Chapter 1. Introduction

1.1 Enzymes

Enzymes are among the most active biological molecules, participating as catalysts in nearly every chemical reaction that takes place within a cell, controlling and accelerating most biological pathways. Enzyme-catalysed reactions are among the fastest chemical processes, reflecting the tremendous potential that these natural catalysts have developed over millions of years of evolution. Without their involvement, life as we know it would be impossible; some of these reactions, when not catalysed by enzymes, are amongst the slowest measured, with half times approaching the age of the Earth [1]. The sophisticated arrangements of atoms within their active sites makes it possible for these reactions to proceed at rates that are necessary for life. Many characterized enzymes have evolved to work nearly as efficiently as is physically possible, with rate enhancements up to 10^{17} fold with respect to the uncatalysed reactions in water and second-order rate constants approaching the rate of encounter with the substrate in solution ($\sim 10^{9}$ s^{-1} M^{-1}).

Although the importance of enzymes is widely known, their exact mechanisms of action are still a source of debate [2]. If unlocked, the mysteries behind the structure and mode of action of these powerful catalysts will not only contribute to the general understanding of biological processes, but will provide the basic knowledge necessary for the development of artificial catalysts, and the rational design of more potent and effective therapeutic drugs.

Recent advances in X-ray crystallography and NMR spectroscopy [3-5] have facilitated the determination of the three dimensional structure of enzymes and their dynamic behaviour in solution. These improvements in structural biology, combined with traditional kinetic and mechanistic experimental analysis, have contributed to the understanding of enzymes and their reactions. Unfortunately, while experimental methods can clearly measure the catalytic efficiency of a given enzyme and help in the identification of key residues, they cannot provide detailed molecular information of the reaction taking place within the active site. One of the main problems in elucidating the origin of the catalytic power of

enzymes is the difficulty of dissecting different energy contributions by direct experiments and, thus, establishing what the most important factors during catalysis are.

Computer simulations, however, make it possible to study reactions at the molecular level, looking at individual residues and atoms and the way they contribute to the activation of the substrates and the reaction mechanism [6]. Although computationally demanding, the type of calculations required to study large and complex biological systems, such as enzymes, are now possible thanks to the ever-growing computer power [7] and to substantial advances in programs and algorithms [7-10]. The new and fast expanding field of computational biology, or more specifically computational enzymology [6], has already achieved important goals in the simulation of enzyme reactions; however, it is still unclear what the most important factors in enzyme catalysis are.

1.2 Enzymatic Catalysis

The basic model used to describe enzyme catalysis is the Michaelis-Menten formalism [11]. The reaction is divided into two main steps, first the binding of the substrate (S) to the enzyme (E), forming the reactive complex (ES), followed by the transformation of the substrate into product (P),

$$E + S \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} ES \xrightarrow{k_{cat}} EP \longrightarrow E + P$$

where K_d (k_{-1}/k_1) represents the affinity of the enzyme (E) for the substrate (E), and $k_{cat,}$ is the rate constant for the catalysed transformation of substrate (S) into product (P). The parameter K_M, defined as $(k_{-1}+k_{cat})/k_1$, is the dissociation constant of the ES complex, and is commonly used for the description of enzymes and their catalytic properties.

To facilitate the interpretation of enzymatic reactions it has been assumed that there is a transient species, the transition state (TS), which is in thermodynamic equilibrium with reactants and decays to products,

$$E + S \xrightarrow{\Delta G_{bind}} ES \xrightarrow{\Delta G^{\neq}} ES^{\neq} \longrightarrow E + P$$

where $S^{\neq}$ represents the TS and ΔGs are the free energy changes associated with each step of the reaction, the binding (ΔG_{bind}) and the formation of the TS ($\Delta G^{\neq}$). However, despite its practical utility and acceptance, the idea of a thermodynamically stable transition state is physically impossible. A recent analysis by Schramm [12], supports the idea that binding equilibrium does not exist at the transition state, it is the synchronized dynamic motions within the enzyme that lead to an instantaneous arrangement for the formation and stabilization of the transition state that only lasts for the duration of bond breaking. Although the transition state is clearly not a real thermodynamic entity, it is many times treated as such in order to facilitate the mathematical description of enzymatic reactions.

1.2.1 Substrate binding

The first step in every single enzyme-catalysed reaction is the binding of the reactant/s within the active site. It is now accepted that the old idea of the 'key and lock' interaction of an enzyme and its substrate is not an accurate description of most catalytic complexes. The ligand-protein interactions resemble more a 'hand and glove' association, where both parts are flexible and adjust to complement each other. They can modify their shape and mould their complementarity so as to increase favourable contacts and reduce adverse interactions, maximising the total binding free energy [13].

Proteins in solution are mobile molecules. They do not exist in a single conformation, but in a manifold of different conformational states separated by low- and high-energy barriers. The distribution and stability of each conformational state will depend on the physicochemical properties of the environment and the protein itself (e.g. free or ligand-bound) [14]. Moreover, not all these conformations will be equally able to bind a given ligand productively, i.e. energetically poised to lead to reaction. Some will be more likely to accommodate the reactant molecule within the binding site without having to undergo large changes, while others will be less likely, or even incapable, of accommodating the ligand due, for example, to loop conformations that block the access to the binding site [15] (Figure 1-1). It has been found that active-site regions present areas of both low and high conformational stability [16]. Mobile loops that close over the ligand upon binding are included within

the flexible parts, while catalytic residues, for example, are usually structurally stable. This dual character of the active-site environment appears important for optimum binding.

