3. Simulation of enzyme catalyzed reactions

The thermodynamic, kinetic and dynamic features of a chemical process are determined by its free energy surface, *i. e.* the energy of a system as a function of the nuclear coordinates. The knowledge of the complete free energy surface of a reaction in an enzyme environment would reveal the reaction path followed by the enzyme when converting substrates to products. The comparison of the free energy surfaces of the enzyme reaction and a corresponding reference reaction in aqueous solution could show, by which means the enzyme accelerates a chemical reaction. It is however not possible to establish a full free energy surface of an enzyme catalyzed reaction by experimental techniques. This task could, at least in principle, be fulfilled by computer simulation studies.

In this chapter I will present the basic concepts of enzyme catalysis and kinetics. Then the principles and main problems of computer simulations of enzymatic reactions are discussed. In the last section the widely used empirical valence bond theory of Arieh Warshel is described in detail. An application of this theory on the enzyme Acetylcholinesterase is presented in chapter 4.

3.1. Theory of enzyme catalysis

Catalysis is in general defined as a process where the rate of a chemical reaction is increased by the interaction of the reacting partners with another substance which is called the catalyst. The equilibrium constant of the reaction is not changed by such a process. The catalyzed reaction shows a free energy surface with an activation barrier on the reaction path that is lower than in the uncatalyzed reaction. The modification of the free energy surface results from the interaction with the catalyst.

The catalytic efficiency of enzymes is outstanding. The rates of enzyme catalyzed reactions can be enhanced by a factor of up to 10^{17} compared to the uncatalyzed reaction.[33] The proposition by Emil Fischer that an enzyme and its substrate can be seen in analogy to a lock and a key was the first attempt to account for the exceptional power and specificity of enzymes as biological catalysts.

The rate enhancement of an enzyme catalyzed reaction is today explained by the fact, that the transition state of the reaction (rather than the substrate) *fits* very well to the enzyme active site.[34] Though the term *fit* goes beyond a simple geometric description. It is widely accepted that enzymes work by stabilizing the transition states of chemical reactions, but it is not clear in full detail by which means this task is fulfilled.

Chemists use in general acids, bases and nucleophiles as catalysts. Enzymes use the same chemistry but are more efficient due to a very subtle and equilibrated binding of substrates, transition states and products. The catalytic effect of enzymes is assumed to consist of several contributions but their relative significance is still discussed. In the next section I will introduce the basic parameters of enzyme kinetics and thermodynamics and show how they can be used to explain the principles of enzyme catalysis.

3.1.1. Basic enzyme kinetics

In contrast to an uncatalyzed chemical reaction in solution an enzyme catalyzed reaction occurs after the substrate has bound to the enzyme. Therefore, the binding of the substrate, the actual chemical conversion and the release of products determine the overall reaction rate. The reaction scheme of a typical enzymatic reaction is given in Fig. 3.1:

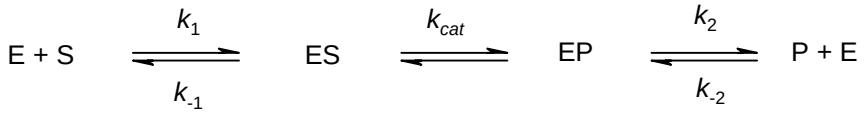

Figure 3.1.: Schematic representation of a typical enzymatic reaction. Key: E=enzyme, S=substrate, ES=enzyme-substrate complex, EP=enzyme-product complex, P=product. Each process has a reaction rate constant k_i, where k_1 and k_{-1} determine the binding of the substrate and k_2 and k_{-2} determine the release of the product. The rate constant k_{cat} determines the chemical conversion from substrate to product.

