Simulation Studies of Aromatic Amine Dehydrogenase bound Phenylethylamine Analogues

A thesis submitted to The University of Manchester for the degree of MPhil in the faculty of Engineering and Physical Sciences

2011

Philip N A Peartree

School of Chemical Engineering and Analytical Sciences
# Table of Contents

Table of Contents ................................................................. 2  
Table of Figures ........................................................................ 4  
Abstract .................................................................................... 6  
Declaration ................................................................................ 7  
Copyright Statement ............................................................... 8  
Acknowledgements .................................................................... 10  
The Author ................................................................................ 11  
1. Introduction and Background ............................................. 12  
   1.1. Preface ............................................................................. 13  
   1.2. History of Enzyme Catalysis ........................................... 13  
   1.3. Transition state theory .................................................... 14  
      1.3.1. Transition state theory in Enzymes ........................... 16  
      1.3.2. Semi-Classical TST .................................................... 17  
   1.4. Beyond TST ..................................................................... 19  
      1.4.1. Quantum Tunneling .................................................. 20  
      1.4.2. Bell Correction ........................................................ 22  
      1.4.3. VTST/MT ................................................................. 22  
   1.5. Dynamics ....................................................................... 23  
   1.6. Vibrationally Assisted Tunneling .................................... 23  
      1.6.1. Failure of the Bell Model ....................................... 23  
      1.6.2. Vibrationally Enhanced Ground State Tunneling ....... 25  
      1.6.3. Kuznetsov and Ulstrup .......................................... 27  
   1.7. Aromatic Amine Dehydrogenase .................................... 28  
      1.7.1. Structure and Chemistry ....................................... 28  
      1.7.2. Tunneling and Dynamics ................................. 31  
   1.8. Aims and Perspectives .................................................... 32  
2. Methods .............................................................................. 35  
   2.1. Preface ............................................................................. 36  
   2.2. Molecular Mechanics .................................................... 36  
      2.2.1. Force-fields ............................................................ 37  
      2.2.2. CHARMM Force-field .......................................... 39  
   2.3. Quantum Mechanics ..................................................... 43  
      2.3.1. Semi-Empirical Methods ...................................... 44  
      2.3.2. Density Functional Methods ................................ 46  
   2.4. Computational Techniques ........................................... 48  
      2.4.1. Energy Minimization ............................................. 48  
      2.4.2. Molecular Dynamics ............................................. 49  
      2.4.3. Analysis Methods .................................................. 52  
3. Quantum Mechanical Studies of Phenylethylamine Analogues 55  
   3.1. Preface ............................................................................. 56  
   3.2. Systems Modelled ......................................................... 57  
   3.3. Bond Lengths ............................................................... 58  
   3.4. Bond Dissociation Energies ........................................... 60  

### Table of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Schematic of the effect of ZPE on classical activation energies (Ranaghan and Mulholland)</td>
<td>18</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>The effect of Tunneling on reactions with a static barrier (Basran et al. 1999)</td>
<td>21</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>AADH heterotetramer, showing the α subunits in blue, β subunits in red, and the location of the active sites in green. (Johannisen 2006)</td>
<td>29</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>TTQ Prosthetic Group (Johannisen 2006)</td>
<td>29</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Proposed reaction mechanism for AADH reacting with tryptamine (Pang et al.)</td>
<td>30</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Definition of an Improper torsion (Brooks et al. 1983)</td>
<td>40</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Diagram indicating transferring hydrogen in AADH-bound phenylethylamine</td>
<td>56</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Core system</td>
<td>57</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Extended system</td>
<td>57</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>PM3 Basic System, log K&lt;sub&gt;obs&lt;/sub&gt; data taken from experiment</td>
<td>59</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>PM3 Extended System, log K&lt;sub&gt;obs&lt;/sub&gt; data taken from experiment</td>
<td>59</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>B3LYP/6-311++g(d,p) Basic System, log K&lt;sub&gt;obs&lt;/sub&gt; data taken from experiment</td>
<td>59</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>PM3 Basic System</td>
<td>60</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>BDE vs KIE for PM3 and B3LYP/6-311++g(d,p)</td>
<td>61</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>BDE vs KIE for extended system (at PM3 theory level)</td>
<td>62</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>BDE vs Activation enthalpy for PM3 theory level</td>
<td>63</td>
</tr>
<tr>
<td>Figure 3.11</td>
<td>BDE Vs Activation enthalpy with B3LYP/6-311++g(d,p)</td>
<td>63</td>
</tr>
<tr>
<td>Figure 3.12</td>
<td>BDE vs log k&lt;sub&gt;obs&lt;/sub&gt; at PM3 theory level</td>
<td>64</td>
</tr>
<tr>
<td>Figure 3.13</td>
<td>BDE vs log k&lt;sub&gt;obs&lt;/sub&gt; at B3LYP/6-311++g(d,p)</td>
<td>64</td>
</tr>
<tr>
<td>Figure 3.14</td>
<td>Vibration frequency vs activation enthalpy at PM3 Level</td>
<td>65</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Backbone Cα RMSD A: Fluorophenylethylamine, B: Phenylethylamine, C: Methylphenylethylamine, D: Hydroxyphenylethylamine</td>
<td>71</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Diagram indicating bond of rotation for dihedral measured in figure 4.3</td>
<td>72</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>CH2-NT-CI2-CI1 Dihedral A) From AH Dimer, B) From BD Dimer</td>
<td>73</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>D-H-A Angle A) from AH Dimer B) From BD Dimer</td>
<td>74</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Hydrogen- acceptor bond length A) from AH Dimer B) from BD Dimer</td>
<td>75</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>SD Distance against ΔΔH‡ A: AH Dimer, B: BD Dimer, ΔΔH‡ taken from experiment</td>
<td>76</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>SD Distance against ΔH‡ A: AH Dimer, B: BD Dimer</td>
<td>77</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Donor carbon spectral density</td>
<td>79</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>Spectral Density of donor carbon expressed using effective frequencies</td>
<td>80</td>
</tr>
<tr>
<td>Figure 4.10</td>
<td>Effective frequency against temperature dependence of the KIE for AH Dimer</td>
<td>80</td>
</tr>
<tr>
<td>Figure 4.11</td>
<td>Effective frequency against temperature dependence of the KIE for BD dimer. ΔΔH‡ from experiment</td>
<td>81</td>
</tr>
</tbody>
</table>
Abstract


A series of para-substituted phenylethylamine analogues bound to the enzyme aromatic amine dehydrogenase have been simulated using quantum mechanical electronic structure calculations and molecular mechanical molecular dynamics simulations. Trends have been verified connecting bond dissociation energy (and thus driving force) to observed rate constants and activation enthalpy. Trends have been identified in connecting statistics drawn from molecular dynamics simulations and the temperature dependence of the kinetic isotope effect, notably that as the temperature dependence of the kinetic isotope effect increases the flexibility of the promoting vibration decreases. This is explained as being more effected by thermal energy put into the system, and therefore more affected by temperature.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning
Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see
http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on presentation of Theses
Acknowledgements

I would like to thank Mike Sutcliffe for, well, putting up with me is the best way I can put it! I know I’ve not exactly been the ideal student these last few years, and without the calm guidance of a great supervisor, I probably would have fallen of the rails many months ago. As well as Mike I would like to thank my co-supervisor of sorts, Linus. Somehow, he managed to find the time between all his work to check-up on me, and no doubt without his “gentle” cajoling the work represented in this thesis would be of a much worse standard.

Kamilla, Jiayun and all the people who have passed through the lab in the last half-decade, without all your help, I’d probably still be working out how to use SSH clients, and the vagaries of molecular dynamics. Special thanks go to Kamilla for expressing an unshaking belief that I would one day “get it finished”. Thanks as well go the kineticists that have provided many of the data used in this thesis, especially Sam, who’s put up with my often widely off-the-mark ideas with good humour.

I would also like to extend my gratitude to my friends, in particular Dan and Jim for always keeping on at me about how my work was going. It may have made for some interesting nights at the pub, but it always kept me thinking about my work at those occasions when it was furthest from my mind.

My parents also deserve thanks for their emotional, and in some cases financial help. Last of all, but certainly not least of all, my long suffering wife Clare. Throughout all the problems we’ve encountered in the last few years (and there have been many!) she has remained at my side, supporting me financially, emotionally and mentally. It’s not been easy, but it’s over!
The Author

Philip Peartree has been awarded a BSc in Chemistry, and also has an MSc in Cheminformatics. Previous research experience includes an industrial project researching ink adhesion to polymer surfaces, and an MSc dissertation detailing simulations of Total Electron Yield X-Ray Absorption Spectroscopy using custom coded software. Research interests include the application of computer simulation to novel problems, and the effect expanding computer power has on scientific research.
1. Introduction and Background
1.1. Preface

Enzymes are essential for life, since because of their tremendous catalytic power, in some cases giving a rate enhancement of over $10^{23}$ of the equivalent solution reactions (Lad et al. 2003), they make chemical processes vital for life accessible at milder conditions than are attainable by organisms. However it is only recently that in-depth studies of them have been carried out. This section gives a brief overview of the current level of knowledge of enzyme catalysed reactions, and outlines the nature of the work represented in this thesis.

1.2. History of Enzyme Catalysis

The first hypothesis of how enzymes worked that was widely accepted was that of Emil Fischer (Fischer 1894). This stated that the enzyme behaved in the manner of a lock, and its substrate acted as a key, in such a way only the correctly shaped substrate would fit into, and therefore be reacted upon by a specific enzyme. This was the first attempt to address the principle of enzyme specificity. The hypothesis stated that enzymes achieve their phenomenal rate enhancements due to the enzyme bringing reacting groups on both itself and its substrate closer together than in the equivalent solution reaction.

This idea was then extended by J.B.S. Haldane, who stated that “The key doesn’t quite fit the lock but exercises a certain strain on it” (Haldane 1930). This described the current understanding that enzymes work on a range of substrates that have similar but not identical shape. This led to the induced-fit principle (Koshland 1973; Bennett and Steitz 1978). This gives the generally accepted principle that enzymes can and do change shape to accommodate different substrates, within a constrained range of movement, which better explains the different substrates that an enzyme can act upon. In tandem with these insights into how enzymes bind with their substrates, in 1935 Henry Eyring postulated Transition State Theory to describe gas phase reactions. This led Linus Pauling to formulate his theory of enzyme-transition
state complimentarity. This theory states that enzymes preferentially bind to the transition state of a chemical reaction, and this gives a rate enhancement by reducing the energy of the transition state complex relative to the reactant state, thus reducing the activation barrier compared to the equivalent solution reaction.