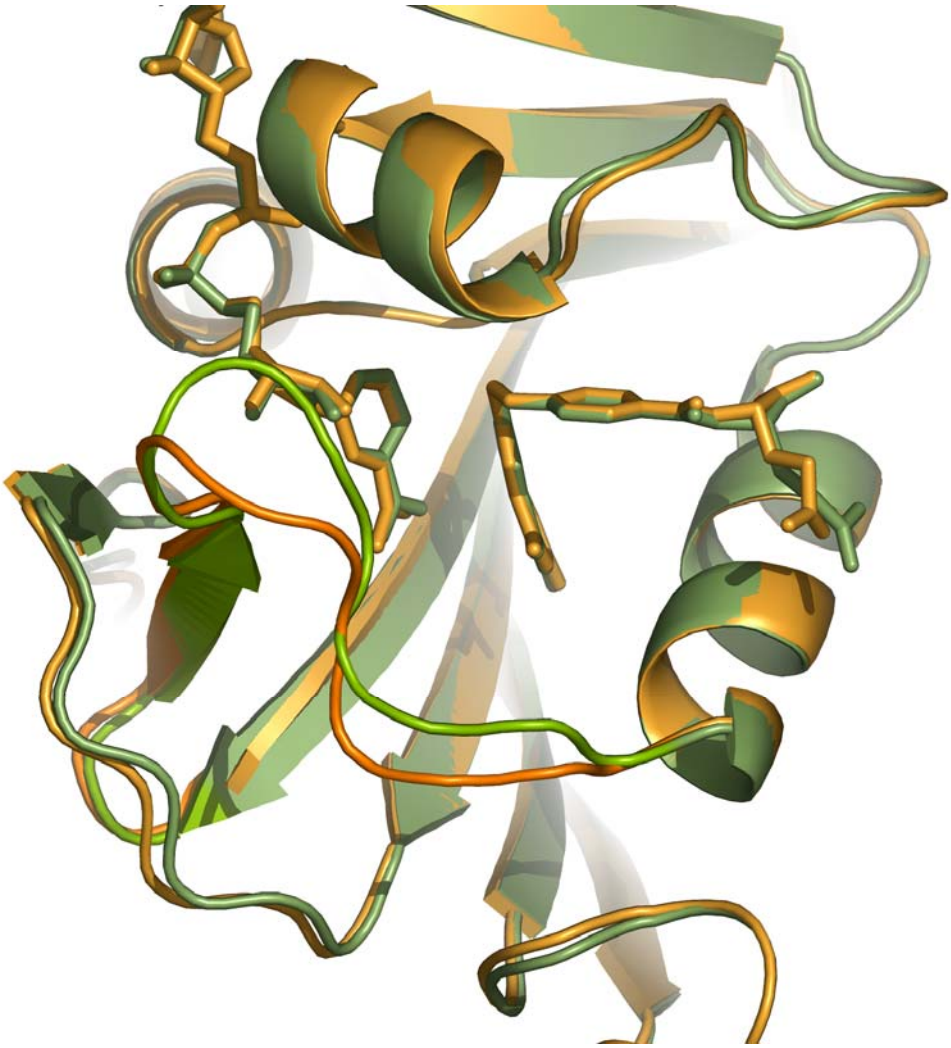

Figure 1-1 Protein flexibility.

The mobility of proteins is clearly shown in these two superimposed ternary complexes of the enzyme dihydrofolate reductase with folate and NADP+ (1RA2 and 1RX2). The flexible loop has been shown to adopt different conformations during the catalytic process; here it is seen "frozen" in the open (orange) and closed (green) states in two different crystal forms of the same complex.

The presence of the ligand itself is expected to affect the structure of the binding site and the dynamic equilibrium between different conformational states of the protein [17]. During a binding event, the protein conformer most likely to accommodate the ligand will be depleted from solution to form a ligand-bound

complex, and other conformers will then adjust to fill the vacated conformational space, driving the binding process forward [14].

Before binding, both ligand and protein are interacting with the solvent, and desolvation of polar and charged parts of the molecule upon complex formation carries a high desolvation penalty. This is partially compensated by the electrostatic interactions and hydrogen bonds that are formed in the complex, together with the increased entropy resulting from the release of water molecules. However, complex formation is also accompanied by translational and rotational losses as the two molecules become one complex. Enthalpy/entropy compensation is commonly observed for many types of host-guest complexes. Increasingly negative entropies of complexation usually go together with increasingly negative complexation enthalpies. Tighter binding leads to both more negative ΔH (favourable) and more negative ΔS values (unfavourable). An overall positive entropy of association may occur if solvent molecules originally inside the active-site cavity are released upon association.

Different factors, including hydrogen bonding, hydrophobic effects, van der Waals forces, and electrostatic interactions, contribute to the binding free energy [18]. Two factors enhance ligand-receptor interactions compared with ligand-water interactions: the hydrophobic effect and the high atom density in macromolecules. Protein densities are 30-40% higher than the density of water, leading to a 50-70% higher density of non hydrogen atoms [19] and increasing the number of van der Waals interactions.

The stability of a particular complex can be measured in terms of the equilibrium constant K_d (k_{-1}/k_1) (see section 1.2). However, the experimentally measured constants usually depend critically on the temperature, ionic strength, pH, pressure, etc. Another constant commonly used to characterize the enzyme-substrate complex strength is K_M, defined as $(k_{-1}+k_{cat})/k_1$. It has been proposed that enzymes have evolved to have K_M values in the mM-μM range, with average values of 10^{-4} M [20], to match the low concentration of substrates present under physiological conditions.

The affinity of an enzyme for a given ligand can also be estimated using computational tools. The most rigorous way currently available for the determination of binding free energy is free energy perturbation (FEP) or thermodynamic integration (TI) methods. In most cases the free energy difference between two similar states is determined by slowly mutating one state into the other, chemical alchemy [21] (see Chapter 2 for further details).

