Reprinted with permission from {J. Phys. Chem. B, 2014, 118 (46), pp 13050–13058}.

Copyright {2014} American Chemical Society.

Deposition into DIVA by permission of publisher 2015.

A Novel Approach for Identifying Key Residues in Enzymatic Reactions: The Case of Proton Abstraction in Ketosteroid Isomerase

*Mika Ito and Tore Brinck**

Applied Physical Chemistry, KTH Royal Institute of Technology, Teknikringen 30, 100 44
Stockholm, Sweden

Abstract: We propose a computationally efficient approach for evaluating the individual contributions of many different residues to the catalytic efficiency of an enzymatic reaction. This approach is based on the fragment molecular orbital (FMO) method and it defines the energy of a deletion form, i.e. the energy of the system when a particular residue is deleted. Using this approach, we found that among ten investigated residues, three, Tyr14, Asp99, and Tyr55, in this order, significantly reduce the activation energy of the proton abstraction from a substrate, cyclopent-2-enone, catalyzed by ketosteroid isomerase (KSI). The relative activation energies estimated in this study are in good agreement with available previous experimental and theoretical data obtained for the similar proton abstraction with a native substrate and substitution mutants of KSI. It was thus indicated that the new approach is efficient for rationally evaluating the catalytic effects of multiple residues on an enzymatic reaction.

Keywords: enzyme design, enzymatic catalysis, fragment molecular orbital method, deletion form, activation energy

Introduction

Enzymes catalyze chemical reactions in vivo at high rates with substrate specificity and stereoselectivity. These properties, together with the benefits of biocatalysts from a green chemistry perspective, make enzymes of potential interest for manufacturing of specialized chemicals, and particularly for synthesis of building blocks for the pharmaceutical industry. There have been many experimental¹⁻⁵ and theoretical studies⁶⁻¹⁰ aimed at designing enzymes to produce efficient catalysts for targeted reactions. Although this research area has seen significant progress in recent years, there is a great need for the development of new and improved methods to aid the design process.

In a recent theoretical study¹¹ with the aim to design a catalyst for the Diels-Alder reactions, it was found that a native enzyme, ketosteroid isomerase (KSI), has the potential to catalyze an acid-base induced Diels-Alder reaction, and it was indicated that its catalytic efficiency can be improved further by rational design. This is an attractive finding, since Diels-Alder is an important reaction for synthesis of cyclic compounds, and cyclic structures are common motifs in drug candidates.

In order to rationally modify KSI with the objective of improving the catalytic efficiency for the acid-base induced Diels-Alder reaction, the residues which are most important for the catalysis should be identified. As a first step of our study, we focused on a prerequisite step (Figure 1) of the KSI-catalyzed Diels-Alder reaction, and tried to devise an efficient method for rationally performing quantitative assessment of the influence of different residues on the catalytic efficiency. In the prerequisite step, Asp38 serves as a general base and abstracts a proton from a substrate, cyclopent-2-enone, inducing the change in the structure of the substrate from an enone form (cyclopent-2-enone) to an enolate form (cyclopenta-1,4-dienolate); the two residues Tyr14 and Asp99 stabilize the enolate form of the substrate (The

residues are numbered according to the sequence of *Pseudomonas testosteroni* KSI throughout the text.). Which are the key residues among the residues surrounding the reactants (the substrate and Asp38) in this proton abstraction step? You can easily guess that residues, like Tyr14 and Asp99, which have direct interactions with the substrate should be most important for the catalysis. However, also residues which do not interact directly with the substrate could have significant effects on the catalytic efficiency.

Unfortunately there are no reference experimental data on the key residues of the KSI Diels-Alder reaction, since this is a novel reaction designed in the previous theoretical study¹¹, which has not yet been examined experimentally. Fortunately, the initial step of the designed KSI reaction and the initial step of the native KSI reaction, which is isomerization of ketosteroids, are almost identical, i.e. proton transfer from a substrate to an Asp. The catalytic reaction and the active site structure of KSI have been studied extensively by experimental techniques¹²⁻²¹ and by theoretical computations²²⁻²⁸. It has been shown that KSI works by a preorganized active site which electrostatically stabilizes the intermediate state and the transition states via hydrogen bonds to the substrate.^{19,23,27,28}

In an experimental study¹⁷, the native KSI reaction, in which Δ^5 -3-ketosteroid (5-androstene-3,17-dione) is transformed into Δ^4 -3-ketosteroid (4-androstene-3,17-dione) via a dienolate intermediate, was examined, and rate constants obtained with wild-type (WT) KSI and some of its mutants were determined. In that study, Tyr14 and Asp99, which are hydrogen bonded to the substrate, and also Tyr55, which does not form a hydrogen bond with the substrate, were targeted for mutations. The rate constants of the Tyr14Phe, Asp99Leu, and Tyr55Phe mutants are about 3 orders, 2 orders, and 1 order of magnitude, respectively, smaller than the rate constant of the WT, indicating that Tyr14, Asp99, and Tyr55 are important for the KSI catalysis. However, it is not obvious from the experimental study whether these residues have

important roles in the initial proton transfer step, because the rate constants were determined for the whole reaction, which includes two proton transfer steps. The individual steps of this reaction in the WT were examined in an earlier experimental study²⁰, and it was shown that the rate constants of the two proton transfer steps of the forward reaction in the WT are likely to be similar, where activation free energies of the first and second proton transfer steps were estimated to be 10.3 and < 10.6 kcal/mol (at 25 °C), respectively. In another experimental study²¹, it was shown that these two rate constants of the same reaction in the Asp38Glu mutant are similar, where activation free energies of the first and second steps were estimated to be 13.9 and 13.7 kcal/mol (at 25 °C), respectively. This study further showed that the activation enthalpy of the first step (7.6 kcal/mol) is much lower than that of the second step (15.5 kcal/mol). Rate constants of the individual proton transfer steps in the Tyr14Phe, Asp99Leu, and Tyr55Phe mutants have not been reported.