At low substrate concentration the velocity of the reaction is given by

$$v = \frac{k_{cat}}{K_M}[E_T][S] \tag{3.1}$$

where $[E_T]$ is the total enzyme concentration and K_M, the Michaelis constant, is defined as

$$K_M = \frac{k_{-1} + k_{cat}}{k_1} \tag{3.2}$$

If k_{cat} is very small compared to k_{-1}, the Michaelis constant can be approximated by the dissociation constant K_D of the enzyme substrate complex:

$$K_M \approx K_D = \frac{[E][S]}{[ES]} \tag{3.3}$$

From Eq. 3.1 the ratio k_{cat}/K_M can be considered as a second order rate constant. It is determined by both processes responsible for the overall reaction rate, which are substrate binding and the chemical conversion. This holds if the release of products is fast enough such that it does not block the binding of new substrate. The smaller K_M, the better the substrate binds to the enzyme. The chemical process of transforming the substrate to product is characterized by the nominator k_{cat}. The transformation does not need to be a one step reaction, but can consist of a sequence of elementary processes, that all contribute to k_{cat}. The term k_{cat}/K_M is the most critical parameter for enzyme kinetics. As the enzyme and the substrate cannot encounter more rapidly than diffusion in their solution permits, the velocity of the enzymatic reaction has an upper limit, which is $\sim 10^9 s^{-1} M^{-1}$.

The significance of k_{cat}/K_M becomes obvious from the following interpretation of Fig. 3.2: The overall barrier of an enzymatic reaction corresponds to the energy difference between $E+S$ and $ES^{\ddagger}$. This energy difference is determined by the values of ΔG_{bind} and ΔG_{cat}. The binding energy ΔG_{bind} is related to the Michaelis constant by the equation

$$\Delta G_{bind} = RT\,ln K_M. \tag{3.4}$$

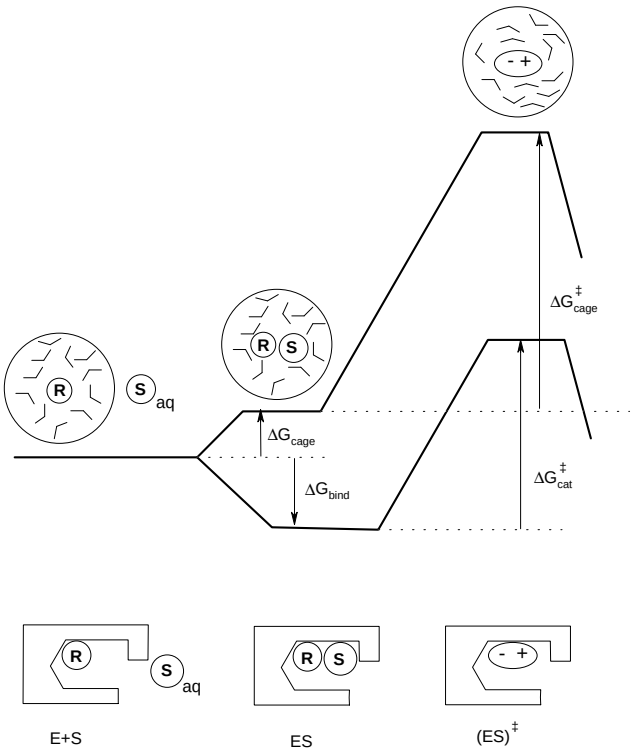

Figure 3.2.: A comparison of the free energy profile for an enzymatic reaction and for a reference reaction proceeding via an identical mechanism in a solvent cage. The upper curve shows the energetics of reference reaction in solution and the lower curve shows the energetics of the reaction in the enzyme. The symbols E,S, and S_{aq} denote the enzyme, the substrate, and the substrate in the bulk solvent, respectively. The symbol R denotes the reacting partner of the substrate. This might be a second substrate or a specific amino acid residue. The reacting particles S and R have to associate, such that a reaction can occur. In solution they have to gather into a solvent cage. This process is denoted by the energy ΔG_{cage}. For the enzymatic reaction the substrate has to bind in the active site. This binding is represented by the binding constant ΔG_{bind}.
The actual chemical conversion of the reacting partners is characterized by the activation barriers: The activation free energy $\Delta G^{\ddagger}_{cage}$ corresponds to a reaction in solution that obeys the same mechanism as assumed for the reaction in the enzyme. This value can be obtained from experimental information or from *ab initio* calculations. The free energy barrier of the enzymatic reaction is $\Delta G^{\ddagger}_{cat}$.

An enzyme-substrate pair has a small K_M value, if the enzyme binds the substrate well.

