flex/Charged	-918.432954363684
Dist-XE-OO (Dimer)	Flex/Charged	-918.333914569428
Dist-XE-OO (Dimer)	Flex/Charged	-918.240930769543
Dist-XE-OO (Dimer)	Flex/Charged	-918.022763173255
Dist-XE-OO (Dimer)	Flex/Charged	-918.710225486258
Dist-XE-OO (Dimer)	Flex/Charged	-918.396120788376
Dist-XE-OO (Dimer)	Flex/Charged	-918.105912096236
Dist-XE-OO (Dimer)	Flex/Charged	-918.074079756966
Dist-XE-OO (Dimer)	Flex/Charged	-918.155286157883

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.287326306076
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.073593977788
Dist-XE-ON (Dimer)	Rigid/Uncharged	-912.913692680357
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.124800636611
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.134329268338
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.059530285524
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.155790346970
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.163365735369
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.187513484661
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.260142792589
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.164153556807
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.209625353635
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.320375694744
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.259751267677
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.240761610085
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.313005354747
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.262312890998
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.180127333955
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.282425731698
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.286468700818
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.222607755551
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.352910345759

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Dist-XE-ON (Dimer)	Rigid/Charged	-917.806014349239
Dist-XE-ON (Dimer)	Rigid/Charged	-918.973218054738
Dist-XE-ON (Dimer)	Rigid/Charged	-919.342448312388
Dist-XE-ON (Dimer)	Rigid/Charged	-919.185994259248
Dist-XE-ON (Dimer)	Rigid/Charged	-918.509031589200
Dist-XE-ON (Dimer)	Rigid/Charged	-918.645009554196
Dist-XE-ON (Dimer)	Rigid/Charged	-918.923062067388
Dist-XE-ON (Dimer)	Rigid/Charged	-918.486381669434
Dist-XE-ON (Dimer)	Rigid/Charged	-918.513853808543
Dist-XE-ON (Dimer)	Rigid/Charged	-918.475652432784
Dist-XE-ON (Dimer)	Rigid/Charged	-918.440817546119
Dist-XE-ON (Dimer)	Rigid/Charged	-918.072885410223
Dist-XE-ON (Dimer)	Rigid/Charged	-918.051518443403
Dist-XE-ON (Dimer)	Rigid/Charged	-918.129921925483
Dist-XE-ON (Dimer)	Rigid/Charged	-918.079878238719
Dist-XE-ON (Dimer)	Rigid/Charged	-918.079853739378
Dist-XE-ON (Dimer)	Rigid/Charged	-918.125586179564
Dist-XE-ON (Dimer)	Rigid/Charged	-918.158522424604
Dist-XE-ON (Dimer)	Rigid/Charged	-918.214719894638
Dist-XE-ON (Dimer)	Rigid/Charged	-917.636397894337
Dist-XE-ON (Dimer)	Rigid/Charged	-918.068832238909
Dist-XE-ON (Dimer)	Rigid/Charged	-917.901457017978

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.702626905246
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.590329690615
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.516695278005
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.611429565433
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.668647275885
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.613175096259
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.640146595720
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.696723311099
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.655515680567
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.728580874769
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.672773243648
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.682764398293
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.722247500465
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.717161874839
Dist-XE-ON (Dimer)	Flex/Uncharged	
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.763747661447
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.735949684064
Dist-XE-ON (Dimer)	Flex/Uncharged	
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.745229783067
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.749542266193
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.720943762024
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.766202847728