In a theoretical study²⁷, the same native KSI reaction was analyzed by molecular dynamics (MD) simulations with a two-state empirical valence bond (EVB) potential and the potential of mean force. Rate constants of each proton transfer step of the reaction catalyzed by WT KSI and its Tyr14Phe and Asp99Leu mutants were calculated. This study showed that the rate constants of the first and second proton transfer steps of the forward reaction in the WT are very similar. Furthermore, the rate constant of the first proton transfer is significantly decreased by the Tyr14Phe and Asp99Leu mutations whereas the rate constant of the second proton transfer is increased by these mutations. It is thus indicated that the main cause for the decrease in the experimental rate constants due to the Tyr14Phe and Asp99Leu mutations is lowered rate constants of the first proton transfer step. In theoretical studies^{26,27}, structural features of the WT were also analyzed, and hydrogen bond distances from the substrate to Tyr14, Asp99, and Tyr55 and a hydrogen bond distance between Tyr14 and Tyr55 were calculated. This analysis showed that, similarly to the substrate-Tyr14 and substrate-Asp99

hydrogen bonds, the Tyr14-Tyr55 hydrogen bond was strengthened during the first proton transfer steps and weakened during the second proton transfer. The substrate does not form a hydrogen bond with Tyr55 during either of these steps. It is clear from these and earlier studies that Tyr14 and Asp99 catalyze the reaction by stabilizing the negative charge on the substrate oxygen that is built up during the formation of the intermediate. The hydrogen bond analysis indicates that the rate enhancing effect of Tyr55 can be explained by the strengthening of hydrogen bond to Tyr14 during the first proton transfer. However, it has been suggested that the main catalytic contribution of Tyr55 comes from the positioning of Tyr14 for hydrogen bonding to the substrate oxygen ¹⁷.

Together with the experimental study, the theoretical studies show that not only Tyr14 and Asp99 but also Tyr55 are important for the catalysis of the initial proton transfer step of the native KSI reaction. However, also other residues may be of importance for reducing the reaction rate of this step.

Generally, the importance of particular residues in enzymatic reactions has been analyzed by site-directed mutagenesis. It takes a lot of time to generate mutants by experimental means. Similarly, the computational time increases linearly with the number of mutations in a theoretical analysis. Consequently, the number of mutants often has to be limited, and only a few residues located at positions adjacent to or close to a substrate are targeted for mutation. Hence, it is easy to miss key residues that are located outside the direct vicinity of the substrate. Additionally, it is known that different mutations can give different effects on an enzymatic reaction; in the case of the KSI isomerization, the decrease in the rate constant caused by the Tyr14Phe mutation is about 2 orders of magnitude larger than the decreases caused by the Tyr14Ala and Tyr14Gly mutations.¹⁸ It was suggested that substituting Tyr14 with smaller residues results in the formation of a water cavity that stabilizes the negative

charge on the substrate oxygen via hydrogen bonds.¹⁸ This example shows that it is not always straightforward to analyze the catalytic contributions of individual residues using site-directed mutagenesis. An alternative is to delete the residues, one by one, from the structure in order to evaluate their function. Experimentally, this is applicable only to residues of N- and C-termini, because deletion of residues of middle regions, where active sites of enzymes are usually located, destroys the structures and functions of proteins. On the other hand, these problems can be overcome in theoretical modeling by freezing the structure to that of the native enzyme. Still, deletions are time consuming and computationally demanding to analyze theoretically, e.g. if one wants to analyze the effect of deletions on a kinetic rate constant it would be necessary to do at least two computations (reactant and transition state) for each deletion to analyze. Thus, the computational time increases essentially linearly with the number of deletions. Furthermore, in a quantum chemical study deletions of covalently bonded residues can create problems due to the necessity to break bonds. In this study we will demonstrate that it is possible to overcome these problems by using the fragment molecular orbital (FMO) method²⁹⁻³².

The FMO method is an approximate molecular orbital (MO) method for calculating large molecules with thousands of atoms. It is known that conventional MO and FMO calculations give almost identical results when used at the same theoretical level and that the computational cost of an FMO calculation for a large molecule is much lower than that of an MO calculation for the same system.^{33,34} The FMO method has been applied to biomolecular systems aiming at better understanding of biological phenomena including enzymatic reactions³⁵⁻³⁸ and at rational drug design etc.. In many FMO studies, binding modes of protein-small molecule (ligand or substrate), protein-protein, and protein-nucleic acid complexes have been investigated by analyzing mostly the following energy aspects: binding energies and energies of interactions between components of the complexes (for example,

residue-residue interactions) at the molecular, atomic, and orbital levels.³⁹ Based on the data on these energy aspects, amino acid residues or nucleic acid bases which are important in the binding modes can be found. However, key residues that are important in accelerating reactions in enzymes may not be found by these types of approaches, because binding energies and interaction energies may not correlate well with activation energies, which are related to reaction rates.