Moreover, the better the stabilization of the actual transition state in the enzyme, the smaller is the activation free energy $\Delta G^{\ddagger}_{cat}$. This activation energy is related to the rate k_{cat} of the reaction via an expression from transition-state theory.:[35]

$$k_{cat} = \frac{k_B T}{h} e^{-\Delta G^{\ddagger} \beta} \tag{3.5}$$

where h is Planck´s constant.

The overall barrier of the enzymatic reaction is therefore small if k_{cat}/K_M is large. Enzymes, that need to perform their catalytic function predominantly fast, have evolved to enlarge the ratio k_{cat}/K_M. As enzymes have to bind the transition state of a reaction as good as possible, the binding of the structurally different substrate cannot be optimal. Accordingly an enzyme should not evolve to yield a perfectly small K_M for its substrate. Preferable it should bind the substrate relatively weakly.

One part of the reduction of the overall barrier in the enzyme compared to aqueous solution can be accomplished via the substrate binding energy ΔG_{bind}. This would stabilize ES and $ES^{\ddagger}$ by the same amount. However, the binding energy alone cannot account for the overall reduction of $\Delta G^{\ddagger}$ as K_M should not be too small. This means, that $\Delta G^{\ddagger}_{cat}$ has to be reduced considerably compared to $\Delta G^{\ddagger}_{cage}$ in aqueous solution. The most important part of the catalytic effect results from this reduction of $\Delta G^{\ddagger}_{cat}$ by transition state stabilization. In general it is not clear what is the nature and mechanism of this stabilization. Enzyme transition states cannot be isolated experimentally and they cannot be understood completely without some quantitative model for structure-function correlation.

3.1.2. The chemistry of enzyme catalyzed reactions

A chemical reaction can be catalyzed in aqueous solution by a base or an acid. In Fig. 3.3 an example for general base catalysis is given.

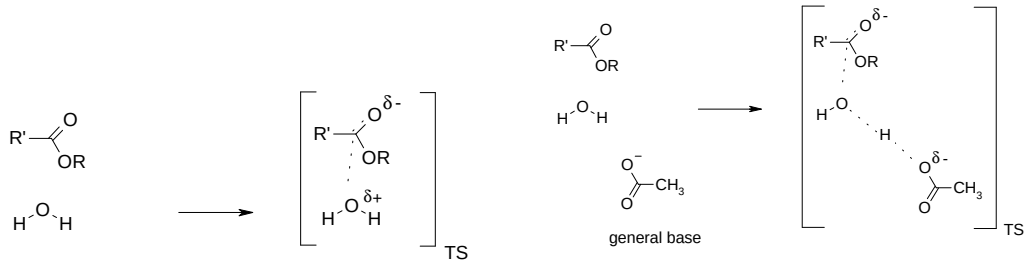

Figure 3.3.: The nucleophilic attack of a water molecule on the carbonyl C-atom of an ester. Figures (a) shows the uncatalyzed reaction. (b) The reaction is catalyzed by a general base that stabilizes the transition state(TS) by partly abstracting a proton from the water molecule. The negative charge is delocalized. The abstraction of the proton increases also the nucleophilicity of the water molecule.

In Fig. 3.3 (a) the uncatalyzed hydrolysis of an ester is shown. The transition state shows a positive and a negative partial charge. The transition state can be stabilized by the transfer of a proton from the water molecule to a general-base as shown in Fig. 3.3 (b). In analogy general-acid catalysis can be applied to stabilize a negative charge by a proton transfer from an acid. The term *general*-base or *general*-acid catalysis is used to distinguish it from *specific*-base or *specific*-acid catalysis, where the catalyst is the hydroxide-ion or the proton.

Acid and base catalysis is applied by chemists in solution as well as by enzymes. Examples of general-acid or base catalyzed reactions can be found in many enzymes *e. g.* in triose phosphate isomerase,[36,37] where a histidine works as general-acid and a glutamate as general-base. Histidine works also as general-base in serine proteases[38] like trypsin, chymotrypsin and subtilisin and in cysteine proteases[39] like papain and actinidin.