1.3. **Transition state theory**

The most common description of the kinetics of a chemical reaction is in terms of transition state theory (TST). In this framework, reactants and products of a reaction are described as being separated by a static potential energy barrier. At the top of this barrier is the transition state, a high energy conformation of which demonstrates both product and reactant like attributes. In applying this framework to real reactions, it is more appropriate to consider the reaction in overall phase space, rather than on an artificial reaction coordinate. Given this phase space, the potential energy barrier now exists on a potential energy surface, with the reactant and products at minima on this surface, with a path connecting the two minima called the minimum energy path (MEP). The transition state exists at the highest energy point on this path (i.e. at a saddle point on the potential energy surface). TST states that the rate of reaction can be described by the Arrhenius equation:

\[
k = Ae^{-\frac{E_a}{RT}}
\]  

(1.1)

where \( k \) is the rate, \( E_a \) is the activation energy, \( R \) is the universal gas constant, \( T \) is the temperature and \( A \) is the Arrhenius prefactor. This prefactor is specific to the reaction/system being studied. This equation shows clearly that the rate of a chemical reaction in the TST framework is governed by two properties, the difference in energy between the reactant and the transition state (the activation energy) and the overall energy of the system (represented by the temperature). This makes sense, as by raising the overall temperature of the system, it makes it more probable that the system can surmount the energy barrier, and similarly, reducing this barrier will make it easier to do so.
The activation energy of the reaction under examination can be found by using the Arrhenius plot, which is a plot of this equation:

\[ \ln(k) = \ln(A) - \frac{E_a}{R} \frac{1}{T} \]  

(1.2)

By plotting \( \ln(k) \) versus \( 1/T \), a straight line plot should be achieved, where the gradient is equal to \(-E_a/R\) and the y axis intercept is equal to the natural logarithm of the Arrhenius prefactor.

According to TST, an enzyme achieves a rate enhancement by reducing the activation energy relative to the same reaction in solution. The transition state complementarity description of this is that the enzyme will stabilize the transition state by binding to it more tightly than to the reactant. This can take the guise of positioning of charged residues to match the electrostatic properties of the transition state or by providing a favourable solvation environment for the TS, the stabilization is found from the differences in the reactant and transition state.

This method is appealing for its simplicity and for many years has been the benchmark in enzyme kinetics, it does not fully encompass all the methods of reducing the difference between the reactant and transition state conformations. The other obvious method is, instead of lowering the energy of the TS, to raise the energy of the reactant (Jencks 1975; Wu et al. 2000). Further to this oversight, the Arrhenius equation does not take into account the entropy of the reaction components. A better representation can be found by using the Gibbs free energy instead of the activation energy, which includes both enthalpic and entropic components:

\[ \Delta G = \Delta H - T \Delta S \]  

(1.3)
where $G$ is the Gibbs free energy, $H$ is the enthalpy and $S$ is the entropy. The rate of the reaction can be found using equation 1.4.

$$k = \frac{k_B T}{h} e^{\frac{\Delta G^\ddagger}{RT}}$$

(1.4)

$\Delta G^\ddagger$ is the free energy barrier that has replaced the activation energy barrier from the Arrhenius equation and $k_B$ is the Boltzmann constant. The activation enthalpy and entropy can be found from the Eyring equation:

$$\ln \left( \frac{k}{T} \right) = \ln \left( \frac{k_B}{h} \right) + \frac{\Delta S^\ddagger}{R} - \frac{\Delta H^\ddagger}{RT}$$

(1.5)

1.3.1. Transition state theory in Enzymes

The theory of a single transition state and single energy barrier is suitable for describing simple reactions such as small molecules in the gas-phase, but for descriptions of more complex reactions, such as those involving enzymes, a more rigorous description is necessary. Reactions involving enzymes can have multiple reactant conformations on the PES and each of these conformations will have its own minimum energy path to the product, and therefore have its own transition state. In this case, the words “transition state” are still used to describe the concept, but instead of referring to a single conformation, it refers to an ensemble of conformations which represent the transition state. On the multi-dimensional PES, this ensemble of structures is represented by a hypersurface which separates the reactant and product states (Pechukas 1981).
1.3.2. Semi-Classical TST

The system referred to above is known as classical TST, which does not take into account any quantum mechanical effects that may be present. This “idealism” works well when any transferring particle is heavy enough for these quantum mechanical effects to be negligible. When the transferred particle is light, as in the case of a hydrogen atom or proton, then the quantum mechanical effects cannot be ignored. Specifically, the Heisenberg uncertainty principle becomes important at this mass scale.

1.3.3. Zero point energy

The Heisenberg uncertainty principle states that it is impossible to know with absolute certainty the position and momentum of a particle. This situation gives rise to the idea of zero point energy (ZPE). The principle of zero point energy is that a particle’s energy must lie above the absolute possible minimum due to the uncertainty from the Heisenberg uncertainty principle. The ZPE can be calculated by \( \frac{1}{2}h\nu_x \) where \( \nu_x \) is the vibrational frequency of the particle X and \( h \) is Planck’s constant. As is obvious, when changing the isotope of a transferring particle, the mass changes, and therefore so does the vibrational frequency and thus the zero point energy. The reaction rate in a semi-classical formalism can be found using:

\[
k_x = A_x e^{\frac{\Delta G^i - \frac{1}{2}h\nu_x}{RT}}
\]  

As is obvious from this equation, the zero point energy has the effect of reducing the activation energy, and therefore increasing the rate. As different isotopes have different ZPE, then they also have different effective ZPE, and thus exhibit different rates. This effect is known as the kinetic isotope effect (KIE).
Figure 1.1 Schematic of the effect of ZPE on classical activation energies, showing that tritium transfer has a higher activation free energy than deuterium transfer which has a higher activation free energy than protium (Ranaghan and Mulholland)

1.3.4. Kinetic Isotope Effects

The kinetic isotope effect is a valuable probe into the kinetics of light particle transfer reactions. It is commonly invoked in hydrogen transfer reactions (either proton, neutral hydrogen or hydride) because hydrogen is a light particle which means that quantum mechanical effects are not negligible and also because hydrogen has readily available isotopes (protium, deuterium and tritium). The KIE for a hydrogen transfer reaction is given by:

\[
\frac{k_H}{k_D} = \left( \frac{A_H}{A_D} \right) e^{\left( \frac{\nu_H - \nu_D}{kT} \right)} \tag{1.7}
\]

Semi-classical TST makes certain predictions regarding the KIE in hydrogen transfer reactions. Firstly, the ratio of protium:deuterium rates is approximately 7 when the reaction is performed at 300K, secondly the ratio of the Arrhenius prefactors is near unity, this being due to the fact that as temperature approaches infinity the KIE approaches 0, finally the rates for the three hydrogen isotopes are connected by the Swain-Schaad relationship (Swain et al. 1958), which states that 

\[ k_H/k_T = (k_D/k_T)^{3.3}. \]
1.3.5. Variational Transition State Theory

TST predicts that the TS is the saddle point on the MEP, and this can be simply found using computational algorithms, since it has defined mathematical properties. The rate can then be calculated by the difference between this saddle point and the reactant, $\Delta G^\dagger$. This works well for systems with few degrees of freedom at low temperatures. In this situation, all reactive paths across the PES pass through the narrow area near the TS, and once on the product side of the TS, the reaction is unlikely to pass back to the reactant side. Large systems however, possess many degrees of freedom, and usually perform their reactions at physiological temperatures. In these circumstances, the energy at the saddle point ceases to be the sole contributor to the rate, as entropy becomes more important. This can lead to the rate calculated from TST being an overestimate of the true rate.

This gives rise to variational transition state theory. In practice the true rate of reaction is found by using the free energy maximum on the MEP, which corresponds to the biggest bottleneck to the reaction, then the position of the “dividing surface” (c.f. hypersurface) is adjusted to minimise the rate of reaction. The TS found using this method is known as the generalised transition state.

1.4. Beyond TST

However, transition state theory cannot be used to explain certain enzyme-catalysed reactions. In these reactions the kinetic isotope effect can be as high as 80, as found with the enzyme Soybean Lipoxygenase-1 (SLO-1) (Rickert and Klinman 1999). As TST predicts a protium/deuterium KIE of approximately 7, there must be another phenomenon forming part of the reaction. The consensus is that to fully explain this, it is necessary to invoke quantum tunneling as part of the reaction mechanism.
1.4.1. Quantum Tunneling

Quantum tunneling refers to the phenomenon of a particle transitioning through an energy state that is forbidden by classical mechanics. This arises from the principle of wave-particle duality. Heisenberg’s uncertainty principle states that if a pair of properties is measured, both cannot be known to a high precision. Interpreted, it means that if the exact energy of a particle is known, its position in space cannot be known to a high degree of accuracy. Instead its location can be given as a probability of the particle existing in different locations. In quantum physics the particle is described by its wavefunction. The indeterminacy of a particle’s position gives rise to a finite probability that a particle can be found on the far side of an energy barrier that classically it does not possess the energy to surmount. A particle’s wave-like properties are described by the de Broglie wavelength:

\[ \lambda = \frac{h}{mv} \]  

Since mass, \( m \), is a divisor in equation 1.8 then it is obvious that the smaller the mass of a particle, the longer the de Broglie wavelength, and therefore the higher the uncertainty in its position. Interpreted in quantum tunneling, this means a lighter particle can tunnel through a wider barrier.

Biochemical electron transfers have long been known to occur by quantum tunneling, which is unsurprising given an electron has a very low mass. Electron transfers are known to occur over distances of up to 25 Å. As a proton is much heavier, the tunneling distance for such a particle will be much less. For an idealized rectangular energy barrier of height \( V \) and width \( \ell \), the tunneling probability can be described using equation 1.9:

\[ P \propto \exp \left[ -\frac{2(\sqrt{2mV})}{h} \right] \]  

(1.9)

---

20
If the probability of tunneling is kept the same as in electron transfer, the transfer distance for a proton is 0.58 Å (Basran et al. 1999), although a more realistic curved energy barrier is likely to increase this distance. As this distance is similar to the width of the energy barrier to hydrogen transfer, it is likely that quantum tunneling plays a significant role in such reactions. Several attempts to introduce quantum mechanical tunneling contributions into the description of enzyme kinetics have been made. Shown in Figure 1.2: The effect of Tunneling on reactions with a static barrier show the different regimes on a sample reaction profile and their effects on a sample Eyring plot. (Basran et al. 1999) is a diagram illustrating the different regimes that have been postulated, and their effects on the temperature dependence of the hydrogen transfer rate.

Regime I shows a traditional TST reaction description, with the reaction passing over the barrier. This type of reaction demonstrates no tunneling contribution. The first attempts to introduce tunneling were made by Bell.

Figure 1.2: The effect of Tunneling on reactions with a static barrier show the different regimes on a sample reaction profile and their effects on a sample Eyring plot. (Basran et al. 1999)
1.4.2. Bell Correction

The Bell correction model (Bell 1980) describes regimes II and III in Figure 1.2, in that it describes tunneling occurring just below the transition state. This gives rise to a temperature dependent KIE, due to the reaction progressing up the reaction barrier before tunneling is invoked. The Bell model allows reactions with shallow energy barriers to be treated classically, i.e. without invoking quantum tunneling, but in cases where the reaction barrier is steeper, tunneling is invoked below the classical transition state.