In our FMO approach, we use an original form of an active site model and its deletion forms in which particular residues are virtually deleted, and we calculate the energies, such as activation energies, heats of reaction, etc., of the original form and the deletion forms by using only one active site model of WT KSI. This approach enables us to evaluate individual contributions of many residues to the catalytic efficiency simultaneously, and does not require an initial guess of which residues are most likely to influence the catalytic efficiency. In order to examine the reliability of this new approach, we estimate the rate enhancing effects of the catalytic residues, and compare them to available experimental and theoretical rate constants obtained for the similar KSI proton abstraction. In addition, we apply another approach which is based on the conventional MO method to see if the same key residues can be identified by the MO and FMO approaches and to demonstrate the rationality of the FMO approach. In the MO approach, we build several small active site models, and compare their properties to the properties of a large active site model. This approach is applicable to the cases like the KSI proton abstraction, where one can guess a few key residues so as to build several small active site models. Otherwise, numerous models of small and medium sizes would be needed, and a lot of time to perform MO calculations for the models would be required. Through this study, we demonstrate that our novel FMO approach is powerful for rationally identifying the influence of individual residues on the enzymatic reaction, and

suggest that this approach can be useful in identifying key residues in other enzymatic reactions.

Methods

MO Calculations

Conventional Kohn-Sham density functional theory (DFT)⁴⁰ calculations were performed for five small and one large active site models (Figure S1). The one large model (Model6), which consists of the substrate (cyclopent-2-enone), the general base (Asp38), and ten surrounding residues including Tyr14, Asp99, and Tyr55 was built using the X-ray crystal structure of *Pseudomonas testosteroni* KSI (PDB code 1QJG)¹⁶ in the previous theoretical study¹¹, and the five small models (Models 1-5), which consist of the substrate, the general base (Asp38), and zero-three surrounding residues (Tyr14, Asp99, and Tyr55), were built using the structure of the large model (Model6). In these models, only the side chains of the residues were considered and ends of the side chains were capped by an H atom or a CH₃ group mimicking the backbone. Geometry optimizations were carried out for the active site models using the M06-2X functional⁴¹ with the 6-31G(d) basis set in the gas phase. In the geometry optimizations, two atoms (a C and an H atoms) at the end of each residue were fixed to keep the position of the corresponding backbone of the KSI active site, because it is known that the KSI has a preorganized active site²⁸ and that the backbone of the KSI active site hardly moves during the two proton transfer steps of the isomerization²⁷. Frequency calculations were performed for all stationary points to verify that they are minima or transition states. Energies were calculated for the optimized structures at the M06-2X/6-311++G(d,p) level in the gas phase, water ($\epsilon = 78.4$), and a nonaqueous solvent, diethyl ether ($\epsilon = 4.2$). Diethyl ether was used to simulate the environment of the protein interior, whose dielectric constant (ϵ) is considered to be low (about 2-5)⁴²⁻⁴⁶. Free energies in solution, computed as sums of gas phase energies and solvation free energies, were obtained by the polarizable continuum

model (PCM)^{47,48} using the SMD parameterization⁴⁹. All MO calculations were carried out with the Gaussian09 program⁵⁰.

FMO Calculations

FMO calculations were performed for the large active site model (Model6). Energies were calculated for the M06-2X optimized structures using the Grimme's SCS-MP2 method^{51,52} with the 6-31G(d,p) basis set in the gas phase. Diffuse *s*- and *p*-functions were added to the basis set for anionic oxygen atoms. The substrate-Asp38 pair was assigned as a single fragment, and all other residues were assigned as single fragments. All FMO calculations were carried out with the PAICS program⁵³.

In the proposed approach, we used the deletion forms, where in each a particular residue is virtually deleted from the model system. To obtain the energy of the deletion form, the energy of the residue targeted for deletion and the energies of all the interactions between this target residue and the remaining components of the model system were subtracted from the total energy of the model system. The deletion form corresponds to a special mutant in which a mutated residue does not interact with other residues and a substrate. This leads to a straightforward method for identifying the key residues of a native enzyme. The energy of the deletion forms can be calculated by taking advantage of the FMO procedure, which divides a large molecule into small fragments. It is not possible to obtain the energy of a deletion form by the regular MO theory in the case where the deleted residue is covalently bonded to other parts of the model system.

The energy of the deletion form was defined as follows using the FMO description of the total energy of a system. When a system is divided into N fragments, the total energy of the system (E_{total}) is expressed with the energies of a fragment (E_I) and a fragment pair (E_{IJ}), which include the environmental electrostatic potential (ESP),^{29,30} as

$$E_{total} = \sum_{I>J}^N E_{IJ} - (N-2) \sum_I^N E_I .$$

By using the approximation of ESP, the total energy of the system is expressed with the energy of a fragment without ESP (E'_I) and the inter-fragment interaction energy (ΔE_{IJ})³² as

$$E_{total} = \sum_I^N E'_I + \sum_{I>J}^N \Delta E_{IJ} .$$

By using this equation, the energy of the deletion form (E_{del}) can be defined as

$$\begin{aligned} E_{del} &= E_{total} - E'_I - \sum_{J \neq I}^N \Delta E_{IJ} \\ &= \sum_{J \neq I}^N E'_J + \sum_{\substack{J>K \\ J \neq I \\ K \neq I}}^N \Delta E_{JK} , \end{aligned}$$

where the residue targeted for deletion is assigned as fragment I .