In the above example of general-base catalysis, the acceleration of the reaction is accomplished without changing the mechanism. Another very efficient means to catalyze a reaction is the modification of the reaction pathway. Examples for this type is the *nucleophilic catalysis* in acyl transfer- or hydrolytic reactions. In Fig. 3.4 the hydrolysis of acetic anhydride is shown. The rate of reaction is enhanced by pyridine, because the highly reactive acetylpyridinium-ion is formed rapidly. *Nucleophilic catalysis* can only be efficient, if the nucleophile that acts as catalyst (in Fig 3.4 pyridine) is more nucleophilic than the one it replaces (in our example acetate). Moreover the intermediate must be more reactive than the parent compound.

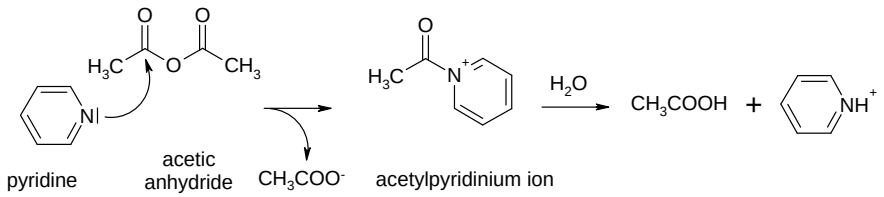

Figure 3.4.: The hydrolysis of acetic anhydride catalyzed by pyridine. This reaction is an example of catalysis, where the mechanism of the reaction is changed an a reactive intermediate is formed. As pyridine attacks the carbonyl carbon as nucleophile, it performs *nucleophilic catalysis*

Another contribution to the catalytic properties of enzymes might come from entropic effects. Uncatalyzed reactions may be unfavorable due to the loss of entropy, when the reacting partners form an adduct that can undergo the corresponding reaction. It was suggested that an enzyme uses the energy of substrate binding to compensate for the loss of entropy.[40,41] This theory is supported[42,43] but also questioned.[44,45] It is difficult to quantify the entropy in enzyme catalyzed reactions by experiments and also by theoretical calculations and simulations. A qualitative examination that investigated the entropic contributions to the catalysis in serine proteases indicated that these contributions are rather small.[45]

3.1.3. Electrostatic contributions in enzyme catalysis

It is widely accepted that electrostatic forces are perhaps the most relevant factors in enzyme catalysis. As many classes of enzymatic reactions involve large changes in the charge distribution of the reacting species it is likely that enzymes evolved to stabilize the charge distribution of the corresponding transition states. This indicates the importance of electrostatic interactions. The stabilization of charge distributions should be accomplished by enzymes better than by water,

which might be surprising, as water molecules can reorient to stabilize charges with their dipoles very efficient. The idea of electrostatic stabilization in enzymes is, that the enzyme has orientationally fixed dipoles, that are optimal for stabilizing transition states. Therefore, the enzyme does not have to provide the so called reorganization energy, whereas water molecules in bulk water have to. The reorganization energy was generated in the protein during to the folding process and is stored in the native protein structure.

Two examples *i. e.* the mechanisms of serine proteases and carbonic anhydrases shall demonstrate the significance of electrostatics in enzyme catalysis.

Serine proteases

Serine proteases (*e. g.* trypsin, chymotrypsin, subtilisin) catalyze the hydrolysis of peptides and esters. As shown in Fig. 3.5 the mechanism consists of the acylation of an active site serine residue. The serine residue is activated during the acylation step by transferring its hydroxyl proton to a nearby histidine residue. In the last step the acylated enzyme is converted into its native state by hydrolytic cleavage of the serine derivate. Accordingly, the serine protease mechanism involves general-base stabilization as well as nucleophilic catalysis.