1.2.2 Enzyme Kinetics

Following the binding of the ligand to the active site, the reaction can take place. The rate constant for a reaction, whether in water or within a protein environment, can be defined according to the transition state theory [22,23] as:

$$k(T) = \gamma(T)\left(\frac{k_B T}{h}\right)\left(C^0\right)^{1-n} e^{\left[\frac{-\Delta G^{\neq}(T)}{RT}\right]}$$

Where $k(T)$ is the temperature-dependent rate constant, C^0 the standard state concentration, n the order of the reaction, R the gas constant, T the temperature, and $\Delta G^{\neq}$ the activation free energy. The factor ($k_B T/h$) is a frequency factor for crossing the transition state, and $\gamma(T)$ is the generalized transmission coefficient, which can be expressed as [23]:

$$\gamma(T) = \Gamma(T)\kappa(T)g(T)$$

The first factor, $\Gamma(T)$, represents the dynamical recrossing and is usually equal to or less than 1, as some trajectories crossing in the direction to products may return to the reactant region [24] (Figure 1-2). The second, $\kappa(T)$, arises from the contributions of quantum mechanical tunnelling, and is usually 1 or greater [23]. The third and last factor, $g(T)$, arises from deviations of the equilibrium distribution in phase space [25,26].

Perturbations of the free activation energy, $\Delta G^{\neq}$, or any of the factors in the generalized transmission coefficient, $\gamma(T)$, are expected to affect the rate of a reaction. The catalytic power of enzymes has been associated with changes to both $\Delta G^{\neq}$ and $\gamma(T)$; however, the exponential dependency of $\Delta G^{\neq}$ changes result in larger contributions to the rate of the reaction than changes of the transmission coefficient [7].

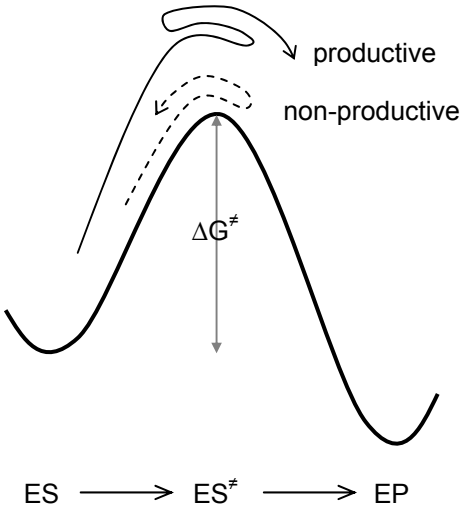

Figure 1-2 Dynamical recrossing of the transition state.

During an enzymatic reaction, the enzyme-substrate (ES) complex has to reach the transition state ($ES^{\neq}$) prior to formation of products (EP). This is a dynamic process, and not all trajectories passing though the transition state will lead to products (non-productive). Therefore, acceleration of a reaction can be produced not only by reduction of the activation energy ($\Delta G^{\neq}$), but also by increasing the proportion of productive trajectories.

1.2.3 Catalytic efficiency

The efficiency of an enzyme is usually defined as:

$$(k_{cat}/K_M)\Big/k_{wat}$$

where k_{cat} is the rate constant for the enzymatic reaction; K_M is the enzyme-substrate dissociation constant ($(k_{-1}+k_{cat})/k_1$); and k_{wat} represents the rate constant of the uncatalysed reaction in water a pH 7 (Figure 1-3).

The catalysed and uncatalysed reactions are usually compared according to the following thermodynamic cycle,

$$\begin{array}{ccc} E+S & \xrightarrow{\Delta G^{\neq}_{non}} & E+S^{\neq} \\ \Delta G{=}0\downarrow & & \downarrow\Delta G_{rel} \\ E+S \xrightarrow[\Delta G_{bind}]{} ES & \xrightarrow[\Delta G^{\neq}]{} & ES^{\neq} \end{array}$$

When the reaction is faster within the enzyme ($\Delta G_{bind}+\Delta G^{\neq} < \Delta G^{\neq}_{non}$), ΔG_{rel} is negative, and the transition state $S^{\neq}$ is more stable when it is bound to the

enzyme than when free in solution. This stabilization of the transition state within the protein environment is claimed to be the key mechanism by which enzymes accelerate reactions. However, this is not sufficient to explain how this differential binding is accomplished.

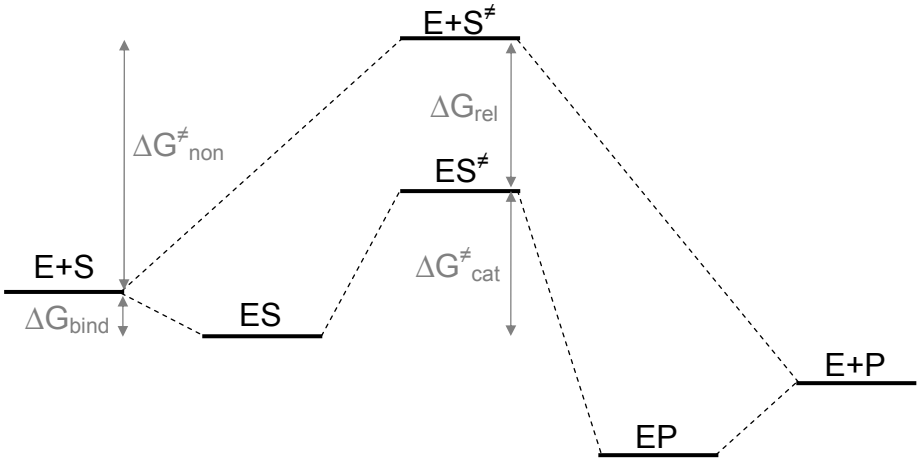

Figure 1-3 Schematic representation of a catalysed reaction ($\Delta G^{\neq}_{cat}$) and its uncatalysed counterpart ($\Delta G^{\neq}_{non}$).

The active-site environment (E) is expected to provide a larger stabilization to the transition state ($ES^{\neq}$) than to the substrate (ES). This preferential stabilization of the transition state ($|\Delta G_{rel}| > |\Delta G_{bind}|$), proportional to the difference in velocity between the catalysed and the uncatalysed reaction, has been claimed to be the key mechanism by which enzymes accelerate the formation of products (P).