In this study, the energies of R, TS, and INT of the original (no deletion) form and the ten deletion forms of the model system, that is 33 different energies in total, were obtained from the FMO calculations for only the 3 (R, TS and INT) structures of the original form. Using a conventional MO approach, calculations for 33 structures would be needed to obtain the 33 different energies. Thus, significant saving of computational time was obtained by the FMO approach.

To estimate the individual energy contributions of the residues to the reaction, the energies of TS and INT relative to the energy of R, which are the activation energy ($\Delta E^\ddagger$) and the heat of reaction (ΔE), respectively, were calculated for the each deletion form, and were compared to those energies of the original form. In order to relate computed energies to kinetic rate

constants we applied the transition state theory^{54,55}. According to this theory, the rate constant (k) for a reaction is given as

$$k = \kappa \frac{k_B T}{h} e^{-\frac{\Delta G^\ddagger}{RT}} = \kappa \frac{k_B T}{h} e^{\frac{\Delta S^\ddagger}{R}} e^{-\frac{\Delta H^\ddagger}{RT}},$$

where κ is the transmission coefficient ($\kappa \approx 1$), h is the Planck constant, k_B is the Boltzmann constant, T is the thermodynamic temperature, R is the gas constant, $\Delta G^\ddagger$ is the activation free energy, $S^\ddagger$ is the activation entropy, and $\Delta H^\ddagger$ is the activation enthalpy. In our calculations, the activation energies ($\Delta E^\ddagger$) that we obtain for the deletion forms do not include effects due to zero-point vibrations and enthalpy and entropy corrections, and thus the $\Delta E^\ddagger$ are not directly comparable to $\Delta H^\ddagger$ or $\Delta G^\ddagger$. However, it can be expected that such effects show small variations upon deletions or mutations, and thus that the relative values $\Delta \Delta E^\ddagger$ and $\Delta \Delta G^\ddagger$ should be comparable. The entropic contributions to the free energies are assumed to be relatively small, since it is known that KSI has a preorganized active site²⁸ and that the KSI active site exhibits highly restricted motion during the two proton transfer steps of the isomerization²⁷.

Results and Discussion

Results from MO Calculations

Geometry optimizations of the reactant state (R) and intermediate state (INT), where the substrate adopts the enone and enolate forms, respectively (Figure 1), were carried out for the five small active site models (Models 1-5) (Table 1, Figure S1). In the case of Model1, which consists of only the substrate and the general base (Asp38), a stable INT structure was not obtained, since all attempts at optimizing the enolate form resulted in transfer of a proton from Asp38 to the substrate and formation of the enone form. For Models 2-5, stable INT structures, as well as stable R structures, were obtained (Figure S2), though the substrate adopts an enol form instead of the enolate form in the INT structure for Model3 due to transfer of a proton from Asp99 to the substrate. This indicates that Tyr14 hydrogen bonded to the substrate is needed to stabilize the enolate form in INT. Geometry optimizations of the transition state (TS) were carried out for Models 2-5, and their TS structures were obtained (Figure S2). Optimized structures of R, TS, and INT for Model6 (Figure S2) were taken from the previous theoretical study¹¹.

The relative energies (Table 1) of Model2 and Model3 show that Tyr14 and Asp99 have similar effects on the stabilization of TS, and that the effect of either Tyr14 or Asp99 is not sufficient to make INT more stable than R. The INT of Model3 is more stabilized than that of Model2, which most likely is an effect of the proton transfer in Model3. The combined effect of the two residues in Model4 stabilizes TS and INT further, and results in INT being higher in energy than R by only 0.6 kcal/mol. The calculations on Model5 show that also Tyr55 has a substantial effect on the stabilization of TS and INT, and with the inclusion of Tyr55 INT becomes more stable than R by -3.6 kcal/mol, even though there is no hydrogen bond between Tyr55 and the substrate. The observed effects are most probably due to the hydrogen

bond between Tyr55 and Tyr14, which becomes stronger along the reaction coordinate. The relative energies of TS and INT for Model5 are similar to those for Model6, indicating that the remaining active site residues have minor effects on the catalytic reaction.

Additional analyses of structures and energies were performed for Model5. The hydrogen bond distances of Model5 (Figure 2) show that the Tyr14-Tyr55 hydrogen bond, as well as the substrate-Tyr14 and substrate-Asp99 hydrogen bonds, is strengthened in TS and INT compared to R. This structural feature of Model5 is also seen in Model6, and the hydrogen bond distances of Model5 are similar to those of Model6. The relative free energies in solution of TS and INT for Model5 (Figure 3) show that TS and INT are more stable in diethyl ether than in water. The reason for this is probably that, in contrast to the delocalized negative charge of the substrate in TS and INT, the negative charge of Asp38 in R is localized and less stabilized in diethyl ether than in water. A similar effect is found for Model6, and the relative free energies in solution of TS and INT for Model5 are similar to those for Model6.

It is thus suggested that the main catalytic effects come from the three residues (Tyr14, Asp99, and Tyr55) among the ten residues that were considered in our study. To confirm this assumption, we need to estimate the energy contributions from all the ten residues. As we will demonstrate, the FMO method is very useful for such an analysis.