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Dist-XE-ON (Dimer)	Flex/Charged	-918.082142394877
Dist-XE-ON (Dimer)	Flex/Charged	-919.254720107360
Dist-XE-ON (Dimer)	Flex/Charged	-919.635706136993
Dist-XE-ON (Dimer)	Flex/Charged	-919.449109722449
Dist-XE-ON (Dimer)	Flex/Charged	-918.885919604835
Dist-XE-ON (Dimer)	Flex/Charged	-918.967585414053
Dist-XE-ON (Dimer)	Flex/Charged	-919.177671950656
Dist-XE-ON (Dimer)	Flex/Charged	-918.793308396003
Dist-XE-ON (Dimer)	Flex/Charged	-918.805816954548
Dist-XE-ON (Dimer)	Flex/Charged	-918.798054708615
Dist-XE-ON (Dimer)	Flex/Charged	-918.845364604376
Dist-XE-ON (Dimer)	Flex/Charged	-918.453808228640
Dist-XE-ON (Dimer)	Flex/Charged	-918.371170271062
Dist-XE-ON (Dimer)	Flex/Charged	-918.501685834636
Dist-XE-ON (Dimer)	Flex/Charged	-918.500581371728
Dist-XE-ON (Dimer)	Flex/Charged	
Dist-XE-ON (Dimer)	Flex/Charged	-918.498451573316
Dist-XE-ON (Dimer)	Flex/Charged	
Dist-XE-ON (Dimer)	Flex/Charged	-918.567641789426
Dist-XE-ON (Dimer)	Flex/Charged	-918.010025855620
Dist-XE-ON (Dimer)	Flex/Charged	-918.512939253600
Dist-XE-ON (Dimer)	Flex/Charged	-918.297756710964

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.287924742603
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.053397790483
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.061459480867
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.077928795007
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.114757822903
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.108689685710
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.172103675653
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.222348184746
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.248262897650
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.231380645109
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.307409936175
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.266354184996
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.287159743948
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.294333487662
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.270089649080
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.283974339972
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.309788368023
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.291743730821
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.286708966979
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.341533810646
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.277238327728
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.342858776141

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Prox-XE-ON (Dimer)	Rigid/Charged	-917.832021775921
Prox-XE-ON (Dimer)	Rigid/Charged	-919.235581614071
Prox-XE-ON (Dimer)	Rigid/Charged	-918.977200857801
Prox-XE-ON (Dimer)	Rigid/Charged	-919.146785515577
Prox-XE-ON (Dimer)	Rigid/Charged	-918.766589499494
Prox-XE-ON (Dimer)	Rigid/Charged	-918.821851315111
Prox-XE-ON (Dimer)	Rigid/Charged	-918.821515226026
Prox-XE-ON (Dimer)	Rigid/Charged	-918.249403110994
Prox-XE-ON (Dimer)	Rigid/Charged	
Prox-XE-ON (Dimer)	Rigid/Charged	-918.513951392500
Prox-XE-ON (Dimer)	Rigid/Charged	-918.052238856748
Prox-XE-ON (Dimer)	Rigid/Charged	-917.881498008068
Prox-XE-ON (Dimer)	Rigid/Charged	-918.379586566815
Prox-XE-ON (Dimer)	Rigid/Charged	-917.897595548236
Prox-XE-ON (Dimer)	Rigid/Charged	-918.010299808114
Prox-XE-ON (Dimer)	Rigid/Charged	-918.420258504756
Prox-XE-ON (Dimer)	Rigid/Charged	-917.930991079122
Prox-XE-ON (Dimer)	Rigid/Charged	-917.931529933305
Prox-XE-ON (Dimer)	Rigid/Charged	-918.510266274730
Prox-XE-ON (Dimer)	Rigid/Charged	-917.681969528092
Prox-XE-ON (Dimer)	Rigid/Charged	-917.913096917531
Prox-XE-ON (Dimer)	Rigid/Charged	-918.031312746416

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.706064957218
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.589437302464
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.560432223843
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.552048182211
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.642254753466
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.595401816123
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.633117446705
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.723629842283
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.668207924518
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.679208289640
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.773779334504
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.649602558533
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.731596704827
Prox-XE-ON (Dimer)	Flex/Uncharged	
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.706377825371
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.745237226237
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.749686892060
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.731608604537
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.753180438238
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.765312536743
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.692808018620
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.794582790813