The Bell model exhibits certain characteristic deviations from semi-classical TST (Kim and Kreevoy 1992); an inflated KIE (above 7), a ratio of Arrhenius pre-factors of less than 1, and an elevated Swain-Schaad exponent (Swain et al. 1958). In addition, the curvature shown in regime III is rarely seen, due to these regions being inaccessible at physiologically relevant temperatures.

1.4.3. VTST/MT

The best theoretical model to address reactions that cannot be explained with the Bell correction to TST is the multidimensional tunneling correction to VTST as developed by Truhlar and coworkers (Truhlar et al. 1996). The multidimensional tunneling corrections allow for the potential for tunneling to occur at all regions on the potential energy surface, rather than just below the transition state as in the Bell corrected TST model.

The rate at temperature $T$ is determined using the following equation:

$$k(T) = \kappa(T)k^{VTST}(T)$$

(1.10)
where \( k_{VTST}(T) \) is the rate calculated from VTST calculations, and \( \kappa(T) \) is known as the transmission coefficient. This transmission coefficient is calculated for multiple points on the MEP. The transmission coefficient for a configuration of energy \( E \) is given as:

\[
\kappa(T) = \beta \int_{0}^{\infty} d\beta P^G(E) \exp(\beta[V[s^*(T)] - E])
\]  

Where \( \beta \) is the potential energy at the variational transition state, and \( P^G \) is the probability of ground state tunneling. The energy where \( \kappa(T) \) is maximised is known as the representative tunneling energy.

\textbf{1.5. Dynamics}

The theoretical treatments outlined above assume the enzyme to be a static entity, but this is known to be untrue. Careri proposed the idea of the fluctuating enzyme (Careri 1974), and since then, research has sought to ally the idea of a flexible molecule with the catalytic effects of enzymes. Karplus and McCammon first suggested that enzyme motions can affect the chemistry that they perform (Brooks et al. 1983). Due to enzymes being flexible, there is no reason to discount the possibility that this dynamic nature may have an effect on the inclusion of tunneling in enzyme kinetic descriptions.

\textbf{1.6. Vibrationally Assisted Tunneling}

Many enzyme-catalysed hydrogen transfer reactions have been found to deviate from simple tunneling correction models. These reactions seem to fit a model where dynamics can affect the tunneling probability.

\textbf{1.6.1. Failure of the Bell Model}

The Bell correction model assumes that tunneling occurs just below the classical transition state, but as is shown in section 1.4.1 the de Broglie wavelength of a
proton is such that it is conceivable that a hydrogen transfer reaction could proceed purely via tunneling, without a need to climb the barrier first. The proviso to this is that the distance the hydrogen atom transfers is shorter than the de Broglie wavelength.

It has been suggested that this could be achieved by low frequency motions of the protein (Basran et al. 1999; Knapp and Klinman 2002). This has been contested by others (Ball 2004) who argue that enzymes cannot achieve a high-energy configuration that could drive tunneling.
1.6.2. Vibrationally Enhanced Ground State Tunneling

Theoretical frameworks have been developed to describe such reactions. This has been driven by the discovery of enzyme reactions that appear to violate traditional descriptions of reaction kinetics. The first such study was that of Klinman and co-workers into thermophilic alcohol dehydrogenase (Kohen et al. 1999). They examined the steady state kinetics over a wide range of temperatures (5-65 °C). This range of temperature is not normally accessible for most enzymes, as most denature at temperatures above 40 °C.

The wide range of temperatures proved useful, as a breakpoint at 30 °C in the Arrhenius plot was found. Below this point, kinetics followed a description that was allowed under the Bell correction model. Above this temperature a ratio of Arrhenius prefactors of 2 was calculated. This is not allowed by the Bell correction model, and suggests that the level of tunneling increases as temperature increases. This leads to the conclusion that dynamics must be involved in the tunneling step, as enzyme motions increase as temperature increases.

Another enzyme requiring environmentally coupled tunneling to explain its kinetics is soybean lipoxygenase-1 (SLO). This enzyme exhibits a very high 1° KIE (>80) at room temperature and a low activation energy of the transfer event (8.8 kJ mol⁻¹). This cannot be explained by any correction model (including the Bell model) (Rickert and Klinman 1999). The entropic barrier was found to be larger than the enthalpic barrier (-TΔS‡ = 53.6 kJ mol⁻¹, ΔH‡ = 6.3 kJ mol⁻¹), the KIE was also found to be only weakly temperature dependent, this means the isotopic effects rely mainly on the entropic term, which Klinman and co-workers claim is contrary to semi-classical theory (Rickert and Klinman 1999).

Further evidence of the requirement for environmentally coupled tunneling comes from the effect of amino acid mutations on the Arrhenius prefactors (Rickert and
Klinman 1999). Mutations carried out proximal to the active site changed the
temperature dependence of the KIE (i.e. changed the prefactor ratio to 3). Distal
mutations lead to an inverse prefactor ratio (0.2). Despite these changes, the KIE
remained large (>80). This implies a fundamental change in the tunneling
coordinate, but with similar tunneling components in all systems (Knapp et al. 2002)

This behaviour, where the type of tunneling is a function of mutational position
(changing the Arrhenius prefactor ratio from >1 to <1) is reminiscent of the
thermophilic alcohol dehydrogenase previously studied by Klinman and co-workers.
The reason why the work on SLO is important is the magnitude of the KIE, because
it is so large, it precludes interpretation with a model that encompasses both
tunneling and over the barrier reactions. This type of reaction requires a model
including pure tunneling (i.e. no over the barrier transitions).

Another successful study into vibrationally enhanced tunneling was performed by
Scrutton and co-workers (Basran et al. 1999). The enzyme selected for study was
methylamine dehydrogenase (MADH). This is closely linked to aromatic amine
dehydrogenase (AADH) and has the advantage that the oxidative and reductive half-
reactions can be separated, and because of this, the transfer step can be studied
directly using stopped-flow kinetic techniques, rather than the steady-state
techniques used by Klinman et al. This reaction showed a large temperature
independent KIE of 16.8 and a high ratio of Arrhenius prefactors of 13.3. This data
is consistent with an extreme tunneling regime, and this is backed up by
computational work of Truhlar and co-workers (Alhambra et al. 2001) which
indicated that the reaction proceeds via 99% tunneling. The overall reaction
activation energies are not zero as would be expected with this level of tunneling and
a static barrier, and the energies give a strongly temperature dependent rate. This
was rationalised through the use of a dynamic free energy surface, and thus,
dynamically linked tunneling.
A number of models have been developed to explain vibrationally enhanced tunneling. The Kuznetsov and Ulstrup model has already been mentioned, but before this came the model of Bruno & Bialek (Bruno and Bialek 1992). This model introduced the concept of the fluctuating barrier width, and stated that this fluctuation modulates the transfer, and thus tunneling, distance. This model, however, does not fit most extreme tunneling reactions, such as those performed by MADH and AADH.

### 1.6.3. Kuznetsov and Ulstrup

Kuznetsov and Ulstrup presented a more complex model (Kuznetsov 1999) that incorporates varying levels of gating motion. This model made use of the Franck-Condon principle, which states that a light particle transition (such as proton transfer) can only occur at heavy atom nuclear configuration where the light particle energies in the product and reactant states match. This is a corollary of the Born-Oppenheimer approximation, where the fast motion of a light particle (e.g. transferring hydrogen) can be separated from the heavy atom motion (e.g. protein motion). Put simply, there must be degeneracy between the reactant and product wavefunctions for nuclear tunneling to occur.

#### 1.6.3.1. Passive and Active dynamics

This model was modified by Klinman and co-workers, who used it as the basis of the model of environmentally coupled hydrogen tunneling (Knapp and Klinman 2002). This model prescribes that dynamics can be split into two types, Passive and Active. Passive dynamics (reorganisation) are motions that lead to degenerate reactant and product states, i.e. a nuclear configuration that is compatible with the tunneling event. Active dynamics (gating or vibrational enhancement) are motions that enhance the tunneling probability by moving the reactant and product wells together. In this model, the Passive dynamics already described are similar to the nuclear reorganisation required by the Marcus theory of electron transfer (Marcus
and Sutin 1985). Active dynamics are similar to those of the fluctuating barrier in the Bruno and Bialek model (Bruno and Bialek 1992).

1.7. **Aromatic Amine Dehydrogenase**

Aromatic amine dehydrogenase is the enzyme under examination in this work and is the second enzyme known to possess the tryptophan tryptophylquinone (TTQ) prosthetic group, the first being MADH. These enzymes are similar in structure and mechanism; the main differences in the active site being an altered position of a phenylethylamine and the presence of an isoleucine in MADH which is shown as an asparagine in AADH. Both catalyze the deamination of primary amines; the main difference is the substrate specificity of each enzyme, with MADH preferring short chain aliphatic amines (e.g. methylamine) and AADH preferring larger aromatic amines. The system has been extensively studied (Basran et al. 2001; Roujeinikova et al. 2006; Johannissen et al. 2007; Hothi et al. 2008; Johannissen et al. 2008) and the presence of a large number of crystal structures in the protein data bank means that it is a good system on which to study the effects of quantum tunneling on enzyme catalysis, in addition to the abundance of substrates with which AADH will react, and the availability of experimental data.

1.7.1. **Structure and Chemistry**

Aromatic amine dehydrogenase was first isolated from the bacterium *Alcaligenes faecalis* by Iwaki and co-workers (Iwaki et al. 1983). The enzyme has a $\alpha_2\beta_2$ heterotetrameric structure, with the $\alpha$ subunit having a molecular weight of approximately 39 kDa and the $\beta$ subunit having a molecular weight of approximately 18 kDa. The main feature of the $\alpha$ subunit is a long tail-like loop region that hydrogen bonds onto part of the opposing $\beta$ subunit. The enzyme heterotetramer has two active sites, that lie between the $\alpha$ and $\beta$ subunits, and each active site contains a TTQ prosthetic group (see Figure 1.4)
The enzyme catalyses the oxidative deamination of a primary amine to its aldehyde and ammonia, the reductive half-reaction can be described as such:

$$RCH_2NH_2 + H_2O \rightarrow RCHO + NH_3 + 2e^- + 2H^+ \quad (1.12)$$

The oxidative half-reaction is completed with the physiological co-enzyme azurin. The mechanism of action of the enzyme was first outlined by Hyun and Davidson (Hyun and Davidson 1995) and has been examined and refined in further work (Pang et al.; Masgrau et al. 2006). The most current reaction scheme is shown in Figure 1.5
The initial steps (I and II) show substrate binding, which ends in the iminoquinone Schiff base at intermediate III. Asp128β then abstracts a proton from the donor carbon to form intermediate IV. This step is now known to involve quantum mechanical tunneling. The steps following this involve the release of the product aldehyde via hydrolysis of the imine bond.
1.7.2. Tunneling and Dynamics

AADH has been extensively used to probe hydrogen tunneling and its role in enzyme catalysis. The main reasons for this are that the reductive and oxidative half reactions can be separated, which means that the reaction can be studied via stopped flow methods as opposed to steady state methods, and that as previously mentioned (section 1.7), a number of key intermediates have been isolated and characterized using X-ray crystallography (Masgrau et al. 2006).