Results from FMO Calculations

The FMO calculations were performed for Model6. The relative energies of TS and INT calculated by the FMO method at the SCS-MP2 level are similar to those calculated by the MO method at the M06-2X level (Figure 4). Due to the favorable scaling of computer time with molecular size for the FMO method, the large Model6 could be analyzed using the computationally demanding SCS-MP2 method at a reasonable computational cost. The

activation energies for the proton abstraction from the substrate (cyclopent-2-enone) are 8.0 and 7.2 kcal/mol at the SCS-MP2 and M06-2X levels of theory, respectively. According to an experimental study²⁰, the activation free energy for the first proton transfer step in the reaction of WT KSI with the native substrate (5-androstene-3,17-dione) is 10.3 kcal/mol (at 25 °C). According to another experimental study²¹, the Asp38Glu mutant, which is believed to react by the same mechanism, has an activation free energy of 13.9 kcal/mol (at 25 °C) and an activation enthalpy of 7.6 kcal/mol for the same step. On the basis of the assumption that the entropy effects are similar for the two proteins, the activation enthalpy for the WT can be estimated to be around 4 kcal/mol. This value is circa 4 kcal/mol lower than our SCS-MP2 activation energy. However, vibrational effects, and in particular the zero point effect, are expected to lower the theoretical barrier further by around 2 kcal/mol. Furthermore, the substrate used in our study is slightly different in structure. Thus, there is satisfactory agreement between experiment and theory.

The interaction energies between the central fragment, which consists of the substrate and the general base (Asp38), and the fragments of the surrounding residues were calculated (Figure 5). The interactions of Tyr14 and Asp99 with the central fragment are quite strong compared to those of the other residues. This is because Tyr14 and Asp99, unlike the other residues, are hydrogen bonded to the substrate. The interactions of Tyr14 and Asp99 are strengthened as the reaction proceeds. This is consistent with the changes in the hydrogen bond distances (Figure 2). The interaction of Tyr55 with the central fragment is relatively weak. This is expected, since Tyr55 is hydrogen bonded neither to the substrate nor to Asp38. It is noteworthy, though, that the interaction of Tyr55 becomes strengthened as the reaction proceeds.

The energies of TS and INT relative to the energy of R were calculated for the original form and the deletion forms. The deletions of Tyr14, Tyr55, and Asp99 destabilize TS and INT (Figure 6) in line with the observed effects from the MO computations. It should be noted though that the deletion of residues results in larger energy changes in the FMO computations than in the MO models. This is particular true for the reaction energies (formation of INT), where deleting Tyr14 and Asp99 increases the reaction energy by 18.7 and 11.5 kcal/mol, respectively in the FMO computations (Table 2). These unphysically large energy differences are due to an effect of the lack of structure relaxation in the FMO approach. The deletions Tyr14 and Asp99 will significantly affect the structure of the INT intermediate and this will result in overestimation of the corresponding INT energies. This effect is particularly large for the Tyr14 deletion, where the MO model indicates that the INT exists in the enol form and the FMO method uses the unrelaxed enolate structure. The problem is less severe for the activation energy, since deletions has a smaller effect on the transition state structure.

In contrast to the deletions of the residues Tyr14, Tyr55 and Asp99, the deletion of Phe54 slightly stabilizes TS and INT (Figure 6). However, this effect is rather small as are the effects of deleting other residues (Figure 6). It is thus confirmed that the main energy contributions to the catalysis come from the Tyr14, Tyr55 and Asp99 residues.

The relative activation energies ($\Delta\Delta E^\ddagger$) of the Tyr14, Tyr55, and Asp99 deletion forms calculated by the FMO method were compared to the relative activation free energies ($\Delta\Delta G^\ddagger$) of the Y14F, Y55F, and D99L mutants calculated by the transition state theory using the rate constants obtained in an MD study²⁷ as well as from an experimental study¹⁷ (Table 2). Note again that the substrate (cyclopent-2-enone) used in our study is different from the native substrate (5-androstene-3,17-dione) used in the other studies. However, the proton abstraction reactions for these substrates are similar in the sense that a proton bonded to a carbon atom

which is adjacent to the carbonyl group of the cyclic substrate is transferred to Asp38. In addition, one should be careful when comparing the uncorrected Born-Oppenheimer activation energies with free energies of activation. Still, it is likely that the zero-point vibrational effects and the thermal effects (including entropy effects) are similar in the wild type and the mutants and thus they should have relatively small effects on the relative barriers.

We first compare our FMO results with the results of the MD study (Table 2), which refer to first proton transfer step of the native reaction. The results of the MD study show that the Tyr14Phe and the Asp99Leu mutations lower the activation free energy by 4.9 and 3.4 kcal/mol, respectively. These results are in relatively good agreement with the corresponding deletions with the FMO method that lower the barrier by 5.1 and 4.3 kcal/mol, respectively. The results of the experimental study are not directly comparable to our FMO results since the experimental activation free energies are computed from the k_{cat} rate constant and refer to the overall reaction. However, a comparison can be justified considering that k_{cat} is similar in magnitude to k_1 and the MD study²⁷ showed that the mutations have a significant effect on only the rate constant of the first step. There is also a good agreement in the relative activation free energies of the Tyr14Phe and the Asp99Leu mutants between the two studies. The experimental results for all three mutants agrees well with the corresponding FMO results, though the FMO approach consistently predicts slightly larger effects by 0.7-1.5 kcal/mol (Table 2). The latter result is most likely an effect of the lack of structure relaxation in the FMO approach, as was discussed earlier. It is particularly interesting that the FMO approach predicts a relatively large effect of the Tyr55 deletion. It has previously been suggested based on structural data that the main effect of Tyr55 is to position Tyr14 for hydrogen bonding with the oxygen of the substrate. However, our results rather indicate that the catalytic effect of Tyr55 is of electronic character, as the positioning effect is not

accounted for by the FMO approach due to the lack of structure relaxation. To our knowledge, there has been no previous attempt to quantify the catalytic contribution of Tyr55 based on theoretical calculations.