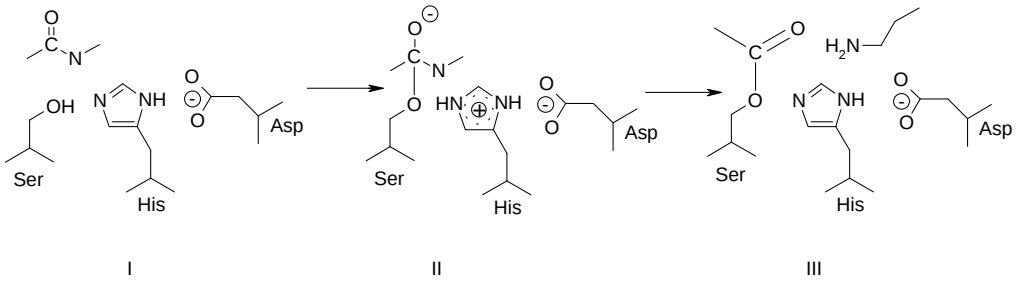

Figure 3.5.: The acylation step of the hydrolysis of a peptide group catalyzed by a serine proteases. The first state (I) shows the catalytic triad of the enzyme, consisting of Ser-His-Asp, in its initial state. The second state (II) is represented by the negatively charged tetrahedral intermediate. In the third state (III) the serine residue is acylated.

The active site of serine proteases consists mainly of a catalytic triad, that is built of serine, histidine and aspartate or glutamate. The reaction involves a proton transfer from serine to histidine yielding a positively charged histidine, which is stabilized by the acidic residue in the active site, namely aspartate or glutamate. It was discussed for a long time whether the stabilization of histidine is accomplished by another proton transfer from histidine to aspartate or glutamate, a process named the *double proton transfer* mechanism. Alternatively the stabilization could be performed by the electrostatic interaction between the charged groups, corresponding to the *charge relay* mechanism. It could have been shown both by experiments[46,47] and theoretical calculations[48,49] that in fact the charged groups are responsible for the stabilization, namely that the *charge relay* mechanism is more likely.

In theoretical studies on the reaction catalyzed by trypsin it was found that the energetic difference between the enzyme catalyzed reaction and an uncatalyzed reference reaction in aqueous solution results almost entirely from electrostatic effects.[49,50] The major aspect of the catalytic effect of serine proteases is associated with the electrostatic stabilization of the (- + -) charge distribution, represented by the Ser-His-Asp triad in the active site.

Carbonic Anhydrase

Carbonic anhydrase catalyzes the reversible hydration of carbon dioxide and the dehydration of hydrogen carbonate. The mechanism involves two processes: The activation of a zinc-bound water molecule, that donates a proton to His64 and a nucleophilic attack of the resulting hydroxide ion on carbon dioxide yielding hydrogen carbonate. The active site of carbonic anhydrase contains a zinc ion. (see Fig. 3.6) Its crucial role was determined quantitatively by showing that the cation reduces the pK_a-value of the bound water molecule by providing electrostatic stabilization of the hydroxide ion. The pK_a value of the bound water molecule in carbonic anhydrase is between 7 and 8.[51] Additionally the active site of the enzyme stabilizes the HCO_3^--ion by electrostatic interactions.[52]

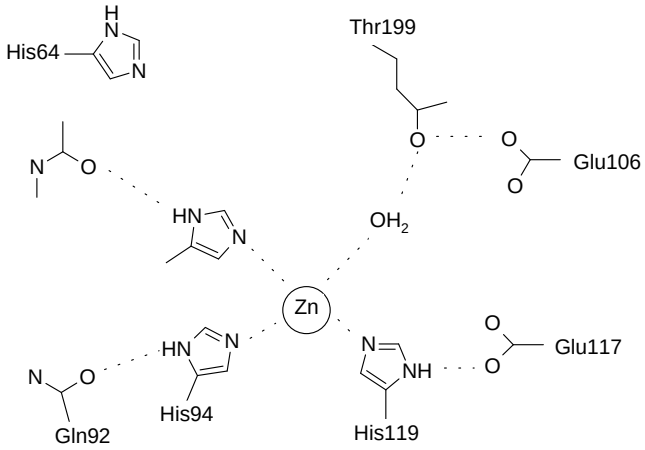

Figure 3.6.: A schematic representation of the active site of carbonic anhydrase. The zinc bound water molecule is activated and transfers a proton to His64. Then the resulting hydroxide ion attacks a carbon dioxide molecule yielding hydrogen carbonate. The active site stabilizes the hydroxide ion as well as the hydrogen carbonate by electrostatic interactions.