The reaction of tryptamine and AADH has been widely studied, and has revealed that the reaction proceeds via 99% tunneling (Masgrau et al. 2006). In addition to this, a significant rearrangement of the donor and acceptor atoms is required to attain a tunneling ready configuration. This rearrangement results in the donor carbon adopting a more planar configuration, with the transferring hydrogen perpendicular to the plane. It has been suggested that this assists the transfer (Johannissen et al. 2007), as it is a more product like configuration, and thus, once the hydrogen transfer has taken place by quantum tunneling, it causes a break in degeneracy, which means that the reverse transfer cannot occur (as would otherwise be likely as the transfer is isoenergetic).

Analysis of the dynamics of the reaction identified a motion that rotates the transferring proton towards the acceptor oxygen (O2 on Asp128β) (Masgrau et al. 2006). Further analysis of this motion characterized it as a motion intrinsic to the substrate and the frequency of these motions is approximately 165 cm⁻¹ (Johannissen et al. 2007).
1.8. Aims and Perspectives

Whilst the reaction of AADH with tryptamine has been widely studied (Pang et al.; Govindaraj et al. 1994; Hyun and Davidson 1995; Hyun and Davidson 1995; Hyun and Davidson 1995; Masgrau et al. 2006; Roujeinikova et al. 2006; Johannissen et al. 2007; Johannissen et al. 2008) the reaction with other substrates has been less widely examined. As previously discussed, AADH has a wide range of specificity, but shows a distinct preference for aromatic primary amines. Kinetic data has been made available for analogues of both phenylethylamine (Hothi et al. 2008) and benzyamine.

Studies of benzyamine analogues have been hampered by kinetic complexities introduced because of the shorter length of the aliphatic chain. This causes the aromatic ring clash sterically with Phe97α (Hothi et al. 2008). The kinetics with phenylethylamine analogues are not hampered in such a way, and a full range of kinetic data are available for these analogues. These are shown in tables 1 and 2:

<table>
<thead>
<tr>
<th>p substituent</th>
<th>( k_{\text{lim}}^\text{H} ) (s(^{-1}))</th>
<th>( k_{\text{lim}}^\text{D} ) (s(^{-1}))</th>
<th>KIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>45.6 ± 0.3</td>
<td>2.97 ± 0.03</td>
<td>15.4 ± 0.3</td>
</tr>
<tr>
<td>OH</td>
<td>412.7 ± 7.0</td>
<td>30.9 ± 0.25</td>
<td>13.4 ± 0.3</td>
</tr>
<tr>
<td>CH(_3)</td>
<td>44.7 ± 0.3</td>
<td>2.22 ± 0.01</td>
<td>20.1 ± 0.2</td>
</tr>
<tr>
<td>OCH(_3)</td>
<td>417.6 ± 11</td>
<td>26.4 ± 0.6</td>
<td>15.8 ± 0.5</td>
</tr>
<tr>
<td>NO(_2)</td>
<td>29.4 ± 0.2</td>
<td>1.77 ± 0.01</td>
<td>16.6 ± 0.2</td>
</tr>
<tr>
<td>F</td>
<td>93.1 ± 0.7</td>
<td>5.39 ± 0.04</td>
<td>17.3 ± 0.3</td>
</tr>
<tr>
<td>Cl</td>
<td>65.5 ± 0.3</td>
<td>3.25 ± 0.02</td>
<td>20.2 ± 0.2</td>
</tr>
<tr>
<td>Br</td>
<td>73.8 ± 0.3</td>
<td>3.76 ± 0.02</td>
<td>19.6 ± 0.2</td>
</tr>
</tbody>
</table>

Table 1.1: Kinetic parameters from stopped flow studies of phenylethylamine analogues (Hothi et al. 2008)
<table>
<thead>
<tr>
<th>$p$ substituent</th>
<th>$\Delta H^{\text{IH}}$ (kJ mol$^{-1}$)</th>
<th>$\Delta H^{\text{ID}}$ (kJ mol$^{-1}$)</th>
<th>$\Delta \Delta H^{\text{I}}$ (kJ mol$^{-1}$)</th>
<th>A$^\text{IH}$:A$^\text{ID}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>55.2 ± 0.8</td>
<td>54.5 ± 0.9</td>
<td>0.7 ± 1.7</td>
<td>19.6 ± 0.5</td>
</tr>
<tr>
<td>OH</td>
<td>45.2 ± 1.3</td>
<td>51.5 ± 0.4</td>
<td>6.3 ± 1.7</td>
<td>0.86 ± 0.05</td>
</tr>
<tr>
<td>CH$_3$</td>
<td>58.8 ± 0.8</td>
<td>70.0 ± 0.8</td>
<td>11.2 ± 1.6</td>
<td>0.21 ± 0.01</td>
</tr>
<tr>
<td>OCH$_3$</td>
<td>44.2 ± 1.3</td>
<td>56.1 ± 0.6</td>
<td>11.9 ± 1.9</td>
<td>0.13 ± 0.01</td>
</tr>
<tr>
<td>NO$_2$</td>
<td>67.1 ± 1.3</td>
<td>72.6 ± 1.8</td>
<td>5.5 ± 3.1</td>
<td>2.12 ± 0.12</td>
</tr>
<tr>
<td>F</td>
<td>56.7 ± 0.9</td>
<td>63.7 ± 1.1</td>
<td>7.0 ± 2.0</td>
<td>0.97 ± 0.04</td>
</tr>
<tr>
<td>Cl</td>
<td>58.6 ± 0.6</td>
<td>69.4 ± 0.9</td>
<td>10.8 ± 1.5</td>
<td>0.26 ± 0.01</td>
</tr>
<tr>
<td>Br</td>
<td>59.6 ± 0.9</td>
<td>69.1 ± 0.9</td>
<td>9.5 ± 1.8</td>
<td>0.43 ± 0.01</td>
</tr>
</tbody>
</table>

Table 1.2: Parameters from temperature dependence studies of phenylethylamine analogues (Hothi et al. 2008)

The work performed by Hothi and coworkers (Hothi et al. 2008) reports a driving force analysis of $p$-substituted phenylethylamine analogues, but does so only at a low level of theory. In addition, it does not address any dynamic effects directly. This poses several questions:

1. Does increasing the level of theory increase agreement with experiment, as would be expected, as the PM3 level of theory used is basic at best?
2. Does the change in $p$-substituent have any other effect on the hydrogen transfer step that could effect the presence of tunneling?
3. What effect does the change in $p$-substituent have on the dynamics of AADH?
4. If so, does this support the vibrationally enhanced tunneling model previously envoked (Johannissen et al. 2007)?

Answering these questions will require different techniques. Chapter 3 describes electronic structure calculations performed on the full range of phenylethylamine analogues to study any effects that these have on the reactivity of the phenylethylamine analogues with AADH.
Chapter 4 makes use of molecular dynamics to study the effect that changing $p$-substituent has on the dynamics of AADH bound phenylethylamines, attempts to determine whether any changes support the model of vibrationally enhanced tunneling employed in previous work.
2. Methods
2.1. **Preface**

Chemical processes often occur on scales much smaller than can be directly observed in physical experiments. Computational chemistry techniques, such as those detailed in this chapter, are an essential compliment to physical experiments in studying chemical and biochemical processes.

Chemical systems can be investigated in depth by using quantum mechanical methods, which examine the systems’ electronic structure, and can thus be used to examine chemical reactions and other processes that are affected by changes in electronic density. Unfortunately these techniques are computationally intensive, which makes them only applicable to smaller systems.

Larger systems can still be examined computationally, but certain simplifications are necessary. Electronic density calculations can be simplified by using approximations such as assuming that only the valance shell of electrons is important to the calculations. When examining molecular motion, such as in molecular dynamics simulations of proteins, further simplification is necessary as these systems have many degrees of freedom. This introduces a problem where the sheer volume of data makes drawing conclusions difficult. A range of analytical techniques are available to reduce the dimensionality and complexity of data to enable meaningful analysis.

2.2. **Molecular Mechanics**

Molecular mechanics techniques are lower-level techniques that can be used when either the system being studied is too large or complex to make higher level techniques accessible or when the speed of the calculation is of more importance than the accuracy.

Molecular mechanics rely on the validity of a set of assumptions. The most important of these is the Born-Oppenheimer approximation. This states that, when calculating the energy of a system, it is possible to separate the contribution of the electrons from that of the nuclei of the system. This allows a calculation of the energy of a system to be based upon its nuclear coordinates, rather than its electron
density as in the case of more complex techniques. The second assumption is that of transferability which states that a force-field (the set of parameters used to describe molecules and atomic interactions; see below) that is correctly developed upon a certain set of molecules can be applied to other molecules and produce valid results. It is this assumption which makes the use of molecular mechanics on large systems possible.

2.2.1. Force-fields

Molecular Mechanics techniques are implemented using force-fields. These consist of a set of parameters which govern the way the nuclei in a simulation interact, based upon the functional form of the particular force-field. As mentioned previously (section 2.2) the Born-Oppenheimer approximation allows the treatment of a system based upon nuclear coordinates rather than the electron density. In this form, the atoms are treated as a sphere with a radius equivalent to the atoms van der Waals radius and have a fixed charge. The bonds are treated using a Newtonian approach, which models them as springs. Through space interactions are usually treated using a combination of Lennard-Jones potentials for van der Waals interactions and Coulomb potentials. A generalised equation is shown below:

\[
\begin{align*}
\nu(r^N) &= \sum_{\text{bonds}}^{N} \frac{k_i}{2} (l_i - l_{i,0})^2 + \sum_{\text{angles}}^{N} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 + \sum_{\text{torsions}}^{N} \frac{V_i}{2} (1 + \cos(n\omega - \gamma)) \\
&+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right) + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}
\end{align*}
\]

(2.1)

Taken from (Leach 2001)

This functional form varies depending on what the force-field is designed to simulate. The bond and angle potential are governed by a modified form of Hooke’s Law, with a separate force constant for each bond type and angle type. These
Parameters are the key to the functionality of the force-field. They vary between individual versions of a certain force-field and between different force-fields.