Conclusions

In this study we have presented a new method for evaluating the effects of deleted residues on reaction rates and energies of enzymatic reactions using the FMO method. The great advantage of using FMO compared to regular MO theory is that it significantly reduce the computational effort. The effects of all possible single-deletions on a particular reaction rate can be estimated from two single point calculations on the wild-type protein, i.e. one calculation on the reactant structure and one on the transition state. In contrast, using regular MO theory the wild type calculations would have to be supplemented with two additional calculations for each single-deletion. Considering that single-point calculations on large systems, which consists of hundreds of atoms, are faster with the FMO method compared to regular MO method, the computational savings can become very large and they increase linearly with the number of deletions. In addition, the regular MO method cannot, without modifications, handle deletions that involve the breaking of covalent bonds.

In order to investigate the usefulness of this new approach we have performed MO and FMO calculations on the KSI-catalyzed proton abstraction of the cyclopent-2-enone substrate. Active site models of differing sizes have been used to analyze the catalytic effects of a few selected residues with conventional MO theory. In the FMO calculations we have used a large active site model to simultaneously investigate the effects of all possible single deletions. The data from both MO and FMO calculations indicate that not only Tyr14 and Asp99, which form hydrogen bonds with the substrate, but also Tyr55, which does not form a hydrogen bond with the substrate, are the key residues for stabilization of TS and INT in the proton abstraction. Moreover, the data from the FMO calculations clearly show that Tyr14, Asp99, and Tyr55, in this order, are important in the stabilization of TS and INT, and that the main catalytic effects come from these three residues among the ten residues which were

considered in this study. The relative activation energies for the different deletions are in good agreement with relative activation free energies for substitution mutants involving the same residues that were analyzed in an experimental study of the native reaction. It is particularly interesting that the FMO study reproduces the effect of mutating the Tyr55 residue considering that this residue is not directly interacting with the substrate, and since it has been suggested that the main effect of Tyr55 is to position Tyr14 for forming a hydrogen bond to the substrate. The latter is an effect that is not considered by our FMO approach. Thus, our calculations have provided physical insight into the catalytic function of Tyr55.

This study has demonstrated that our new approach, which is based on the FMO method, is useful for quantitatively evaluating the individual contributions of many residues to the enzymatic catalysis at the same time. This rational evaluation cannot be done by either the conventional MO method or site-directed mutagenesis. It could be argued that the kind of analysis we have presented is of limited value for rational design, since this study has provided little information on how to improve the catalytic efficiency of the enzyme. After all, the deletions that had significant effects in all cases raised the activation energy. However, it should be remembered that this is a reaction that is optimized by evolution and it is expected to be very difficult to improve the catalytic properties of the enzyme. The new method may be useful when attempting to change the reaction-specificity of an enzyme and for improving the catalytic efficiency of promiscuous reactions.

Associated Content

Supporting Information: Images of the model systems (Figure S1) and all geometry optimized structures of the model systems (Figure S2). This material is available free of charge via the Internet at <http://pubs.acs.org>.

Author Information

Corresponding Author: Phone: +46 8 790 8210, E-mail: tore@physchem.kth.se

Acknowledgments

This work was supported by the Carl Trygger Foundation and the Swedish Research Council (VR). The authors thank Dr. Takeshi Ishikawa of Nagasaki University for technical support with the PAICS program.

Figure 1. Proton abstraction catalyzed by KSI. The substrate, the general base (Asp38), and three surrounding residues (Tyr14, Asp99, and Tyr55) in the reactant state (R) and the intermediate state (INT) are displayed. The substrate changes from cyclopent-2-enone (in R) to cyclopenta-1,4-dienolate (in INT) by the proton (H in red) abstraction.

Figure 2. Optimized structures of R, TS, and INT obtained for Model5 at the M06-2X/6-31G(d) level by the MO method. Hydrogen bond distances (in Å) between the substrate O atom and the Tyr14 H atom, between the substrate O atom and the Asp99 H atom, and between the Tyr14 O atom and the Tyr55 H atom calculated for Model5 and Model6 (in parentheses) are shown. The transferring proton is marked with a blue circle.

Figure 3. Free energy diagram along the intrinsic reaction coordinate of Model5 and Model6 (in parentheses) in diethyl ether (red) and water (blue) obtained at the (SMD-PCM)-M06-2X/6-311++G(d,p) level by the MO method. Gas phase free energy corrections due to translation, rotation and vibration are not included.

Figure 4. Energy diagram along the intrinsic reaction coordinate of Model6 in the gas phase obtained at the SCS-MP2/6-31G(d,p) level by the FMO method and at the M06-2X/6-311++G(d,p) level by the MO method (in parentheses).

Figure 5. Interaction energies between the fragment of the substrate-Asp38 pair and the fragments of the surrounding residues calculated at the SCS-MP2/6-31G(d,p) level by the FMO method.

Figure 6. Energies of TS (red) and INT (blue) relative to energies of R (black) for the original form and the deletion forms calculated at the SCS-MP2/6-31G(d,p) level by the FMO method. No residue is deleted in the original form (the first data points), and the listed residues are deleted in the deletion forms. The relative energies of TS and INT of the original form are indicated with red and blue dotted lines, respectively.