3.1.4. Reference reactions

The efficiency of enzyme catalysis can only be evaluated by comparing the catalyzed reaction with a suitable reference reaction. In section 3.1.1 we saw that the reduction of $\Delta G^{\ddagger}_{cat}$ is very important for enzyme catalysis. See also Fig. 3.2 that illustrates this effect. A recent analysis of the reaction of ribonuclease indicated for example that this enzyme provides a transition state stabilization $\Delta G^{\ddagger}_{cage} - \Delta G^{\ddagger}_{cat}$ of 18 kcal/mol.[53] In my work I concentrate on the investigation of this kind of stabilization *i. e.* the difference between the activation barrier of a reaction in a solvent cage and the reaction in the enzyme active site. The reaction in the solvent cage is a very good reference state because

- one can choose a reaction that has the same mechanism as the reaction in the enzyme. A possible difference in mechanism between the uncatalyzed reaction in water and the enzyme catalyzed reaction is known as a catalytic effect, but its investigation is not the aim

of my work as this effect is well understood in terms of nucleophilicity or pK_a-values (see section 3.1.2).

- possible concentration effets of reactions in aqueous solution cancel. The concentrations of the reacting species are the same in a solvent cage and in the enzyme active site.
- the focus is then concentrated on the actual task of enzymes: the stabilization of transition states and the advantage of the enzyme compared to water in this respect.

In this work I apply computer simulation studies to compare chemical reactions in aqueous solution and in enzymes. The theoretical tools for the simulation of chemical reactions in enzymes by computers are far from being usable as black box programs. The field is under intensive development and no standard algorithms or software have evolved. In the next section I will outline the general principles applied in these simulation studies.

3.2. How to simulate an enzymatic reaction?

To obtain the free energy surface of a reacting system it is necessary to sample all relevant conformations of the system. Such a search in the configurational space can be done with Monte-Carlo (MC) or Molecular Dynamics (MD) simulations. These methods generate conformations corresponding to energy potentials *e. g.* a mechanical force field.[54,55,56]

Enzymes are very large molecules and MD or MC studies on them are very time consuming even when simple potential functions are used as force field. To sample the conformational space of an enzyme substrate complex sufficiently to draw reliable conclusions, it is necessary to use a thoroughly parameterized classical molecular mechanics (MM) force field as it is described in appendix C. This involves always the evaluation of a large number of interactions between all atoms at each conformation. With fast computers and efficiently implemented algorithms it is nevertheless possible today to simulate an average protein in solution up to several ns.[57] Even simulations of up to 1 ms have been performed, but this is still an exception.[58]

Another situation arises with the simulation of chemical reactions. Their study requires quantum mechanical (QM) methods, as only quantum mechanics is able to describe bond breaking and bond formation processes. The computational time that is needed for QM calculations increases with the number of explicitly treated electrons N proportional to N^3 or N^4 . From the viewpoint of computational costs it is impossible until today and also for the near future to apply full QM calculations on a whole enzyme-substrate complex involving at least some thousand atoms.

A way out of this problem is the combination of quantum mechanical (QM) and molecular mechanical (MM) methods. This concept was first introduced in 1976 by Warshel and Levitt.[59] It devides the system into two regions: The reacting fragment is described by quantum chemical methods, whereas the environment is described classically. Since the initial idea of the QM/MM approach has been presented, various models of this principle have been published and were applied to enzymatic reactions.[60,61,62,63] The QM/MM method has recently been reviewed by Field.[64] The individual models differ in the special QM and MM methods used as well as in the treatment of the QM/MM interactions. The general Hamiltonian for a system consisting of differently described parts can be formulated as

$$H = H_{QM} + H_{MM} + H_{QM/MM} \tag{3.6}$$

where H_{QM} represents the Hamiltonian for the quantum system, H_{MM} the interaction within the surrounding medium and $H_{QM/MM}$ the interaction between the quantum region and the surrounding medium. The level of quantum chemical methods employed range from rather accurate *ab*

initio Hartree-Fock and Density Functional methods to semiempirical approaches and empirical valence bond descriptions. Due to their computational costs, with *ab initio* calculations it is only since recently possible to calculate actual free energies, instead of just enthalpies at several points along the reaction coordinate.[65,66] Semiempirical QM methods have lower computational demands, but they are also less accurate. The most widely applied QM/MM approach to simulate enzymatic reactions is the empirical valence bond method (EVB), that is discussed in the next section.