Parameters are usually obtained from experimental or highly accurate simulations of small molecules, and contain the basic parameters of force constants for bond stretches and angle bends, as well as their equilibrium lengths and angles. They also contain periodicity of dihedral interactions and the equilibrium angles and force constants for this potential term. They also include the mass, van der Waals radius and partial charges for each atom in the force-field. Most force-fields make use of a fixed charge model, where no attempt is made to model the perturbation of a partial charge by its electrostatic neighbourhood. The reason given for this by most force-field developers is that the computational expense of modelling polarizability is prohibitive and is usually contrary to the stated goals of molecular mechanics (that is to reduce computational complexity to allow large systems to be modelled). These parameters are specific to each atom and in most cases to each chemical environment, e.g. the parameters for a carbonyl oxygen are different to that of a hydroxyl oxygen. Parameterisation is a very important part of the use of molecular mechanics, as inaccurate parameters will influence any simulations carried out using these parameters.

There are many different molecular mechanics force-fields, such as AMBER (Cornell et al. 1995) and CHARMM (Brooks et al. 1983) for simulations of proteins and biological macromolecules and UFF (Rappe 1992) and MMFF (Halgren 1996) for simulations of smaller molecules. Parameters are included in each force-field for its designed application, e.g. the CHARMM force-field comes with parameters for amino acid residues and common derivatives, such as an acylated C terminus. The forcefield used in the work represented by this thesis was the CHARMM forcefield.
2.2.2. CHARMM Force-field

The current version of the CHARMM force-field for proteins is 22, which is an all-atom force-field, i.e. unlike unified forcefields the hydrogens are explicitly included rather than included in their bonded heavy atom.

The CHARMM forcefield follows a similar form to that shown in equation 2.1 but with some differences. The basic form is as follows:

\[
E = E_b + E_\phi + E_\theta + E_{vdW} + E_{sb} + E_b + E_{\psi} + E_{\gamma} + E_{\phi}
\]  

(2.2)

Taken from (Brooks et al. 1983)

As can be seen there are a number of differences between this functional form and the generalized one shown in equation 2.1 the inclusion of two torsion terms, an additional nonbonded term, and a set of constraint terms. Presented below are expansions of each term, and definitions of each symbol.

2.2.2.1. Bonded terms

The bond and angle potentials are in the same form as those shown in the generalised functional form (equation 2.1)

The torsion set consists of two separate potentials, the proper and improper potentials. The dihedral angles are called proper torsions and the potential is given as follows:

\[
E_\phi = \sum |k_\phi| - k_\phi \cos(n\phi)
\]  

(2.3)

Charmm Torsion potential (Brooks et al. 1983)
In this form $E_\Phi$ is the energy summed across all proper torsions, $k_\Phi$ is the torsion force constant, $n$ is the periodicity of the torsion (where $n$ is 1,2,3,4 or 6) and $\Phi$ is the torsion angle between four atoms. Across a torsion there may be multiple contributions for each set of four atoms, and these contributions may have different periodicities and force constants, hence the energy for each individual torsion angle is given by the sum of all applicable torsion parameters.

The improper torsion was designed to maintain planarity of planar atoms, such as the carbonyl carbon and the carbons in aromatic rings, such as found in phenylalanine. The functional form is given as:

$$E_\omega = \sum k_\omega (\omega - \omega_0)^2$$

Charmm improper potential (Brooks et al. 1983)

The energy, $E_\omega$ is given in terms of a force constant $k_\omega$, a torsion angle $\omega$ and an equilibrium torsion angle $\omega_0$. The angle $\omega$ is given as the angle between planes ABC and BCD in the following diagram.

![Diagram of an improper torsion](image)

**Figure 2.1: Definition of an Improper torsion (Brooks et al. 1983)**
2.2.2.2. Nonbonded terms

The nonbonded interactions in the CHARMM forcefield are characterised by two potential energy terms: van der Waals and electrostatic. The van der Waals potential is given as:

\[
E_{vdW} = \sum_{excl(i,j)=1} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) sw\left( r_{ij}^{-2}, r_{on}^{-2}, r_{off}^{-2} \right)
\]  \quad (2.5)

CHARMM van der Waals potential (Brooks et al. 1983)

This follows the form of the Lennard-Jones 12-6 potential, and is calculated in for every pair of atoms except those joined to each other by a bond. The sw(…) term is a switching term which is necessary when the number atoms in a system becomes large. The switching function limits the number of pairs by distance, so as to limit the number of pairwise interactions, in an effort to make the calculations tractable. This makes the process of calculating pairwise interactions such as van der Waals interactions expensive as it scales with the square of the number of atoms.

The electrostatic potential is usually of the form shown in the generalized functional form, with a dielectric value of 1 for simulations which include explicit water molecules. This potential can be augmented with a range of switching functions that can be used to make large and complex calculations tractable on modest hardware. A more complex method of treating long-range electrostatic interactions in simulation is to use the technique of Ewald summation. This method can only be used in periodic systems and involves treating the summation of long-range electrostatic interactions in Fourier space, rather than real space. In addition, the system must be charge neutral, which is commonly achieved in simulations involving a charged molecule by adding a number of oppositely charged ions to bring the overall system charge to zero. The most frequently implemented Ewald summation algorithm is that of Particle Mesh Ewald (Darden et al. 1999); by utilizing a fixed lattice model of
charge density it allows the use of a fast fourier transform to further speed up the calculation.

2.2.2.3. Constraints

When performing minimizations or molecular dynamics, constraining the structure in some way is often useful, e.g. to prevent large atom displacements or to examine the potential energy surface of a particular motion. To accomplish this two constraint terms are built into the CHARMM forcefield, distance constraints and dihedral constraints. The distance restraints can either be rigid or of a harmonic nature:

\[
E_{c, \text{d}} = \sum K_i \left( r_i - r_{i,0} \right)^2 
\]  

CHARMM harmonic distance constraint (Brooks et al. 1983)

A rigid constraint does not allow any difference in the position for the constrained atoms. \( K_i \) is the force constant of the harmonic constraint (this is specified in the software) and \( r_i \) is the atom coordinates both currently and in the case of \( r_{i,0} \) at equilibrium. Dihedral constraints follow the following form:

\[
E_{\phi, \text{d}} = \sum K_i \left( \phi_i - \phi_{i,0} \right)^2 
\]  

CHARMM dihedral constraint (Brooks et al. 1983)

This equation follows a similar form to equation 2.6, but using the dihedral angle instead of atomic positions as a basis for the constraint. There is a special type of constraint used mainly in molecular dynamics simulations called a SHAKE (van Gunsteren and Berendsen 1977) constraint which is used to constrain the bonds to light atoms such as hydrogen, because their vibration frequency is high, due to a high force constant. The SHAKE constraint allows the simulation to have a slower
update frequency, whilst still being accurate, and thus allow a longer period of time to be investigated in a given simulation length.

2.2.2.4. **Parameterisation**

Parameters for molecules can be chosen in two ways. The simplest is by analogy. This involves finding chemically similar molecules and adapting the parameters to the novel molecule. This is made possible by the nature of parameter development, where the original parameters for proteins were developed from small molecule studies. This is the least accurate method, but for most cases the compromises made by choosing molecular mechanics dictate that at best data acquired will be qualitative.

The more complex and accurate method is by using electronic structure calculations. These are used to find equilibrium geometries and vibrational frequencies, which are then used to calculate force constants. Charges are obtained by an iterative process of adjustment to give an equal energy of interaction with a water molecule in the CHARMM forcefield as found from higher level electronic structure calculations.

2.3. **Quantum Mechanics**

Often the study of interest involves bond formation or cleavage, and in this case, a quantum mechanical method will be chosen, since most MM forcefields do not allow bonds to break and form. Quantum mechanical methods can be broken down into three general formalisms: Semi-Empirical methods, Hartree-Fock methods, and Density Functional Methods. The work represented in this thesis used two different methods, semi-empirical, and density functional.
2.3.1. Semi-Empirical Methods

The idea of any quantum mechanical method is to solve the Schrödinger equation:

\[ H\Psi = E\Psi \]  

The Schrödinger equation

Where H is Hamiltonian matrix, \( \Psi \) is wavefunction and E is energy. The wavefunction as described in section 1.4.1 is a function of the probability of finding a particle at a given point in space. In electronic structure calculations as described here these particles are electrons.

Electrons in atoms exist in defined regions of space called orbitals, each of which are governed by a wavefunction. A hydrogen atom's wavefunction can be calculated exactly, due to it containing only one electron, more complex atoms cannot be calculated exactly due to each electron affecting the other electrons. These are usually solved using Hartree-Fock theory, which describes the orbitals of atoms as spheres.

This theory becomes extremely inaccurate when addressing arrangements of multiple atoms such as molecules. In this situation, the atoms in each molecule are described by a model called Linear Combination of Atomic Orbitals (LCAO). This model constructs molecular orbitals by combining atomic orbitals. A complete solution would require an infinite set of atomic orbitals, so basis set approximation, as described by the Roothaan-Hall equations, is used.

Semi-empirical methods, like Hartree-Fock methods, use the Roothaan-Hall equations:

\[ FC = SCE \]  

The Roothaan-Hall Equation (Leach 2001)
Where $F$ is the Fock Matrix, $C$ is the Coefficient matrix, $S$ is the overlap matrix, and $E$ is the matrix of orbital energies. The main problem with Hartree-Fock methods is that obtaining the Fock matrix requires many calculations of integrals, which are computationally expensive. A way to simplify this is to address only the valence electrons, and treat the core electrons as being part of the nucleus. This still allows bonding considerations to be addressed, but reduces the computational complexity considerably.

There are a number of semi-empirical models that can be applied to calculations, but the more popular ones are Austin Model 1 (AM1) (Dewar 1985) and Parameterized Model 3 (PM3) (Stewart 1989; Stewart 1989). These techniques are similar, in that they both adapt from a previous method, MNDO (Dewar and Thiel 1977). This technique uses the zero differential overlap approximation to simplify calculations. This states that if two atomic orbitals overlap then the integral of the two electrons can be set to zero. This approximation requires that the integrals which were ignored are treated empirically. AM1 attempts to remedy some of the short-comings of the MNDO method, specifically a tendency to overestimate repulsions of atoms that are approximately in contact, i.e. separations approximately equal to the sum of their van der Waals radius.

AM1 was parameterized by hand, using chemical knowledge, and so only a limited number of compounds could be used. PM3 is a similar method, but was parameterized by an automated procedure which allowed for a much larger set of reference molecules. This leads to discrepancies between the forcefields parameters despite a similar functional form. Despite this, in general they perform equivalently well at predicting thermodynamic and structural properties. One minor problem with PM3 is that it significantly underestimates the rotational barrier of the amide bond.

Previously mentioned was the fact that PM3 is parameterized by an automated procedure. A key difference between semi-empirical techniques and Hartree-Fock
techniques, is that a semi-empirical method often has empirical parameters defined, which helps alleviate some of the potential inaccuracies brought about by the approximations used. The usual reason these techniques are used is for efficiency as the number of atoms is too great to simulate using higher level techniques. They are also used in molecular dynamics simulations, where calculations need to be performed multiple times in short succession, and hence speed is of great importance.