It is important to define what contributions are responsible for the differences in activation energy between the water and protein environments, and the way in which they operate (i.e. stabilization of the TS or destabilization of the reactants).

1.3 Origin of the Catalytic Power of Enzymes

It is very well known that many enzymes have evolved by optimising k_{cat}/K_M [1], and that the key question relates to the reduction of the activation barrier, the energy required to transform reactants into products ($\Delta G^{\neq}$). However, it is still unclear to what extent different factors, such as electrostatic, steric, hydrogen-bonding, or differential solvation effects, contribute to the catalytic power [12,27,28]. While several more or less reasonable microscopic mechanisms have been proposed to explain the origin of the catalytic power of enzymes [7,8,29],

none seems best to explain every single reaction, but rather the best description varies with the particular system. The following factors are those that appear to be most important for enzyme catalysis:

1.3.1 Desolvation Hypothesis

For a reaction in solution to occur, the bulk solvent must reorganize around the reactive species to produce a configuration that will allow the transition state to form [30,31] (Figure 1-4).

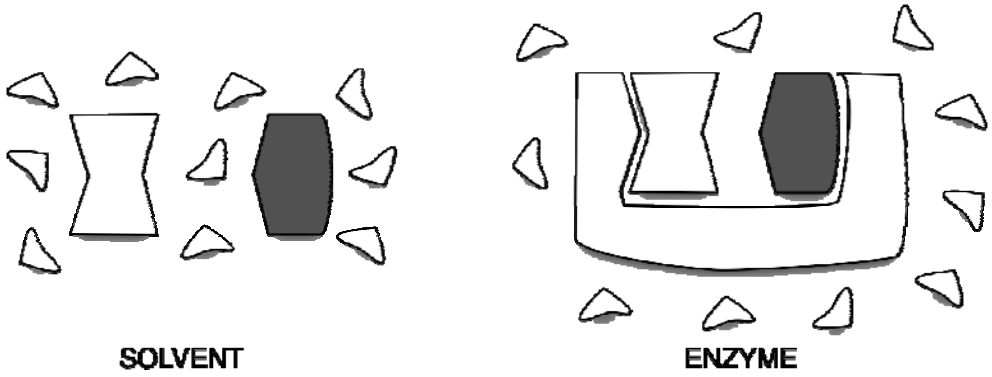

Figure 1-4 Desolvation hypothesis.

The desolvation hypothesis suggests that the reacting species lose their solvation shell upon enzyme binding, accelerating the reaction process. No energy needs to be spent in reorienting solvent molecules, and the low dielectric nature of the active site further assists in stabilizing the transition-state complex.

Cannon and Benkovic [29] suggested that the large rate enhancements observed for enzyme-catalysed reactions result from the large solvent reorganization energies required for the reference reaction in water. Reactions in solution appear retarded by the solvent when compared with a similar reaction in the gas phase [32]. Extrapolation to the low dielectric environment typical of an active site led Dewar and Storch [33] to propose that enzymes act through a desolvation mechanism. Enzymes active sites are usually buried within the protein structure, sheltered from the polar water solvent. As the substrates bind to the enzyme, they are desolvated and water molecules are forced out of the active site, resulting in a change of the solvent environment that has been suggested to accelerate the reaction [34,35]. The enzyme orotidine monophosphate decarboxylase has been said to use a general acid catalysis

mechanism only effective in the low dielectric of its active site [36], and the enzymatic action of haloalkane dehalogenase has been attributed to a combination of desolvation effects and transition-state stabilization [37]. As some studies have shown [38,39], the organized dipolar nature of the active site seems to "solvate" the transition state many times much more strongly than bulk solvent, without the reorganization costs associated with reactions in solution.

1.3.2 Electrostatic Effects

The effect of the environment on a given reaction rate depends on both the dielectric constant and the polarity of the medium. Electrostatic effects cannot be large in aqueous solution because of the large dielectric constant even at short distances [40]. Although enzyme active sites generally have a low dielectric constant when compared with bulk solvent, they are usually very polar heterogeneous environments that can create intense electric fields. This can significantly affect the properties of the substrate, such as its pK_a. Moreover, during an aqueous reaction the solvent requires significant reorganization energy to orient the polar environment towards the transition-state charges, while the enzymatic environment is already preorganized [40,41] (Figure 1-5).

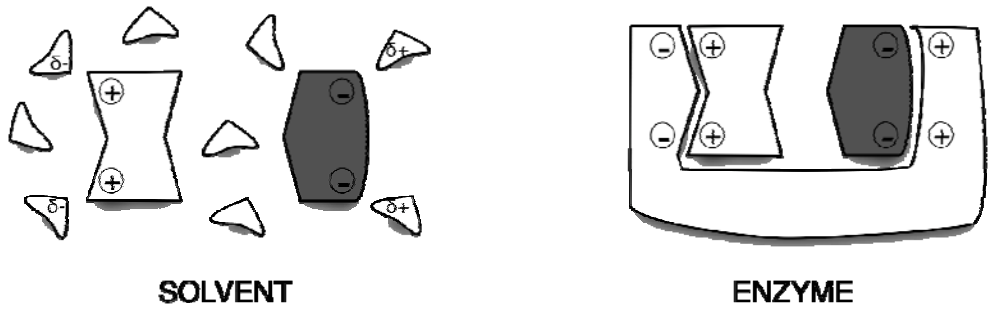

Figure 1-5 Electrostatic effects.

It has been claimed that the active-site region provides a pre-organized environment, with dipoles oriented to complement the transition-state configuration of the reactants. Not only there is no reorganization penalty, as in the solvent environment, but also the low dielectric constant of the protein environment can result in more intense electric fields that significantly affect the properties of the reacting species.