3.3. The empirical valence bond method

The wish to obtain quantitatively reliable results of free energy surfaces from computer simulations of enzymatic reactions in an acceptable time, lead to the formulation of the empirical valence bond model.[45,52] The most prominent feature of this method is the calibration procedure, that allows to fit some inherent EVB parameters, such that the free energy curve of a simulated reference reaction in aqueous solution is in agreement with energies obtained from experiments. This unique parameterization scheme makes the EVB method until today the most powerful tool in simulations of enzymatic reactions.

The EVB method models a chemical reaction in the simplest case as a transition between two valence bond structures or resonance states corresponding to the reactants and products of the reaction. Each of the two states is described by a separate molecular mechanics force field. The transition from the reactant state to the product state is accomplished by a linear combination of the two energy functions, yielding the mapping potential:

$$\varepsilon_{map}^{1,2}(\lambda) = (1.0-\lambda)\varepsilon_1 + \lambda\varepsilon_2 \quad \text{where}\, \lambda\varepsilon[0,1] \tag{3.7}$$

With this mapping potential $\varepsilon_{map}(\lambda)$, the reaction is guided from the reactant state to the product state by varying the parameter λ from 0 to 1. The potentials ε_1 and ε_2 are described by the energy functions:

$$\begin{aligned} \varepsilon_\mu &= \sum_j \Delta M_j^{(\mu)}(b_{0,j}^{(\mu)}) + \frac{1}{2}\sum_l \gamma_l^{(\mu)} k_l^{(\mu)}(\theta_l^{(\mu)} - \theta_{0,l}^{(\mu)})^2 \\ &+ \frac{1}{2}\sum_m \gamma_m^{(\mu)} K_m^{(\mu)} \cdot [1 + cos(n_m^{(\mu)}\phi_m^{(\mu)} - \delta_m^{(\mu)})] + V_{nb,rr}^{(\mu)} + \alpha^{(\mu)} + V_{nb,rs}^{(\mu)} + V_{ss}, \mu = 1,2 \end{aligned} \tag{3.8}$$

The subscripts r and s stand for the reacting fragments and the surrounding of the molecular system, which can be aqueous solution or the protein/water environment, respectively. The first term in Eq. 3.8 represents the Morse potential of forming and breaking bonds relative to its minimum value $b_{0,j}^{(\mu)}$ for the jth bond in the valence bond state μ. The second and third term denote the corresponding potentials of bond angles and torsional angles. The energies of those bond or torsional angles, that are involved in forming or breaking bonds, are coupled to the corresponding bond strength by the coupling factor $\gamma_l^{(\mu)} = |\Delta M_j^{(\mu)}/D_j^{(\mu)}|$, where $D_j^{(\mu)}$ is the dissociation energy of the bond j. $V_{nb,rr}$ and $V_{nb,rs}$ are the nonbonded energies (van der Waals and electrostatic) among the reacting atoms and between reacting atoms and the surrounding. The van der Waals interaction between atoms, that form a new bond during the reaction or whose bond is broken, is modified to an exclusive repulsive interaction in the state where the bond is not present. The interaction has the form $V_{rep} = C_{ij}e^{-ar}$, where the constants C and a depend on the atom types. The last term V_{ss}, lumps together all energy contributions of the non reacting part of solvent and protein.

Conventional force fields describe molecular systems with an energy function corresponding to one chemical state only. They do not provide any relation between the energies of resonance states possessing different bonding schemes, that result in alteration of Lennard-Jones or charge parameters. The energy difference between two resonance states is given by the difference of heat of formation in the gas phase and is thus introduced in the energy function by the constant energy parameter $\alpha^{(\mu)}$.