\subsection{Density Functional Methods}

Density functional theory is a remarkably efficient way of performing full quantum mechanical calculations. It is often used because it offers the greatest accuracy with the least computational power and thus the best results in the least time. Density functional theory makes use of a theoretical treatment that allows the user to calculate energies and other properties without taking into account each electron in a system. All that is required is the electron density at each point. The work that brought about this technique was performed by Kohn, Hohenberg and Sham (Hohenberg 1964; Kohn 1965).

The key considerations for this technique are selection of the functional, which governs how the electron density is treated, and selection of the basis set, which governs how the electron density is calculated. Density functional theory includes electron correlation, which means that even relatively fast calculations can be accurate. Electron correlation describes how an individual electron in a molecule is influenced by the other electrons in the molecule; most Hartree-Fock calculations do not include this, and those that do are computationally expensive. The addition of electron correlation increases the accuracy of the calculations compared to experimental data. The most popular basis sets for general calculations are B3LYP (Becke 1993; Stephens et al. 1994) and BP86. The calculations represented in this thesis use B3LYP. There are more specific functionals available for treatment of
certain cases, such as BH&H which gives better results when the system of interest includes areas of $\pi$-stacking.

The choice of basis set usually involves a trade-off between computational accuracy and calculation speed. A popular basis set is 6-31G*, as this is relatively accurate, but is still quick to calculate, thus allowing its application to larger systems. This basis set is of a Pople split-valence type, the 6 represents the 6 Gaussian functions used to approximate the Slater type orbital of the core orbitals, the 31 represents the fact that 3 Gaussians of one type and 1 of another are used to approximate the valence orbitals. The * represents that the basis set includes a set of polarization functions to model the polarization of the lighter atoms’ electronic orbitals. This gives increased accuracy.
2.4. Computational Techniques

Using these methods, a variety of techniques can be applied to molecular data. These techniques are used to various ends, and some are essential precursors to other techniques. In most cases the input data for these techniques are atomic coordinates, usually in Cartesian format, but in some cases internal coordinates are required. Most techniques start by requiring energy minimization of some sort.

2.4.1. Energy Minimization

Energy minimization, or geometry optimization, is a procedure that can be performed in many formalisms, such as MM, QM and QM/MM. The purpose of energy minimization is to reduce the energy of a system to the minimum possible. For large molecules such as proteins, it is never possible to reach a global, or total minimum, since most methods use first or second derivatives of the energy surface, and therefore, should the energy rise, to leave a local minimum, the algorithms judge this to be leaving a minimum energy conformation, and stop.

Energy minimization is usually required because of the inherent inaccuracy of both computational modeling techniques and experimental methods of structure determination. This inherent incompatibility can lead to bad contacts and poor geometry being “frozen” into the modelled structure. This would cause the system to possess an inflated amount of internal energy, and when performing later operations, such as molecular dynamics, this excess energy can cause structures to unfold or in some cases appear to explode.

Various algorithms are implemented depending on the type of software being used, the method used in the NAMD package (Phillips et al. 2005) is known as conjugate gradient, whereas the method commonly used in the Gaussian03 (Frisch et al. 2003) package uses the Berny optimization algorithm, which is based on the work of Schlegel(Schlegel 1982).
At a minimum energy state, the system is at an ideal equilibrium geometry at a temperature of 0 K and there should be no forces acting on any of the bonds in the system, other than those caused by zero point energy. This brings the system into a state where further calculations should not be affected by any artifacts introduced in the initial data acquisition stage.

2.4.2. Molecular Dynamics

Many processes in proteins are due to their inherent flexibility. These processes are often of great interest, as these dynamic processes are often used to explain discrepancies between current models and experimental results, as stated in (see section 1.6). Therefore studying the dynamics of the system, using molecular dynamics simulations, is very important for understanding enzyme reactions. The basic technique involves the application of Newton’s Laws to the atoms in a system, most crucially the Second Law, i.e.:

\[ F = ma \]  

**Newton’s Second Law of Motion**

So from equation 1, given the force, which can be calculated using a molecular mechanics force-field (section 2.2) and knowing the mass of the atom involved, the acceleration can be calculated. By using this acceleration, and a given time step, the velocity and position change of the atom can be calculated. Different algorithms are available which offer different methods of applying the time step and different methods of treating the thermal energy in the system, but all adhere to the same essential principle above. A procedure for performing a molecular dynamics simulation can be broken down into three distinct areas, thermalization, equilibration, and production.
2.4.2.1. Thermalization

For most molecular dynamics (MD) algorithms to function, a system requires forces and accelerations to begin with. Following minimization, the atoms have no forces and accelerations, due to being at a notional 0K, so it is necessary to assign velocities to mimic the effect of temperature on the kinetic energy of each atom. This is usually performed by a thermalization routine, which assigns kinetic energy to atoms based on a Boltzmann distribution which equal those expected at the temperature of interest. In protein simulations the temperatures requested are usually biologically relevant temperatures, and 300K is a common temperature. If this assignment was done in one step, the sudden increase in velocities in random directions can cause proteins to lose structural integrity, effectively denaturing \emph{in silico}. To avoid this situation an iterative process is used where the systems kinetic energy, and therefore temperature, is increased stepwise and in between the system is allowed to relax to that temperature. During this process, constraints are often applied to the protein structure, which prevents the increases in kinetic energy and potential energy from changing the structure artificially.

This thermalization stage gets a system to the correct kinetic and potential energy for a given temperature, but does not ensure a realistic distribution of this energy throughout the system, which can give large fluctuations of highly flexible regions. In order to allow the molecules to distribute the energy correctly through the system, a period of equilibration is performed. In the work represented in section 4 thermalization was performed for 20 ps with a temperature increase of 25 K every 300 fs, and a final temperature of 300 K.

2.4.2.2. Equilibration

Equilibration is the next stage of the molecular dynamics procedure. In this stage the velocities are left to equilibrate the required temperature. Often this is performed with certain constraints (page 42) on the system to hold it in a particular
conformation. These constraints are then gradually relaxed as the system equilibrates. The equilibration can be performed under several types of formalism, the most popular being NVE and NPT. NVE describes a formalism where the number of particles the volume of the system and the total potential energy is kept constant. This can result in a system where the potential energy remains the same, but the effective temperature of the system can change. The NPT formalism again keeps the number of particles in the simulation the same, but stabilises the pressure and temperature of the system as opposed to the volume and potential energy. This is usually achieved using a periodic box in which the system is situated. This periodic box exists in a group of identical image boxes, which allow the number of particles in the box to remain constant whilst allowing complete translational motion in and out of the box. To allow the pressure to remain the same, the box volume is allowed to shrink or grow. This pressure and temperature consistency is achieved through a thermostat and either a barostat or simulated piston dependent on the scheme selected. Once the system is correctly equilibrated, which can be assessed using the method outlined in section 2.4.2.4, a production run can be performed.

### 2.4.2.3. Production

A production run refers to the part of the simulation used for data acquisition and subsequent analysis. This is usually over a time scale of 10 – 20 ns but on modern supercomputers and cluster based architectures times of up to 100 ns are possible for real world systems (Rueda et al. 2007), indeed recent work has shown with custom hardware times longer than 1 ms are accessible for small systems (Shaw et al. 2010). The timescale chosen reflects the motions of interest and also takes into consideration computational expense for larger systems. Production is normally carried out in the same formalism as equilibration. For the work represented in section 4 production runs of 10 ns were carried out after equilibration. The simulations were monitored using the techniques in section 2.4.2.4 to ensure that the simulations were valid.
2.4.2.4. Verifying the validity of MD data

When performing molecular dynamics, it is common practice to monitor certain variables derived from the simulation. This is the means by which a system is judged to have equilibrated, when these statistical variables become stable across a given time period, usually 0.5 ns, but longer periods are sometimes used for longer simulations. The variables usually chosen are: (i) Root mean square deviation (RMSD) compared to the starting structure (the usual RMSD value for proteins is approximately 1.5 Å, but often it can be more for larger systems, the importance is stability of the value). RMSD measures the average deviation of a structure from a given reference structure, in this case the starting structure; (ii) Root mean squared fluctuations (RMSF), which gives an indication of the mobility of regions of the protein and can be directly compared to the crystallographic b-factors to ensure that the system is behaving as predicted by crystallographic data. RMSF is in many ways the counterpart to RMSD, whereas RMSD averages over all the atoms in a given system and presents the data according to each frame, RMSF averages over all the frames for a given atom, thus giving a numerical measure to the flexibility of a given region of protein; (iii) radius of gyration, which is a measure of the volume of the system, by means of the average distance of each atom from a given point (usually the centre of mass). The change of this over the course of the simulation gives information on whether the system is a constant volume throughout; (iv) values such as density in the case of constant volume or pressure simulations, stability of pressure and temperature, and potential energy. A consensus is usually sought between several of these variables in order to decide whether a simulation has reached, and remains at, equilibrium.

2.4.3. Analysis Methods

Molecular dynamics simulations can often generate large data-sets, a 10 ns simulations can contain up to $1 \times 10^7$ structures. These data can be visualised in a
 qualitative manner using packages such as VMD (Humphrey et al. 1996), but in most cases, quantitative analysis is necessary. This requires tools to perform statistical analyses on the trajectories produced. There are many packages available to perform such analysis, some are built into software packages such as CHARMM (Brooks et al. 1983) and the ptraj tool in the AMBER software suite (Case et al. 2005), or as add-ons to other software such as plug-ins for VMD and the matdcd (Gullingsrud) plug-in for MATLAB®. Most packages allow basic analysis such as plotting how a dihedral angle or bond length changes throughout a simulation, or how a statistical quantity such as RMSD changes over the course of a simulation. More complex analyses are available in some packages however, such as the ability to calculate a 2D RMS matrix.

2.4.3.1. Spectral Density Analysis

A spectral density is a representation of the vibrational frequencies present in a signal and their contributions to that signal. In this case the spectral density is of the motion of specific atoms or groups within the enzyme or substrate. The spectral density is generated from the Fourier transform of the velocity autocorrelation function. The velocity autocorrelation function is:

$$C(\tau) = \left\langle x(t) \cdot x(t + \tau) \right\rangle$$

(2.11)

The autocorrelation function $C(\tau)$ measures the self-similarity of datapoints separated by a given time interval $\tau$. This by nature is a function in the time domain. This is then Fourier transformed to the frequency domain.

The velocities are calculated from the difference between coordinates taken from a molecular dynamics simulation at regular intervals. The magnitude of the sampling interval governs the smallest frequency obtainable by the spectral density process. The spectral densities represented in this thesis were computed using software...
provided by Linus Johannissen and has been used in published works (Johannissen et al. 2007).
3. Quantum Mechanical Studies of Phenylethylamine Analogues
3.1. Preface

As discussed in section 1.8 a range of kinetic parameters have been found and derived for a range of phenylethylamine analogues. It has been shown that by approximating the driving force of the reaction, a number of trends can be seen that re-enforce the opinion the hydrogen abstraction from an aromatic amine by AADH proceeds via environmentally coupled tunneling. This chapter hopes to reproduce these trends, and provide additional validity by using higher levels of theory. In addition, other properties will be examined to provide additional insight into the reactivity series of phenylethylamine analogues with AADH. Figure 3.1 shows the location of the hydrogen that is abstracted.

