KEYNOTE

QM and QM–FE simulations on reactions of relevance to enzyme catalysis: trypsin, catechol *O*-methyltransferase, β-lactamase and pseudouridine synthase

Tai-Sung Lee, Irina Massova, Bernd Kuhn and Peter A. Kollman*

Department of Pharmaceutical Chemistry, University of California San Francisco, California 94143-0446, USA

Received (in Cambridge, UK) 27th September 1999, Accepted 18th November 1999

The application of the quantum mechanics–free energy hybrid technique (QM–FE) to calculate the free energy changes in two enzymatic reaction systems, trypsin and catechol *O*-methyltransferase (COMT) is reported. The results rationalize the observed rate enhancements by comparing the reactions in enzyme and in aqueous solution. Quantum mechanical studies of the model systems of β-lactamase and pseudouridine synthase systems are also presented. For β-lactamase, the effect of solvation on hydrolysis and methanolysis of β-lactams has been investigated. For pseudouridine synthase, the first steps of two different proposed mechanisms have been modeled.

Contents

Introduction

Quantum mechanical methods for determining the electronic structure of molecules are crucial for a reliable description of complex chemical processes that are inaccessible to conventional empirical models, such as electronic density polarization, charge transfer, bond formation, and bond breaking in chemical reactions. Quantum mechanical methods are, however, limited to fairly small systems because the computational expense scales as the cube of the number of atoms (or higher), making applications of these methods to large biomolecules unfeasible.

Instead of using quantum mechanical methods to model the whole enzyme system, these methods may be used to model only a small part of the enzyme, assuming that the contribution from the rest of enzyme is small and negligible. While this approach is useful to obtain a rough energy profile along the reaction coordinate and other electronic structure information, the contribution from the rest of enzyme is ignored and this approach could lead to wrong results if long-range interactions are important and missing.

This problem was addressed by Warshel and Levitt, who introduced the concept of hybrid quantum mechanics and molecular mechanics (QM/MM) methods.[1] This concept involves treating only a small part of the system quantum mechanically while the rest is treated by molecular mechanics.[2–4] The quantum mechanical treatment provides accurate electronic information for the interesting region and the molecular mechanical calculation represents the non-bonded interactions due to the environment. Various models have been reported subsequently, which differ in the particular quantum mechanical and molecular mechanical methods used, in the treatment of the interactions between quantum mechanical and molecular mechanical regions, and the way to define the QM/MM boundary. Popular hybrid methods include link atom methods,[5,6] modified orbital or atom methods,[3,7–9] energy subtraction methods,[10–12] and self-consistent methods.[13] Nevertheless, there are still no hybrid methods that can provide rigorous definitions of the interactions between quantum mechanical and molecular mechanical regions and the QM/MM boundary. Some problems are still very difficult to handle by current hybrid methods, *e.g.*, the electron transfer between QM and MM regions.

To apply these hybrid methods to free energy simulations of enzyme systems, the computational expense of the calculation for the region treated by quantum mechanics must be much reduced. Semiempirical quantum mechanical potentials have been used for this purpose but are limiting because of the lower accuracy of currently available semiempirical methods than the best *ab initio* approaches.[14] The empirical valence bond model, which fits valence parameters to quantum mechanics results, is computationally efficient and has been widely used.[3]

We recently developed a computational technique to study the energetics of enzyme catalyzed reactions.[15] It is based on a combination of high level quantum mechanical calculations of a model of the reaction center and classical free energy calculations for the interactions between the region treated by quantum mechanics and its environment as well as within the environment itself. This quantum mechanics–free energy hybrid technique (QM–FE) requires only few points of quantum calculations, thus allowing a high level of *ab initio* theory to be used. In this approach, the region treated by quantum mechanics is represented by a set of charges during classical free energy calculations to determine the free energy of interaction between the quantum and classical atoms.[15] In this way the free energy changes of an enzymatic reaction can be calculated at relatively low expense. In this approach, however, the geometries of the region modeled by quantum mechanics along the reaction coordinate are not easy to define. This approach also ignores the entropy contribution of the region modeled by quantum mechanics.

The QM–FE approach can be summarized as follows. First, we generate a model of the Michaelis complex by equilibration and energy minimization of an appropriate structure (deter-

DOI: 10.1039/a907793f

mined by X-ray or NMR techniques) of the enzyme–substrate complex. This system is divided into a molecular mechanical and a quantum mechanical part, which, at least, should include all atoms that affect the electron distribution of the reaction center and whose interactions cannot be described classically in an adequate manner. Subsequent optimization of the fragments treated by quantum mechanics for different points along the reaction path yields the energy profile for the reaction *in vacuo*. During the quantum mechanics optimizations, we retain the relative orientation of the model system from the enzyme by imposing appropriate coordinate constraints, thereby taking into account the enzymatic preorganization of the substrates. The free energy of interaction between the regions treated by quantum mechanics and molecular mechanics is obtained by first assigning force field parameters and atomic partial charges, using the restrained electrostatic potential fit (RESP) methodology,[16] to the QM model for each point along the reaction path and then performing molecular dynamical free energy calculations. The free energy between two structures along the reaction path, ΔG_{tot}, can be calculated as in eqn. (1), where

$$\Delta G_{tot} = \Delta E_{QM} + \Delta G_{int} \quad (1)$$

ΔE_{QM} denotes the difference in *ab initio* energy and ΔG_{int} is the difference of the free energy of interaction.