The described energy function is used to sample the configurational space of the system by molecular dynamics. Subsequently, the adiabatic electronic ground state energy $E_g^{1,2}$, mediating the transition between the states 1 and 2, is calculated at all configurations:

$$E_g^{1,2} = E_g^{1,2}(\varepsilon_1 - \varepsilon_2) = \frac{1}{2}(\varepsilon_1 + \varepsilon_2) - \frac{1}{2}\sqrt{(\varepsilon_1 - \varepsilon_2)^2 + 4{H_{12}}^2} \tag{3.9}$$

E_g is the lower of the two eigenvalues of a 2×2 Hamiltonian matrix, where the diagonal elements are the potentials ε_1 and ε_2, from Eq. 3.8 and the off-diagonal element H_{12} is represented by

$$H_{12} = A_{12} exp[\mu_{12}(r_{XY} - r_{XY}^{\ddagger}) - \eta_{12}(r_{XY} - r_{XY}^{\ddagger})^2] \tag{3.10}$$

where r_{XY} is the distance between two atoms that are involved in a bond breaking or forming process in the reaction considered and can be used to monitor to which extent the reaction has evolved. The distance $r_{XY}^{\ddagger}$ is a constant, chosen such that H_{12} adopts a maximal value close to the transition state.

The free energy difference of the transition of state 1 and state 2 is evaluated with the free energy perturbation approach (FEP):[67]

$$\Delta G_{map}(\lambda_j) = G_{map}(\lambda_j) - G_{map}(\lambda_0) = -k_B T \sum_{i=0}^{j-1} ln\langle exp[-(\varepsilon_{map}(\lambda_{i+1}) - \varepsilon_{map}(\lambda_i))/k_B T]\rangle_i \tag{3.11}$$

$\Delta G_{map}(\lambda_j)$ represents the free energy associated with moving on the mapping potential $\varepsilon_{map}^{1,2}(\lambda)$ defined in Eq. 3.7.

The generalized reaction coordinate X, is defined as the energy gap $\varepsilon_1 - \varepsilon_2$ of the two energy functions at each conformation of the trajectory. The reaction coordinate X is partitioned in a number of bins X_m, typically 50. The free energy, $\Delta G^{1,2}(X)$, corresponding to the trajectories moving on the adiabatic energy surface of the electronic ground state E_g (Eq. 3.9) mediating the transition from the reactant state to the product state, is obtained with the umbrella sampling expression

$$\begin{aligned} &exp[-\Delta G^{1,2}(X_m)/k_B T] = \\ &\quad \frac{1}{N(X_m)} \sum_j n(j, X_m) exp[-\Delta G_{map}^{1,2}(\lambda_j)/k_B T] \cdot \langle exp[-(E_g^{1,2}(X_m) - \varepsilon_{map}(\lambda_j))/k_B T]\rangle_{j,X_m} \\ &\quad \text{where } N(X_m) = \sum_j n(j, X_m) \end{aligned} \tag{3.12}$$

The Boltzmann factor of the free energy $\Delta G^{1,2}(X_m)$ is obtained as a weighted sum over the different λ_j ensembles, $n(j, X_m)$ is the number of conformations from the λ_j ensemble, where the reaction coordinate X belongs to bin m. The ensemble average in Eq. 3.12 considers all conformations, where the reaction coordinate corresponds to bin m and λ_j. This statistical average accounts for the energy difference between the mapping potential $\varepsilon_{map}^{1,2}$ used for the MD simulation and the adiabatic potential energy surface $E_g^{1,2}$ of the reaction considered.

An important aspect of the EVB procedure is, as stated above, the calibration of the EVB parameters in H_{12} and of α. The parameters are chosen such that in a simulated reference reaction in aqueous solution the resulting free energy profile reproduces the experimental value of the activation barrier and of the free reaction energy. In practice, the reaction of interest is simulated in a water sphere and the obtained energy data are evaluated within the FEP procedure, in which the parameters α and H_{12} can be varied to yield the proper free energy curve. A detailed discussion about setting up the reference energy profile is given in section 4.3.2, within the discussion of the simulation of the deacylation step in acetylcholinesterase. These calibrated parameters are then also used to evaluate the energy profiles from a simulation in which the water sphere has been replaced by the enzyme and water molecules to fill the protein cavities. With this approach, the effect of the enzyme environment on the free energy reaction profile is obtained.