![Diagram indicating transferring hydrogen in AADH-bound phenylethylamine](image)

Figure 3.1: Diagram indicating transferring hydrogen in AADH-bound phenylethylamine
3.2. Systems Modelled

Two systems were modelled and results from both are represented in this chapter. These systems were modelled from the crystal structure of phenylethylamine bound to AADH (PDB id: 2HKM). These systems are shown in figures Figure 3.2 and Figure 3.3.

Figure 3.2: Core system

Figure 3.3: Extended system
Figure 3.2 shows just the TTQ and bound phenylethylamine, whereas Figure 3.3 shows this and surrounding residues, most significantly the acceptor aspartate. These systems were modelled in GaussView and simulated in Gaussian03, at both PM3 and B3LYP/6-311++g(d,p) theory levels. The first system shown in Figure 3.2 was chosen as it contains the bare minimum of atoms to simulate the TTQ moiety, whereas the extended system shown in Figure 3.3 was chosen to include some nearby residues that are postulated to stabilise the TTQ residue. These residues are truncated to their sidechains only to minimise the computational cost of the simulations. To prevent movement of these sidechains relative to their position within the crystal structure, they were anchored by a fixed atom.

3.3. Bond Lengths

As part of the geometry optimization process, bond lengths were optimized to match the appropriate substituent. In a traditional hydrogen transfer reaction, the donor-hydrogen bond length is an indicator of bond strength, and thus the ease of transfer. Due to this it is sensible to look at this in relation to the reactivity series presented in section 1.8. The bond length measured in this section is the carbon to transferring hydrogen, as this is most likely to be an indicator of the donor-hydrogen bond strength.
Figure 3.4: PM3 Basic System, log $K_{obs}$ data taken from experiment

Figure 3.5: PM3 Extended System, log $K_{obs}$ data taken from experiment

Figure 3.6: B3LYP/6-311++g(d,p) Basic System, log $K_{obs}$ data taken from experiment
As can be seen from these plots, there is little correlation between the log of the observed rate and the bond length. This is to be expected, as the transfer step is known to occur via quantum tunneling, and the difference between bond lengths is unlikely to affect the width of the activation barrier.

![Figure 3.7: PM3 Basic System ΔH‡ taken from experiment](image)

From Figure 3.7 there is a noticeable trend between activation enthalpy and bond length, showing that activation enthalpy decreases with increasing bond length. Again, this is expected, as the increased bond length would mean less energy is required to reach a tunneling ready configuration. The correlation is not found in the extended system or in the high level B3LYP/6-311++g(d,p) calculations.

### 3.4. Bond Dissociation Energies

As described by Hothi et al (Hothi et al. 2008), if the reaction fits an environmentally coupled tunneling model, then the reaction can be described by an extension of the Marcus theory of electron transfer (Marcus and Sutin 1985). A classical validation of Marcus theory is that the reaction rate should increase with increasing driving force, and that the activation enthalpy will decrease with increasing driving force.
In a traditional electron transfer reaction, the driving force is altered by changing the reduction potential of either the donor or acceptor. The analogous method for a hydrogen transfer reaction is to change the bond dissociation energy of either the donor or acceptor. In the case with the phenylethylamine analogues, these should all have different bond dissociation energies. Therefore the bond dissociation energy of the acceptor (Asp128β) remains unchanged.

It should be noted that the calculations of bond dissociation energies shown here are approximations using the equality, where \( E^R \) is the calculated energy of the hydrogen dissociated fragment, \( E^H \) is the energy of the lone hydrogen and \( E^{R+H} \) is the energy of the fragment with the bound H:

\[
\text{BDE} = (E^R + E^H) - E^{R+H} \quad (3.1)
\]

The energies calculated must include zero point energy corrections, and therefore must be found from frequency calculations which include explicit zero point energy corrections. Bond dissociation energies are shown relative to that of unsubstituted phenylethylamine. Figure 3.8 shows the trend between relative BDE and measured KIE as calculated using the smaller system.

![Figure 3.8: BDE vs KIE for PM3 and B3LYP/6-311++g(d,p)](image)

Figure 3.8: BDE vs KIE for PM3 and B3LYP/6-311++g(d,p)
The R-squared values, which measure goodness of fit of trendlines, show that neither PM3 nor the higher B3LYP level of theory show good correlation, with values of 0.033 for the PM3 level of theory and 0.06 for the higher B3LYP level of theory. Increasing the system size has a negligible effect, increasing the R-squared value to 0.09, as shown in Figure 3.9

**Figure 3.9: BDE vs KIE for extended system (at PM3 theory level)**

The lack of correlation with KIE is not unexpected, as the KIE arises due the effects of an isotopic substitution, which by nature does not affect anything to do with the electronic structure.

More appropriate is to measure the correlation of BDE with observed rate constant and with activation enthalpy, as described above. Plotting relative BDE against should agree with the work of Hothi et al (Hothi et al. 2008).
Figure 3.10: BDE vs Activation enthalpy for PM3 theory level

The R-squared value is approximately 0.11 which is indicative that the trend is not particularly strong. Removing the outlier (which is the data for fluorophenylethylamine) improves the R-squared value to 0.92. The same result is found when simulating the larger system. Using a higher level of theory reverses the trend (Figure 3.11)

![Graph](image1)

Figure 3.11: BDE Vs Activation enthalpy with B3LYP/6-311++g(d,p)

Again, the trend is not strong, being an almost horizontal line. Plotting against the log of the observed rate constant should yield the opposite trend, i.e. a negative trend should be apparent. Figure 3.12 shows an R-squared value of approximately 0.11. Figure 3.13 shows the same data plotted against the relative BDE calculated using B3LYP level of theory. This shows a reversal of the trend shown in Figure 3.12, again showing a very low level of correlation.
3.5. Frequency Calculations

It was noted by Johannissen et al (Johannissen et al. 2007) that a promoting motion of approximately $165 \text{ cm}^{-1}$ was present in the iminoquinone moiety when studying the reaction of tryptamine with AADH. This was studied both with Spectral Density analysis of molecular dynamics trajectories, and by analysis of frequencies from electronic structure calculations. Since such frequencies are available from the calculations performed for this chapter, examining these frequencies for any shift could prove useful.
As noted by Johannissen, the when calculated using an unrestrained molecule, the frequency was found to be approximately 174 cm$^{-1}$ using B3LYP/6-31G*. Using PM3 correlation has been found between activation enthalpy and the shifts in frequency across the reactivity series. This is shown in Figure 3.14.

![Figure 3.14: Vibration frequency vs activation enthalpy at PM3 Level](image)

The R-squared value is 0.47, which is the strongest correlation found from these calculations. A trend is not present with either the extended system (most likely due to the system partially moving towards a transition state, as suggested by the presence of negative frequencies) or the higher level calculations.

### 3.6. Conclusion

Attempts have been made to replicate the work of Hothi et al (Hothi et al. 2008). The trends present in the work have not generally been identified. Trends that have been identified have, for the most part, been weakly correlated with R-squared values not rising above 0.11. Discounting fluorophenylethylamine causes R-squared values to increase to 0.92. Several trends have been identified that are novel however.
A trend has been identified showing that increasing bond donor acceptor bond length correlates with decreasing activation enthalphy. A trend has also been identified in increasing frequency of a motion identified by Johannissen as being important to the promoting vibration and increasing activation enthalpy.

Attempts to improve on the work of Hothi et al, by changing the level of theory used, and by including surrounding residues in the calculation have shown no improvement over the original work, and in most cases lose the original trend entirely. Discussions with the authors of the original work have suggested that it was known that higher levels of theory were deemed to give no discernable trends, but this was not reported in the published work.
4. Molecular Dynamics studies of Phenylethylamine Analogues
4.1. Preface

As discussed in section 1.8 the hydrogen transfer between aromatic amines and AADH has been suggested to fit a model of vibrationally enhanced tunneling. In order to verify this in relation to the reactivity series, the motions of the enzyme bound substrates should be examined. Of particular interest is the promoting motion identified in AADH by Johannissen (Johannissen et al. 2007) and whether changing the para substituent affects this motion in any way. In order to study the dynamics of AADH, unrestrained MM molecular dynamics simulations were performed, and the results are presented here.

4.2. System Setup

Four phenylethylamine analogues were chosen for simulation by molecular dynamics. These were then homology modelled from the crystal structure of AADH bound phenylethylamine (taken from Protein DataBank, id: 2HKM, not yet published) by altering the para substituent. The four chosen were phenylethylamine, p-fluorophenylethylamine, p-hydroxyphenylethylamine and p-methylphenylethylamine. These substituents were chosen for their wide spread of characteristics, both sterically and electronically. In addition, parameters were available or could easily be found by analogy for all chosen substituents.

A full heterotetramer was simulated with explicit solvation. The solvation included crystal waters present in the crystal structure. Initial system setup was carried out in the VMD package, using the automated psf builder. Each system was neutralized by adding a concentration of sodium and chloride ions to give a zero charge and an ionic concentration of 0.5M. The calculations were run with periodic boundary conditions. Solvation included a margin of 12 Å on each side of the heterotetramer. Table 4.1 shows the details of each of the systems.
Table 4.1: System details

<table>
<thead>
<tr>
<th>Substituent</th>
<th>Atoms</th>
<th>Waters</th>
<th>Chloride Ions</th>
<th>Sodium Ions</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>92442</td>
<td>25918</td>
<td>113</td>
<td>131</td>
</tr>
<tr>
<td>F</td>
<td>92442</td>
<td>25918</td>
<td>113</td>
<td>131</td>
</tr>
<tr>
<td>OH</td>
<td>92444</td>
<td>25918</td>
<td>113</td>
<td>131</td>
</tr>
<tr>
<td>Me</td>
<td>92448</td>
<td>25918</td>
<td>113</td>
<td>131</td>
</tr>
</tbody>
</table>

Simulations were started with 50000 steps of conjugate gradient minimization with the enzyme fixed. This is to allow the waters to take up natural, energetically favourable positions, as waters added by the solvation tool are added in a regular manner. The fixed atoms are then released and a further 50000 steps of conjugate gradient minimization are performed to remove bad contacts from the addition of hydrogen to the crystal structure.

Simulations were thermalized using a temperature reassignment procedure, and with atoms constrained to prevent system collapse. The protein backbone was held with a harmonic constraint of 50 kcal mol\(^{-1}\) Å\(^{-2}\) protein sidechains were held with 25 kcal mol\(^{-1}\) Å\(^{-2}\) and waters and ions were held with 10 kcal mol\(^{-1}\) Å\(^{-2}\). The reassignment protocol was started at 0 Kelvin, and temperatures were reassigned every 10 fs until a temperature of 300 Kelvin was reached.