Dipoles within the active site are oriented to complement the transition-state configuration of the reactants and, as a result, there is no reorganization penalty. Much of the reorganization energy is paid during the folding process, and the

catalytic power of enzymes is said to be stored in their folding energy [10,41-44]. However, in a recent study of the electrostatic complementarity of the enzyme ketosteroid isomerase and transition-state analogues [45] it was concluded that the electrostatic complementarity makes at most a modest contribution to catalysis of ~300 fold. The catalytic efficiency was said to result, instead, from the combined modest contributions of several factors, rather than a single dominant contribution.

1.3.3 Entropic Effects

Experimental [46] and theoretical studies [47] suggest that bimolecular reactions in solution are slow, in large part due to the entropic penalty associated with bringing the two reactants together into a reactive complex. The enzyme environment drastically reduces the large configurational space available for the reacting fragments in solution (Figure 1-6).

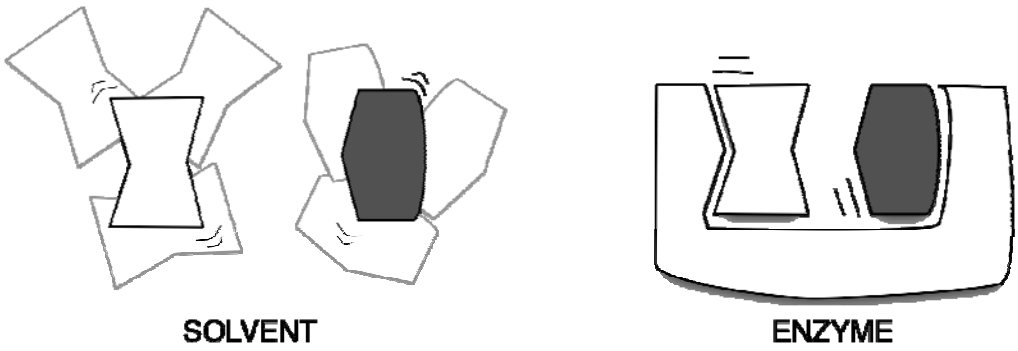

Figure 1-6 Entropic effects.

The entropic theory supports the idea that the loss of movement upon binding facilitates the encounter and reaction of the ligands. The smaller conformational space available within the active site facilitates the formation of a reactive complex.

Although the alignment of the reacting groups in a specific orientation is entropically unfavourable, as it restricts the conformations available to the substrate, the reaction will be entropically more likely because the substrate will have to explore fewer conformations to reach the transition state. The entropy loss on binding is compensated for by the enthalpy of binding between the substrate and the enzyme and the favourable enthalpy and entropy contributions from the liberation of water molecules to the bulk solvent. This entropy effect has been said to be responsible of the rate enhancement observed in several

enzymes, like trypsin [48] and catechol O-methyltransferase [49]. However, the entropic effect has been found to be minimal, or non-existent, in reactions involving single substrates, and hydrolytic reactions where the second reactant is present in abundance [1].

1.3.4 Strain Mechanism

The spatial disposition of residues within the active site provides a confined space for the species to react. This steric strain imposed on the ground state of the reactants could facilitate the reaction by destabilizing the ground state, pushing the ligands towards the transition state (Figure 1-7).

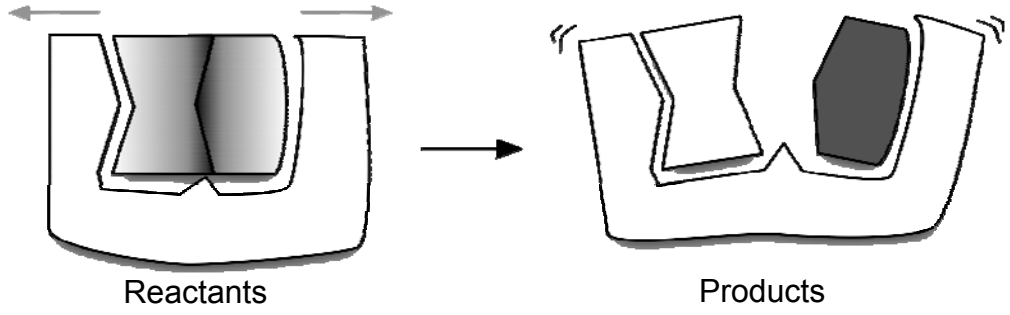

Figure 1-7 Strain mechanism.

The strain mechanism suggests that the restricted space of the active site facilitates reactions by imposing a physical strain on the reacting species, destabilizing the ground state and pushing the system towards the transition state.

This idea was put forward in classical studies of lysozyme [50], and invoked in histidinol phosphate aminotransferase [51], and class A beta-lactamases [52]. In the case of chorismate mutase studies both supporting the strain mechanism [53] and rejecting it [54] have been published. Ferrochelatase, which catalyses the insertion of metal ions into porphyrin, as well as the antibody metal chelatase [55], have been said to function by straining the porphyrin ring towards a distorted conformation in which the two pyrrole N-σ electrons are more accessible to metal ions [56]. Moreover, the chelatase-induced distortion of the porphyrin substrate not only decreases the activation energy but also modulates which divalent metal ion is incorporated [57].

1.3.5 Reactive Near Attack Conformers (NACs)

In order for a reaction to take place, it has been suggested that the reactants should adopt a ground-state conformation that closely resembles the transition state (near attack conformers, NACs) [58,59]. It is thought that the preorganized enzyme environment favours the formation of NACs compared with those in solution, and that this is a critical component of enzyme catalysis [58] (Figure 1-8).