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Reactant Energy / Hartree
Prox-XE-ON (Dimer)	Flex/Charged	-918.102371973604
Prox-XE-ON (Dimer)	Flex/Charged	-919.495203115826
Prox-XE-ON (Dimer)	Flex/Charged	-919.218902335252
Prox-XE-ON (Dimer)	Flex/Charged	-919.363983026683
Prox-XE-ON (Dimer)	Flex/Charged	-919.146207627833
Prox-XE-ON (Dimer)	Flex/Charged	-919.086701966165
Prox-XE-ON (Dimer)	Flex/Charged	-919.073888164845
Prox-XE-ON (Dimer)	Flex/Charged	-918.603049970484
Prox-XE-ON (Dimer)	Flex/Charged	-918.363349601773
Prox-XE-ON (Dimer)	Flex/Charged	-918.750024127951
Prox-XE-ON (Dimer)	Flex/Charged	-918.384680523957
Prox-XE-ON (Dimer)	Flex/Charged	-918.175473799433
Prox-XE-ON (Dimer)	Flex/Charged	-918.662940061828
Prox-XE-ON (Dimer)	Flex/Charged	-918.355877647956
Prox-XE-ON (Dimer)	Flex/Charged	-918.344957091218
Prox-XE-ON (Dimer)	Flex/Charged	-918.710958781928
Prox-XE-ON (Dimer)	Flex/Charged	-918.336335389778
Prox-XE-ON (Dimer)	Flex/Charged	-918.277796909591
Prox-XE-ON (Dimer)	Flex/Charged	-918.749046763490
Prox-XE-ON (Dimer)	Flex/Charged	-918.067085289879
Prox-XE-ON (Dimer)	Flex/Charged	-918.272709404550
Prox-XE-ON (Dimer)	Flex/Charged	-918.353148594903

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.185052708920
Prox-XE-ON (Dimer)	Rigid/Uncharged	-912.890332476066
Prox-XE-ON (Dimer)	Rigid/Uncharged	-912.978362247595
Prox-XE-ON (Dimer)	Rigid/Uncharged	-912.897232308238
Prox-XE-ON (Dimer)	Rigid/Uncharged	-912.987684691191
Prox-XE-ON (Dimer)	Rigid/Uncharged	-912.981629523415
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.071503685879
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.020519877069
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.047615815535
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.113509741945
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.200717376362
Prox-XE-ON (Dimer)	Rigid/Uncharged	
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.152249133925
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.113987703116
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.043045603393
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.193895985615
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.083949579110
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.067831220228
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.161775931201
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.222121705664
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.107731119395
Prox-XE-ON (Dimer)	Rigid/Uncharged	-913.233227602799

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Prox-XE-ON (Dimer)	Rigid/Charged	-918.016480400912
Prox-XE-ON (Dimer)	Rigid/Charged	-919.251411158641
Prox-XE-ON (Dimer)	Rigid/Charged	-918.585932896140
Prox-XE-ON (Dimer)	Rigid/Charged	-918.939065875654
Prox-XE-ON (Dimer)	Rigid/Charged	-918.421457714093
Prox-XE-ON (Dimer)	Rigid/Charged	-918.151852529445
Prox-XE-ON (Dimer)	Rigid/Charged	-918.611035408249
Prox-XE-ON (Dimer)	Rigid/Charged	-917.848946549823
Prox-XE-ON (Dimer)	Rigid/Charged	-918.395123999898
Prox-XE-ON (Dimer)	Rigid/Charged	-917.663768736309
Prox-XE-ON (Dimer)	Rigid/Charged	-917.720915428306
Prox-XE-ON (Dimer)	Rigid/Charged	-918.366179056465
Prox-XE-ON (Dimer)	Rigid/Charged	-917.868733242094
Prox-XE-ON (Dimer)	Rigid/Charged	-917.854800366350
Prox-XE-ON (Dimer)	Rigid/Charged	-917.719029887418
Prox-XE-ON (Dimer)	Rigid/Charged	-917.781915378147
Prox-XE-ON (Dimer)	Rigid/Charged	-917.923819824668
Prox-XE-ON (Dimer)	Rigid/Charged	-917.768237930908
Prox-XE-ON (Dimer)	Rigid/Charged	-917.718968957700
Prox-XE-ON (Dimer)	Rigid/Charged	-917.577591231808
Prox-XE-ON (Dimer)	Rigid/Charged	-917.807215051647
Prox-XE-ON (Dimer)	Rigid/Charged	-917.651485255644