Simulation protocols include 3 ns of equilibration with slowly relaxed constraints, and then 10 ns of production with no constraints. The timesteps used in the molecular dynamics were of the multiple timestep variety, using the n-RESPA integrator implemented in the NAMD software package. A timestep of 1 fs was used for bonded interactions, 2 fs for short through-space interactions, and 4 fs for long range interactions.

The production steps were performed in 0.5 ns sections for more efficient use of computational time. Each simulation required 650.5 cpuhours per ns of simulation time. Running on 32 processors gave approximately 20.3 hours per ns of simulation.
time. The simulations would have scaled to more than 32 processors, but limited number of processors and a queueing system in use on the cluster meant that 32 processors resulted in better use of the available resources.

### 4.3. Simulation Quality

As discussed in section 2.4.2.4 it is important to examine the quality of a molecular dynamics simulation. If the system has not equilibrated sufficiently then any data drawn from the simulation may contain artifacts caused by the relaxing protein. A simulation can be judged to have sufficiently equilibrated when physical and statistical properties such as root mean squared deviation (RMSD) have reached an acceptable level of stability.

RMSD is defined as the root mean squared deviation from a reference structure. This reference structure is arbitrary, but is most commonly selected as the structure at the beginning of the simulation. In this case, the RMSD of the system shows how far, on average the system has deviated from the initial structure. Figure 4.1 shows the RMSD of backbone Cα across all systems studied.
As can be seen in Figure 4.1 the RMSD has stabilized after approximately 6 ns in all cases except for $p$-hydroxyphenylethylamine, which has stabilized after approximately 8 ns. Some simulation drift occurs, but it is no more than approximately 0.1 Å over the course of the equilibrated region. As equilibration has been deemed to have occurred by 8 ns in each simulation, the last 5 ns will be used for analysis.

### 4.4. Analysis of Molecular Dynamics

To analyse the trends present in the simulations multiple statistics were extracted from the last 5 ns of the simulations. Due to there being two active sites in the full heterotetramer, two sets of values are found for each statistic extracted. These are both examined, but there should be little difference in which dimer is analysed.

As mentioned previously a motion exists in the linking region between the aromatic ring system and the quinone system of TTQ. Obviously then, investigating statistics
that are most likely to be affected by changes to this motion is of the upmost importance. Should the motion change, then it is likely that the dihedral angle measured across the donor carbon will be affected. As this is likely to fluctuate throughout the simulation, binning the data, and then plotting gives a better chance of analyzing the data than any other plotting method. The dihedral measured is indicated in Figure 4.2

Figure 4.2: Diagram indicating bond of rotation for dihedral measured in figure 4.3
Figure 4.3: CH2-NT-C12-C1 Dihedral A) From AH Dimer, B) From BD Dimer
Figure 4.4: D-H-A Angle A) from AH Dimer B) From BD Dimer
As can be seen from Figure 4.3 and Figure 4.4 there is a notable shift of the average angle with increasing \(p\)-substituent size, indicating that increasing the size of the substituent changes the angle of the donor carbon in relation to the quinone moiety. Figure 4.4 shows that the distribution of angles is notably less with both the methyl and hydroxyl substituents, as indicated by the narrowing of the peak.

These trends are noticeable in both the AH and BD dimers, but are more pronounced with the AH dimer. Investigating both dimers will provide further insight into differences between the active sites.

Since the focus of this work is driven by the availability of kinetic data for the range of phenylethylamine analogues, it is appropriate to make comparisons of these data with statistics than can be extracted from the simulations. Average lengths and
angles, as well as standard deviations (SD) of the data shown in figs 4.3, 4.4, and 4.5 can be plotted against kinetic values obtained from (Hothi et al. 2008). The average values correspond to the location of the peaks, and SD values are proportional to the width of the peaks.

Figure 4.6: SD Distance against $\Delta \Delta H^\ddagger$ A: AH Dimer, B: BD Dimer, $\Delta \Delta H^\ddagger$ taken from experiment
Figure 4.6 shows that there is a correlation between the SD of the hydrogen acceptor distance and the temperature dependence of the KIE, with an R-squared value of 0.3 for the AH dimer, and 0.99 for the BD dimer. This correlation is not visible against the temperature dependence of the rate (Figure 4.7).

These trends are not visible when comparing the average of the distance. In molecular terms, the average distance can be related to the position of the transferring hydrogen in relation to the donor and acceptor. The SD can be related to
the motion of the hydrogen (and therefore donor-hydrogen bond) throughout the simulation. A larger SD indicates a larger range of motion throughout the simulation.

Since molecular motions are of interest, using spectral densities to analyze these motions is appropriate, especially since such an analysis has been performed on AADH with bound tryptamine (Johannissen et al. 2007).

4.5. **Spectral Density analysis.**

Examining the spectral densities across the reactivity series previously outlined will give additional insight into what effect changing the \( p \)-substituent has on the inherent sub-picosecond motion that has been identified as being part of the hydrogen transfer step. In the AADH-Tryptamine system this motion has been identified at approximately 165 cm\(^{-1}\). This was identified on the donor carbon, so this will be investigated with the phenylethylamine analogues as probes. Spectral densities were calculated using velocities extracted every 10 frames from the trajectory.

The 165 cm\(^{-1}\) motion can be detected using the molecular dynamics techniques represented in this chapter. This can be shown using the Nyquist-Shannon theorem, which dictates that if a frequency \( B \) Hz is to be sampled, then measurements at least \( 1/(2B) \) seconds apart. When applied to MD simulations, one can work out the maximum frequency possible sampled by a simulation (the Nyquist frequency) from the sample rate, in this case, samples are taken every 10 fs. Using equation 4.1 we can calculate the maximum frequency of the motion that can be sampled.

\[
B < \frac{f_s}{2} \quad (4.1)
\]
Since the sampling frequency is $1 \times 10^{14}$ Hz then the maximum frequency that can be sampled is $5 \times 10^{13}$ Hz, which equates to 1667 cm$^{-1}$. So a motion at 165 cm$^{-1}$ is well below the maximum frequency that can be sampled.

![Graph](image)

**Figure 4.8: Donor carbon spectral density**

It has been reported (Johannissen et al, in publication) that to deconvolve the promoting vibrations from the otherwise noisy spectrum a promoting vibration can be treated as a “quasi-harmonic mode that describes the overall compression”. This enables the effective frequency of this mode to be calculated by mass-weighting the spectral densities according to their relative contribution to the spectrum, i.e. that the lower wavenumbers contribute more to motions than higher ones. The effective frequencies thus prioritize the lower more important frequencies over higher less important ones. The spectral density as expressed using effective frequencies for the BD dimer can be seen in Figure 4.9
Figure 4.9: Spectral Density of donor carbon expressed using effective frequencies

Plotting the effective frequencies for the phenylethylamine analogues against the temperature dependence of the KIE shows

Figure 4.10: Effective frequency against temperature dependence of the KIE for AH Dimer
Figure 4.11: Effective frequency against temperature dependence of the KIE for BD dimer. 

ΔΔH‡ from experiment.

Figure 4.10 Figure 4.11 show a very good correlation between increasing temperature dependence and increasing effective frequency. These data are obtained by averaging across the frequencies represented in figure 4.8. The R-squared values are 0.81 for the AH dimer and 0.94 for the BD dimer. This correlation reiterates that a promoting vibration must be driving the tunneling step.

4.6. Conclusion

A link between motions of the donor carbon and the temperature dependence of the KIE has been established. This has been examined by comparing the kinetic parameters of the physical reaction with statistical properties determined from molecular dynamics simulations. A correlation between the SD of the hydrogen acceptor distance and the temperature dependence of the KIE has been found which shows that as the motion of the hydrogen (and thus donor carbon) increases in amplitude the temperature dependence of the rate decreases. A correlation has also been found between the effective frequencies of spectral densities and the
temperature dependence of the rate. This shows that as the effective frequency increases, the temperature dependence also increases.

The two trends are connected, as amplitude of the donor carbon motion increases, the less energy contained within the motion, and therefore the less affected by temperature it is. The same is true of the effective frequencies, as the frequencies decrease, the motion becomes more flexible, and contains less energy, hence there is less temperature dependence of the KIE.
5. Conclusion and future work

The range of phenylethylamine analogues are a useful probe for environmentally coupled tunneling in enzyme reactions. The quantum mechanical calculations presented in section 3 show correlation between bond dissociation energy and both the observed rate constant and the activation enthalpy. The trends predicted by Marcus theory have been confirmed for low level PM3 calculations. Higher level density functional theory calculations, as well as calculations with an extended system show no trends. This is not understood, but investigating and correcting this is beyond the scope of this thesis.

Trends have also been found between bond length and the activation enthalpy, these are not connected to the presence of tunneling in the reaction mechanism, but do confirm that the low level PM3 calculations are correctly predicting that the reducing bond energy (as shown by reduced BDE) results in longer bonds.

Molecular dynamics studies presented in section 4 proved more useful. A trend has been identified showing that increasing amplitude of enzyme motion reduces the temperature dependence of the KIE. The presence of a temperature dependent KIE is known evidence of vibrationally coupled tunneling. A trend has also been found in between the effective frequency of the motions represented in the spectral densities of the molecular dynamics simulations and the temperature dependence of the rate. This adds emphasis to the vibrationally coupled tunneling model invoked for AADH.

Future work investigating vibrationally coupled tunneling making use of phenylethylamine analogues as probes is possible. Further simulations of a similar nature to those already performed in section 4 could be performed, but starting from different starting configurations. This would enable the trends found in section 4 to be verified, by sampling more configurations. Calculations of potential energy surfaces of the hydrogen transfer step could lead to computational calculation of the
KIE. The trend of these against the experimentally derived KIE would enable construction of new probes to further elucidate the role of enzyme dynamics in the hydrogen transfer step. In particular, what ways the active site environment supports the inherent motion within the iminoquinone moiety. Additional experimental work would be desirable to confirm the presence of the 165 cm\(^{-1}\) vibration. This work would require pump-probe spectroscopes that operate in the THz region, and as these are only just beginning to come online, so this work is still some way off.

In addition to these new areas of examination, the link between the quantum mechanical calculations and molecular dynamics simulations could be examined. Such molecular dynamics are increasingly studied to probe the link between the dynamic processes in proteins and the reactions they facilitate. The reactions of AADH and the posited role of dynamics in the reaction make this system an ideal application. Unfortunately full quantum mechanical molecular dynamics is too computationally intensive for studying a large system such as AADH, so a hybrid quantum mechanical/molecular mechanical (QM/MM) formalism must be employed to reduce computational complexity. This would provide the most in depth atomistic view of the dynamic processes and their links to the hydrogen transfer step possible with current technology.
6. Bibliography