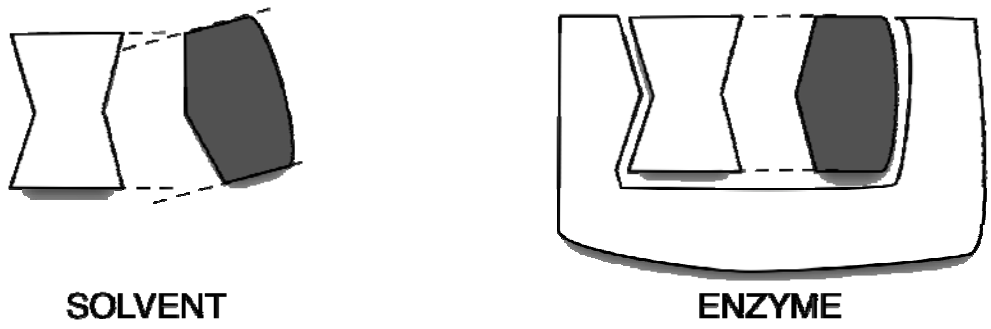

Figure 1-8 Reactive Near Attack Conformers.

The active-site environment facilitates the formation of near-attack conformers compared with bulk solvent. The reduced space within the active site, and its preorganized nature, facilitates the approach of the ligands in the right conformation for the reaction to occur.

Dynamic excursions of enzyme-bound reactants, together with motions of side chains, loops and domains, change the distances and dynamic interactions between the ligand and the enzyme, resulting in an alignment poised to achieve the transition state [12]. The source of the enormous rate acceleration by the enzyme catechol O-methyltransferase appears to be related to the ability of the enzyme to desolvate the substrate and orient the two reactants into a NAC [60]. Molecular dynamics simulations have suggested that the population of NACs within the active site can be significantly larger than that observed in bulk water; 50% in the case of lactate dehydrogenase [61].

1.3.6 Orbital Steering Mechanism

Koshland [62] introduced the idea of orbital steering which suggests that a precise orientation of a substrate to a reactive group was the key to enzyme activity. The active-site geometry restricts the approach of the reacting species to a very narrow angular range, with a steep potential energy surface in the

direction perpendicular to the reaction coordinate. This constraint ensures that the reacting species approach each other in the most favourable orientation. Mechanistic studies of the enzyme aminoimidazole carboxamide ribonucleotide transformylase/ inosine monophosphate cyclohydrolase (ATIC) suggest that it catalyses the ring-closure reaction by an orbital steering mechanism [63]. This mechanism has also been invoked to explain how the hammerhead ribozyme both prevents and enhances RNA autocatalysis [64] and how the flavoprotein D-amino acid oxidase catalyses the dehydrogenation of amino acids to the corresponding imino acids [65]. Experimental studies on isocitrate dehydrogenase indicate that orbital overlap produced by optimal orientation of the reacting orbitals plays a major quantitative role in the catalytic process [66].

1.3.7 Low-Barrier Hydrogen Bonds

Common hydrogen bonds, involving NH or OH groups as donors and N or O atoms as acceptors, are usually about 2.8 Å in length and have energies of approximately 5 kcal/mol. Low-barrier hydrogen bonds (LBHB) are usually shorter and stronger than normal H bonds, with distances of around 2.5 Å and energies of up to 20 kcal/mol. These bonds are partially covalent [67] and show a particular NMR chemical shift. Although their existence has been known for many years, the idea that they could play a role in enzyme catalysis was only put forth in 1993 [68-70]. Although this type of hydrogen bonds is not usually found in water, the non-aqueous active site provides an environment closer to organic solvents, where LBHB are possible [71]. It has been suggested that the enzyme can transform a weak H bond into a strong one by modifying the pK_a value of the substrate so that it is closer to that of the residue to which it is hydrogen bonded. As the distance between donor and acceptor shortens, the barrier to proton transfer decreases, and the hydrogen becomes diffusely distributed between the two centres. Examples of enzymes suggested to exhibit LBHB during catalysis include chymotrypsin [69], ketosteroid isomerase [68], triose-phosphate isomerase [68], citrate synthase [68], mandelate racemase [70], and serine proteases [72]. Cleland *et al.* [71] suggested that most enzymatic reactions that involve proton transfer from a general acid or to a general base are likely to involve LBHBs. Furthermore, LBHBs are claimed to accelerate a reaction by 5 orders of magnitude, at least, with other factors providing the rest of the

catalysis. However, Warshel and coworkers have postulated that LBHB are more likely to prevent rather than favour the reaction, as their diffuse charge distribution will result in smaller electrostatic stabilization of the transition state (TS) structure [73,74].

1.3.8 Dynamic Effects

More than 60 years ago, Kramers suggested that enzymes can use dynamic fluctuations to bring the initial complexes into a reactive conformation, and that these movements lead to passage through the transition state [75] (Figure 1-9).

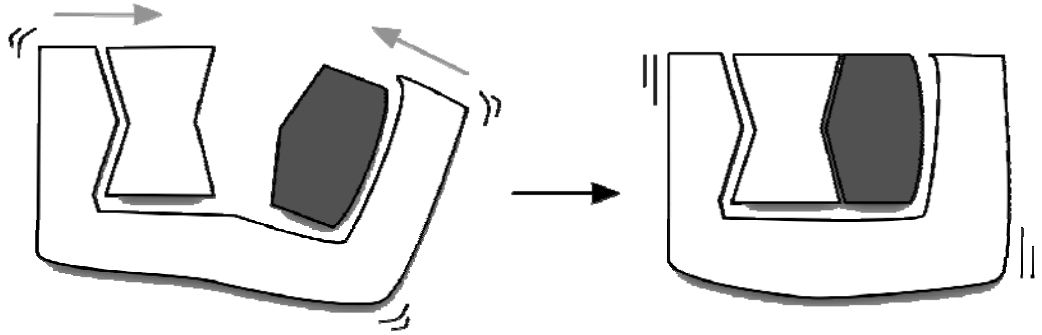

Figure 1-9 Dynamic effects.

Enzymes are not rigid objects, but flexible proteins that can undergo both large-scale and local changes. The idea that specific protein movements facilitate the reaction process has become stronger during the last few years, with several supporters behind it.