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.623929053874
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.471691834532
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.553846159881
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.460618913183
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.545372296637
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.598502593702
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.593952384855
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.592720787221
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.612786858338
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.710802986631
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.714710603017
Prox-XE-ON (Dimer)	Flex/Uncharged	
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.715291051059
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.670745117520
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.592221294313
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.724004350000
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.665994065276
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.626413562941
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.668430416428
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.761848108039
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.643069556747
Prox-XE-ON (Dimer)	Flex/Uncharged	-913.770567755563

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Prox-XE-ON (Dimer)	Flex/Charged	-918.299029352878
Prox-XE-ON (Dimer)	Flex/Charged	-919.576465835640
Prox-XE-ON (Dimer)	Flex/Charged	-918.919127761769
Prox-XE-ON (Dimer)	Flex/Charged	-919.262209189692
Prox-XE-ON (Dimer)	Flex/Charged	-918.761991191932
Prox-XE-ON (Dimer)	Flex/Charged	-918.535994923755
Prox-XE-ON (Dimer)	Flex/Charged	-918.929947704405
Prox-XE-ON (Dimer)	Flex/Charged	-918.275391401308
Prox-XE-ON (Dimer)	Flex/Charged	-918.801969513248
Prox-XE-ON (Dimer)	Flex/Charged	-918.116694112057
Prox-XE-ON (Dimer)	Flex/Charged	-918.175361335356
Prox-XE-ON (Dimer)	Flex/Charged	-918.753822508806
Prox-XE-ON (Dimer)	Flex/Charged	-918.421929768943
Prox-XE-ON (Dimer)	Flex/Charged	-918.272767721099
Prox-XE-ON (Dimer)	Flex/Charged	-918.105800925179
Prox-XE-ON (Dimer)	Flex/Charged	-918.203979004301
Prox-XE-ON (Dimer)	Flex/Charged	-918.341842519767
Prox-XE-ON (Dimer)	Flex/Charged	-918.271212400743
Prox-XE-ON (Dimer)	Flex/Charged	-918.096717482041
Prox-XE-ON (Dimer)	Flex/Charged	-918.019105937911
Prox-XE-ON (Dimer)	Flex/Charged	-918.274603686693
Prox-XE-ON (Dimer)	Flex/Charged	-918.068180412431

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.243762294509
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.007753540390
Dist-XE-ON (Dimer)	Rigid/Uncharged	-912.919346545850
Dist-XE-ON (Dimer)	Rigid/Uncharged	-912.893940442458
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.094968169287
Dist-XE-ON (Dimer)	Rigid/Uncharged	-912.958732738663
Dist-XE-ON (Dimer)	Rigid/Uncharged	-912.978588391039
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.123834757537
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.057788035741
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.117439047417
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.119441561646
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.108906854112
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.191275099070
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.203381769771
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.130584126303
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.162732987409
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.156766926289
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.133948786644
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.232252275880
Dist-XE-ON (Dimer)	Rigid/Uncharged	
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.202150797934
Dist-XE-ON (Dimer)	Rigid/Uncharged	-913.226872970095