Table 1. Energies of TS and INT Relative to Energies of R for the Model Systems

model	contents	TS $\Delta E^{\ddagger a}$	INT ΔE^a
Model1	substrate, Asp38	— ^b	— ^b
Model2	substrate, Asp38, Tyr14	10.8	8.6
Model3	substrate, Asp38, Asp99	10.5	5.3
Model4	substrate, Asp38, Tyr14, Asp99	7.4	0.6
Model5	substrate, Asp38, Tyr14, Asp99, Tyr55	5.9	-3.6
Model6	substrate, Asp38, 10 surrounding residues	7.2	-3.0

^aThe relative energies (in kcal/mol) were calculated at the M06-2X/6-311++G(d,p) level by the MO method. ^bThe relative energies of TS and INT for Model1 are not listed, because the stable TS and INT structures were not obtained for Model1.

Table 2. Theoretical Activation Energies ($\Delta E^\ddagger$), Heats of Reaction (ΔE), and Relative Activation Free Energies ($\Delta\Delta G_I^\ddagger$) for the Initial Proton Transfer Step, and Experimental Relative Activation Free Energies ($\Delta\Delta G_{cat}^\ddagger$) for the Overall Reaction in KSI

deleted residue	This work cyclopent-2-enone substrate				Previous studies native substrate		
	TS		INT		enzyme	TS	
	$\Delta E^\ddagger^a$	$\Delta\Delta E^\ddagger^b$	ΔE^a	$\Delta\Delta E^b$		$\Delta\Delta G_I^\ddagger^c$	$\Delta\Delta G_{cat}^\ddagger^d$
No deletion	8.0	0.0	-4.2	0.0	WT	0.0	0.0
Tyr14	13.1	5.1	14.5	18.7	Tyr14Phe	4.9	4.4
Tyr55	10.1	2.1	2.2	6.4	Tyr55Phe	—	1.1
Asp99	12.3	4.3	7.3	11.5	Asp99Leu	3.4	2.8

$^a\Delta E^\ddagger$ and ΔE (in kcal/mol) were calculated for the original (no deletion) form and the Tyr14, Tyr55, and Asp99 deletion forms with the cyclopent-2-enone substrate at the SCS-MP2/6-31G(d,p) level by the FMO method. $^b\Delta\Delta E^\ddagger$ and $\Delta\Delta E$ are relative to the $\Delta E^\ddagger$ and ΔE of the original form, respectively. $^c\Delta G_I^\ddagger$ (in kcal/mol) were calculated for WT KSI and its Tyr14Phe and Asp99Leu mutants with the native substrate (5-androstene-3,17-dione) by the transition state theory using the rate constants for the initial proton transfer step obtained from MD-EVB-simulations ($k_I = 1.7 \times 10^5$, 44, and 5.7×10^2 (s⁻¹), respectively, at 25 °C)²⁷. $\Delta\Delta G_I^\ddagger$ are relative to the $\Delta G_I^\ddagger$ of WT. $^d\Delta G_{cat}^\ddagger$ (in kcal/mol) were calculated for WT KSI and its Tyr14Phe, Tyr55Phe, and Asp99Leu mutants with the native substrate (5-androstene-3,17-dione) by the transition state theory using the experimental rate constants for the overall reaction ($k_{cat} = 21230$, 13.3, 3510, and 220 (s⁻¹), respectively, at 25 °C)¹⁷. $\Delta\Delta G_{cat}^\ddagger$ are relative to the $\Delta G_{cat}^\ddagger$ of WT.

Table of Contents (TOC) Image:

References

- (1) O'Brien, P. J.; Herschlag, D. *Chem. Biol.* **1999**, *6*, R91-R105.
- (2) Kazlauskas, R. J. *Curr. Opin. Chem. Biol.* **2005**, *9*, 195-201.
- (3) Gerlt, J. A.; Babbitt, P. C. *Curr. Opin. Chem. Biol.* **2009**, *13*, 10-18.
- (4) Hult, K.; Berglund, P. *Trends Biotechnol.* **2007**, *25*, 231-238.
- (5) Humble, M. S.; Berglund, P. *Eur. J. Org. Chem.* **2011**, *2011*, 3391-3401.
- (6) Baker, D. *Protein Sci.* **2010**, *19*, 1817-1819.
- (7) Pantazes, R. J.; Grisewood, M. J.; Maranas, C. D. *Curr. Opin. Struct. Biol.* **2011**, *21*, 467-472.
- (8) Saven, J. G. *Curr. Opin. Chem. Biol.* **2011**, *15*, 452-457.
- (9) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. *Science* **2010**, *329*, 309-313.
- (10) Linder, M.; Johansson, A. J.; Olsson, T. S.; Liebeschuetz, J.; Brinck, T. *J. Comput. Aided. Mol. Des.* **2012**, *26*, 1079-1095.
- (11) Linder, M.; Johansson, A. J.; Manta, B.; Olsson, P.; Brinck, T. *Chem. Commun.* **2012**, *48*, 5665-5667.
- (12) Kuliopulos, A.; Talalay, P.; Mildvan, A. S. *Biochemistry* **1990**, *29*, 10271-10280.
- (13) Li, Y. K.; Kuliopulos, A.; Mildvan, A. S.; Talalay, P. *Biochemistry* **1993**, *32*, 1816-1824.
- (14) Kim, S. W.; Cha, S. S.; Cho, H. S.; Kim, J. S.; Ha, N. C.; Cho, M. J.; Joo, S.; Kim, K. K.; Choi, K. Y.; Oh, B. H. *Biochemistry* **1997**, *36*, 14030-14036.
- (15) Wu, Z. R.; Ebrahimian, S.; Zawrotny, M. E.; Thornburg, L. D.; Perez-Alvarado, G. C.; Brothers, P.; Pollack, R. M.; Summers, M. F. *Science* **1997**, *276*, 415-418.
- (16) Cho, H. S.; Ha, N. C.; Choi, G.; Kim, H. J.; Lee, D.; Oh, K. S.; Kim, K. S.; Lee, W.; Choi, K. Y.; Oh, B. H. *J. Biol. Chem.* **1999**, *274*, 32863-32868.
- (17) Kim, D. H.; Jang, D. S.; Nam, G. H.; Choi, G.; Kim, J. S.; Ha, N. C.; Kim, M. S.; Oh, B. H.; Choi, K. Y. *Biochemistry* **2000**, *39*, 4581-4589.
- (18) Kraut, D. A.; Sigala, P. A.; Fenn, T. D.; Herschlag, D. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1960-1965.
- (19) Schwans, J. P.; Sunden, F.; Gonzalez, A.; Tsai, Y.; Herschlag, D. *J. Am. Chem. Soc.* **2011**, *133*, 20052-20055.
- (20) Hawkinson, D. C.; Eames, T. C.; Pollack, R. M. *Biochemistry* **1991**, *30*, 10849-10858.
- (21) Houck, W. J.; Pollack, R. M. *J. Am. Chem. Soc.* **2004**, *126*, 16416-16425.
- (22) Park, H.; Suh, J.; Lee, S. *Bull. Korean Chem. Soc.* **2002**, *23*, 837-844.
- (23) Feierberg, I.; Åqvist, J. *Biochemistry* **2002**, *41*, 15728-15735.
- (24) Park, H.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **2003**, *125*, 901-911.
- (25) Mazumder, D.; Kahn, K.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 7553-7561.
- (26) Chakravorty, D. K.; Soudackov, A. V.; Hammes-Schiffer, S. *Biochemistry* **2009**, *48*, 10608-10619.
- (27) Chakravorty, D. K.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2010**, *132*, 7549-7555.
- (28) Kamerlin, S. C.; Sharma, P. K.; Chu, Z. T.; Warshel, A. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4075-4080.
- (29) Kitaura, K.; Sawai, T.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *312*, 319-324.

- (30) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *313*, 701-706.
- (31) Nakano, T.; Kaminuma, T.; Sato, T.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. *Chem. Phys. Lett.* **2000**, *318*, 614-618.
- (32) Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. *Chem. Phys. Lett.* **2002**, *351*, 475-480.
- (33) Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2004**, *121*, 2483-2490.
- (34) Fedorov, D. G.; Ishimura, K.; Ishida, T.; Kitaura, K.; Pulay, P.; Nagase, S. *J. Comput. Chem.* **2007**, *28*, 1476-1484.
- (35) Watanabe, H.; Enomoto, T.; Tanaka, S. *Biochem. Biophys. Res. Commun.* **2007**, *361*, 367-372.
- (36) Nakamura, T.; Yamaguchi, A.; Kondo, H.; Watanabe, H.; Kurihara, T.; Esaki, N.; Hirono, S.; Tanaka, S. *J. Comput. Chem.* **2009**, *30*, 2625-2634.
- (37) Polyakov, I. V.; Grigorenko, B. L.; Moskovsky, A. A.; Pentkovski, V. M.; Nemukhin, A. V. *Chem. Phys. Lett.* **2013**, *556*, 251-255.
- (38) Ishikawa, T.; Burri, R. R.; Kamatari, Y. O.; Sakuraba, S.; Matubayasi, N.; Kitao, A.; Kuwata, K. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3646-3654.
- (39) Tanaka, S.; Mochizuki, Y.; Komeiji, Y.; Okiyama, Y.; Fukuzawa, K. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10310-10344.
- (40) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133-A1138.
- (41) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
- (42) Harvey, S. C.; Hoekstra, P. J. *Phys. Chem.* **1972**, *76*, 2987-2994.
- (43) Bone, S.; Pethig, R. *J. Mol. Biol.* **1982**, *157*, 571-575.
- (44) Bone, S.; Pethig, R. *J. Mol. Biol.* **1985**, *181*, 323-326.
- (45) Smith, P. E.; Brunne, R. M.; Mark, A. E.; Vangunsteren, W. F. *J. Phys. Chem.* **1993**, *97*, 2009-2014.
- (46) Simonson, T.; Perahia, D. *Proc. Natl. Acad. Sci. U S A* **1995**, *92*, 1082-1086.
- (47) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117-129.
- (48) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239-245.
- (49) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378-6396.
- (50) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford CT, 2009.
- (51) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095-9102.
- (52) Goumans, T. P. M.; Ehlers, A. W.; Lammertsma, K.; Wurthwein, E. U.; Grimme, S. *Chem. Eur. J.* **2004**, *10*, 6468-6475.
- (53) Ishikawa, T.; Ishikura, T.; Kuwata, K. *J. Comput. Chem.* **2009**, *30*, 2594-2601.
- (54) Eyring, H. *J. Chem. Phys.* **1935**, *3*, 107-115.
- (55) Eyring, H. *Chem. Rev.* **1935**, *17*, 65-77.