The increased interest in the role of protein dynamics in enzyme activity, combined with more accurate experimental techniques and growing computer power, has resulted in the publication of a considerable number of studies during the last few years [76-82]. In recent experimental work [79], it was shown that the characteristic motions of the enzyme prolyl *cis-trans* isomerase cyclophilin A (CypA) during catalysis were already present in the free enzyme, with frequencies corresponding to the catalytic turnover rate. This suggests that the protein motions necessary for catalysis are an intrinsic property of enzymes [77]. Warshel and coworkers [82], on the other hand, have concluded that there is no evidence to date for a dynamical contribution to enzyme catalysis.

As mentioned before (section 1.2.2), the overall rate of a reaction within a protein may be expressed as the product of an equilibrium transition-state rate,

which depends exponentially on the activation free energy, and a transmission coefficient, which accounts for dynamical recrossing of the barrier [83]. Motions that influence the activation barrier are equilibrium properties of the system, thermally averaged, while those influencing the transmission coefficient are the dynamic properties [7].

1.3.8.1.1 Effect on activation free energy

Theoretical studies indicate that equilibrium motions representing conformational changes along the reaction coordinate have an important role in enzymatic reactions [84-87]. Conformational changes of the enzyme between the reactant and the transition state can contribute to lower or increase the activation barrier. Enzymes have been shown to close loops, flaps or domains over the catalytic site before reaching the transition state. Superimposed on these movements, are higher frequency bond vibration modes from amino acid side chains and from the reactants, which can affect hydrogen bonds, angles and distances of interactions [12]. Thus, computational and experimental studies indicate that the protein architecture is designed to promote motions, which facilitate the formation of the transition state [88,89]. In the case of purine nucleoside (PNP), the concerted effects of five catalysis-promoting motions result in the formation of the transition state [90]. In the short-chain acyl-co-enzyme A (CoA) dehydrogenase, the effect of the relaxation of the protein along the reaction path was estimated to provide a stabilization of 18 kcal/mol [91]. In Co corrinoid mutases [92] x-ray crystal data suggests that both conformational changes and steric effects are important in lowering $\Delta G^{\neq}$. A simulation study of xylose isomerase [93,94] showed that motions of Mg^{2+} ions in the active site are correlated with the hydride-transfer reaction coordinate. Correlated motions also appear to participate in the formation of NACs in catechol O-methyltransferase [95].

Several dynamic regions have been associated with the reaction pathway of dihydrofolate reductase (DHFR) [80,96-99]. Structural studies have shown that the M20 loop can adopt several conformations [100] and that it oscillates at frequencies similar to k_{cat} [101]. These movements have been said to be coupled to the hydride-transferred step and provide stabilization to the transition state

[84,85]. Both experimental [102,103] and theoretical [104,105] studies on DHFR have shown that the mutation of key amino acids located far from the active site disrupts the network of coupled motions, decreasing the probability of sampling configurations conducive to hydride transfer and increasing the free-energy barrier.

1.3.8.1.2 Effect on the transmission coefficient

Protein dynamics that affect the transmission coefficient are not expected to contribute to the overall reaction rate as much as those that lower the free-energy barrier [88]. The magnitude of effects produced by changes in the recrossing factor, $\Gamma(T)$, and the nonequilibrium factor, $g(T)$, within the enzyme environment have been said to be unimportant, with only those associated with tunnelling leading to a sizable contribution to rate enhancements [7].

A recent study discovered a network of protein vibrations in cyclophilin A coupled to its catalytic activity [106,107]. Increasing the kinetic energy of selected protein vibrational modes, corresponding to low-frequency network fluctuations, was shown to produce a significant increase in the transmission coefficient, suggesting that not only the fast motions but also conformational fluctuations impact on the barrier recrossing behaviour of the system [76]. These observations lead to the idea that protein dynamics can produce changes in the barrier crossing that result in increases of reactions rates of several orders of magnitude [58]. Klinman and coworkers [108,109] have investigated the recrossing phenomenon as a function of temperature, by comparing thermophilic and psychrophilic enzymes, and mutagenesis of promoting groups, with convincing results.

Enzyme motions in reactions involving hydrogen transfer have been said to increase the probability of quantum mechanical tunnelling, given by the pre-exponential factor $\kappa(T)$ [110,111]. Because experiments cannot provide direct measurements of the magnitude of $\kappa(T)$, simulations including quantum effects are essential. Vibrations that occur at the sub-picosecond level, and are much faster than the chemical turnover, can increase the rate of the reaction by

decreasing the width and the height of the barrier along the tunnelling coordinate [111-113].

1.3.8.2 Quantum Effects

The classical conversion of reactants into products implies the passage over a potential barrier, which involves a certain activation free energy ($\Delta G^{\neq}$). If quantum effects are considered, however, this is not always the case. The probability that the system will tunnel through the barrier and allow the reaction to occur without any energy cost is not zero (Figure 1-10).

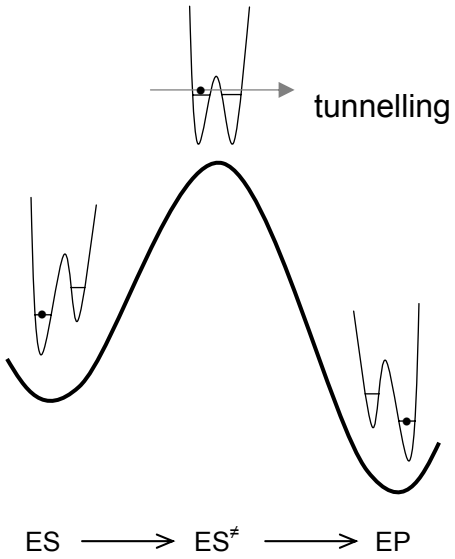

Figure 1-10 Quantum effects.