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Dist-XE-ON (Dimer)	Rigid/Charged	-918.270074031035
Dist-XE-ON (Dimer)	Rigid/Charged	-918.851217693848
Dist-XE-ON (Dimer)	Rigid/Charged	-919.331439797720
Dist-XE-ON (Dimer)	Rigid/Charged	-918.808742022246
Dist-XE-ON (Dimer)	Rigid/Charged	-918.601216614580
Dist-XE-ON (Dimer)	Rigid/Charged	-918.845250476860
Dist-XE-ON (Dimer)	Rigid/Charged	-918.892753189500
Dist-XE-ON (Dimer)	Rigid/Charged	-918.357982643212
Dist-XE-ON (Dimer)	Rigid/Charged	-918.410797536904
Dist-XE-ON (Dimer)	Rigid/Charged	-917.953415529170
Dist-XE-ON (Dimer)	Rigid/Charged	-918.188339257347
Dist-XE-ON (Dimer)	Rigid/Charged	-918.170664302235
Dist-XE-ON (Dimer)	Rigid/Charged	-917.779784149835
Dist-XE-ON (Dimer)	Rigid/Charged	-918.151842987041
Dist-XE-ON (Dimer)	Rigid/Charged	-917.886774090020
Dist-XE-ON (Dimer)	Rigid/Charged	-917.694169241311
Dist-XE-ON (Dimer)	Rigid/Charged	-917.939667223409
Dist-XE-ON (Dimer)	Rigid/Charged	-917.730703617144
Dist-XE-ON (Dimer)	Rigid/Charged	-917.588592171104
Dist-XE-ON (Dimer)	Rigid/Charged	-917.972168284475
Dist-XE-ON (Dimer)	Rigid/Charged	-917.769131907753
Dist-XE-ON (Dimer)	Rigid/Charged	-917.493932828244

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.674394272266
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.554103374000
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.458524098825
Dist-XE-ON (Dimer)	Flex/Uncharged	
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.659138513920
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.599361439689
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.614750567319
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.701013905239
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.634099123972
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.708872875231
Dist-XE-ON (Dimer)	Flex/Uncharged	
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.685363843683
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.738546315550
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.694934831603
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.708173618607
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.714328560210
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.676363868131
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.671080099010
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.776698301415
Dist-XE-ON (Dimer)	Flex/Uncharged	
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.751298289480
Dist-XE-ON (Dimer)	Flex/Uncharged	-913.777137433996

APPENDIX: TABLE S4 -XE-ON ONIOM

Ring Conformation	MM Layer	Product Energy / Hartree
Dist-XE-ON (Dimer)	Flex/Charged	-918.617966233401
Dist-XE-ON (Dimer)	Flex/Charged	-919.175471857774
Dist-XE-ON (Dimer)	Flex/Charged	-919.668984135183
Dist-XE-ON (Dimer)	Flex/Charged	-919.188694408604
Dist-XE-ON (Dimer)	Flex/Charged	-919.113460016351
Dist-XE-ON (Dimer)	Flex/Charged	-919.300793662703
Dist-XE-ON (Dimer)	Flex/Charged	-919.248350947053
Dist-XE-ON (Dimer)	Flex/Charged	-918.892262870855
Dist-XE-ON (Dimer)	Flex/Charged	-918.830777897824
Dist-XE-ON (Dimer)	Flex/Charged	-918.420485799623
Dist-XE-ON (Dimer)	Flex/Charged	-918.710409696312
Dist-XE-ON (Dimer)	Flex/Charged	-918.626553894782
Dist-XE-ON (Dimer)	Flex/Charged	-918.221898339952
Dist-XE-ON (Dimer)	Flex/Charged	-918.686219140501
Dist-XE-ON (Dimer)	Flex/Charged	-918.334054746515
Dist-XE-ON (Dimer)	Flex/Charged	-918.122454658693
Dist-XE-ON (Dimer)	Flex/Charged	-918.461421825663
Dist-XE-ON (Dimer)	Flex/Charged	-918.212512986147
Dist-XE-ON (Dimer)	Flex/Charged	-918.078808815447
Dist-XE-ON (Dimer)	Flex/Charged	-918.417647903896
Dist-XE-ON (Dimer)	Flex/Charged	-918.246514364398
Dist-XE-ON (Dimer)	Flex/Charged	-917.992703702000