Schematic representation of the free-energy profile for a catalysed hydrogen-transfer reaction. The potential energy curve along the hydrogen coordinate and the ground vibrational level is shown for each species (ES, enzyme-substrate complex; $ES^{\neq}$, transition state; EP, enzyme-product complex). The double-well proton potential becomes symmetric at the transition state, making it possible for the hydrogen atom to tunnel through the barrier without any energy cost.

The probability of the tunnelling is proportional to the size of the particle; it is very important in reactions involving electron transfer, but it has also been reported for hydrogen atoms. The pre-exponential factor $\kappa(T)$ is the one that determines the tunnelling contribution to the reaction (see section 1.2.2). Accumulated experimental evidence suggests that hydrogen tunnelling may play an important role in several enzyme-catalysed reactions [111,114-120]. The protein environment, and its dynamic properties, could serve to align the

reactants and product energy wells in a configuration that is suitable for the hydrogen tunnelling process. In order to explain the large kinetic isotope effects and weak temperature dependence of the reaction catalysed by soybean lipooxygenase-1, a rate enhancement of 780 by quantum tunnelling contribution has been suggested (~3.9 kcal/mol) [121,122]. The computational study of horse liver alcohol dehydrogenase gave rise to the concept of 'protein promoting vibration'; low-frequency modes that are dynamically coupled to the hydrogen transfer coordinate [80,86,123,124]. Experimental evidence for the occurrence of hydrogen tunnelling in *E. coli* dihydrofolate reductase (DHFR) has also been reported [125], and associated with specific promoting motions [85,121]. Calculations at room temperature have yielded rate enhancements due to tunnelling of 1.5 for the intermolecular transfer in triosephosphate isomerase [126], and 23 to 114 for methylamine dehydrogenase [122,127,128]. In a recent study of the reaction pathway of tryptamine oxidation by aromatic amine dehydrogenase [129], X-ray structures and computational analysis were combined to show that proton transfer occurs in a reaction dominated by tunnelling over ~0.6 Å, with a kinetic isotope effect of ~55. It was suggested that a short-range motion that modulates the distance between proton and acceptor promotes tunnelling, but none of the long-range coupled motions are involved.

1.3.8.3Covalent Hypothesis

Spector presented the idea of a covalent component of enzymatic catalysis in his book "Covalent Catalysis by Enzymes" [130]. In a recent paper [131] the group of Houk performed a survey of several enzymes and compared binding constants among different complexes (e.g. cyclodextrins binding organic molecules, catalytic antibodies, and enzymes binding their substrates). The strong binding observed for the association of enzymes and transition states, with an average association constant of $10^{16\pm4}$ M^{-1}, led them to conclude that all enzymes that present an efficiency ($(k_{cat}/K_M)/ k_{water}$) greater than 10^{11} M^{-1} involve covalent bonds between the transition state and the enzyme. However, as stated by Bruice *et al*. [27], the conclusions of Houk *at al.* are not entirely valid, as they presumed that the high enzyme efficiency values were result of the strong binding, reflected in large enzymatic rate constant k_{cat}/K_M, without taking into account the

possibility of this being the result of the low rates for the reactions in water (k_{water}).

Overall, these and other catalytic strategies have been actively supported by some and adamantly rejected by others. Although some of the mechanisms appear to have stronger experimental and theoretical support than others, like electrostatic stabilization and dynamic effects, none of them seem to clearly explain on their own the total reduction of $\Delta G^{\neq}$ observed in every single enzyme-catalysed reaction [45]. It is clear that many of the questions regarding enzyme catalytic power cannot be uniquely addressed by current experimental approaches. This is an energy issue, which needs to be resolved by dissecting the individual contributions to the total observed energy. Current advances in algorithms and increasing computational power make it possible to approach the enzymatic catalysis problem from a detailed theoretical point of view.

1.4 Goals of the Thesis

The precise mechanism that enzymes use to catalyse their reactions is still a topic of debate - a long-standing question of critical importance. The understanding of these natural catalysts holds the key to the rational design of man-made catalysts, as well as the development of novel and more potent drugs. Although essential, experimental analysis cannot provide the molecular detail of the reaction taking place within the active site and, therefore, an alternative approach is necessary. Increasing computer power and improvements in computational algorithms during recent years have allowed development of the field of computational enzymology. Computational modelling and simulation techniques permit us to look at reactions at the atomic level, and provide a detailed description of the different energy contributions involved in the catalytic process.

The principal aim of my thesis studies is to contribute to the understanding of enzymes and their mechanisms through the analysis of two particular cases: DfrB or type II dihydrofolate reductase (DHFR) and methyltetrahydrofolate:

corrinoid/iron-sulfur protein methyltransferase (MeTr). For each system, the specific aims are summarized as follows:

DfrB DHFR:

- to study how the enzyme binds and interacts with its substrates, dihydrofolate (DHF) and NADPH,
- to generate a feasible structural model of the DfrB•DHF•NADPH reactive complex using docking programs,
- to analyse the stability of the reactive complex and describe key protein-ligand interactions using molecular dynamic simulations,
- to characterize the DfrB DHFR family of proteins from a sequence and a structural point of view,
- to study the evolution and possible origins of the DfrB DHFR family.

MeTr:

- to study how the enzyme activates its substrate N5-methyltetrahydrofolate (CH_3THF), by protonation,
- to compare and contrast the activation mechanism of MeTr with that of a cobalamin-dependent methionine synthase (MetH),
- to assess the accuracy of quantum mechanical methods in predicting protonation positions and pK_as of CH_3THF and similar molecules in aqueous solution,
- to study the effect of the protein environment in the protonation state and pK_a of CH_3THF,
- to analyse the protonation state of aspartic acid residues located within the active site.

Through the detailed study of these two specific systems I hope to contribute to the body of basic core knowledge, a foundation from which a greater understanding of enzymes will emerge.