An important aspect of our strategy is to compare the enzymatic transformation with the analogous reaction in aqueous solution. Although the actual reaction mechanism in water might differ from the enzyme, it is important to compare the same mechanism in both environments in order to elucidate the key factors for enzymatic catalysis. In the process, we have to consider the "cratic" free energy, ΔG_{cratic}, in the solution reaction, *i.e.* the free energy required to bring the fragments into a reactive geometry.[17,18] If we use as reference state unassociated reactants, we need to add ΔG_{cratic} to eqn. (1) to consider the reaction energetics in solution. One can estimate this contribution, which reflects the change in solvation and entropy upon dimer formation, using different methods, including a polarizable continuum model,[19,20] potential of mean force, and normal mode analysis calculations as well as using the analytic approach of Hermans and Wang.[21] In the case of the enzyme catalyzed reaction, we assume that ΔG_{cratic} is included in the free energy of forming the enzyme–substrate complex, ΔG_{bind}, from the separated enzyme and substrate. This free energy can be approximated as $-RT \ln K_M$, where K_M is the Michaelis constant for the enzyme reaction. We assume that this free energy is a composite of the favorable non-covalent enzyme–substrate interactions and the unfavorable free energy to restrict the substrate into a catalytically favorable geometry.

There are two major differences in what one must do to apply the quantum mechanical–free energy hybrid technique to enzyme-catalyzed reactions rather than organic reactions in solution. The first, the enzymatic restriction of the pathway, must be considered on a case-by-case basis, since ΔE_{QM} in eqn. (1) must correspond to a pathway consistent with the enzyme substrate geometry and which leads to a productive reaction. Secondly, applying eqn. (1) to an organic reaction is straightforward because the ΔG_{int} simply involves the non-covalent free energy of interaction of the classical solvent and quantum mechanical reactants. To apply eqn. (1) to enzyme catalyzed reactions, one must solve the link atom problem, where classical atoms are bonded to quantum mechanical ones. We don't claim to have 'solved' the problem, but our RESP methodology enables us to 'splice' the charge distribution of the atoms in the region treated quantum mechanically to allow classical ΔG_{int} calculations in a way that appears general, robust, and internally consistent with the way we do free energy calculations involving non-covalent protein interactions.

In this review, we report the application of the quantum mechanical–free energy hybrid technique to calculate the free energy changes in two enzymatic reactions with very different charge redistribution during the reaction, *i.e.*, the amide hydrolysis in trypsin and the *O*-methylation of catechol in catechol *O*-methyltransferase. We also tried to rationalize the rate enhancement by comparing the reactions in enzyme and in aqueous solution. Quantum mechanical (non-hybrid) studies of the model systems of β-lactamase and pseudouridine synthase systems are also presented. For β-lactamase, the effect of solvation on reaction barriers for hydrolysis and methanolysis of β-lactams and the stability of their transition states and reaction intermediates have been investigated. For pseudouridine synthase, the first steps of two different proposed mechanisms have been modeled and the solvent effect has also been examined.

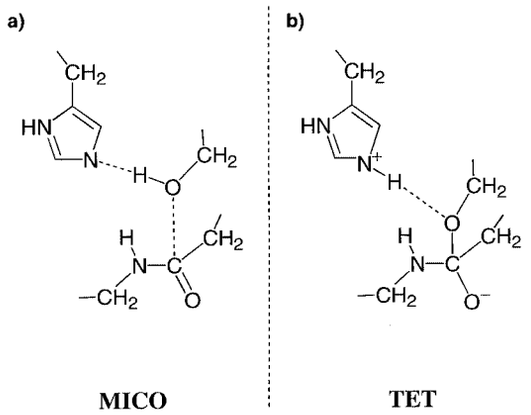


Fig. 1 Quantum mechanical model system for the amide hydrolysis in trypsin and in solution. (a) Michaelis complex (MICO), (b) tetrahedral intermediate for acylation (TET). In the quantum mechanical calculations, the CH_2– groups are replaced with CH_3.

Results and discussion

Quantum mechanics–free energy hybrid technique

Until now, we applied our quantum mechanical–free energy hybrid technique to study the mechanisms of two enzymatic reactions, namely the amide hydrolysis in trypsin[15] and the *O*-methylation of catechol in catechol *O*-methyltransferase (COMT).[22] It is particularly instructive to compare the catalytic mechanisms of these two enzymes because in trypsin charge is being created in the transition state while in catechol *O*-methyltransferase charge is being annihilated.

Trypsin. In the hydrolysis of amides by the serine protease trypsin, previous experimental[23] and theoretical[24,25] studies clearly suggest that the rate-limiting step is the formation of the tetrahedral intermediate for acylation. Furthermore, there is significant circumstantial evidence that the free energy of the tetrahedral intermediate is very similar to that of the transition state in its formation.[23,26] Thus, in the case of trypsin, we used the quantum mechanical–free energy hybrid technique to calculate the ΔG_{tot} difference between the Michaelis complex (MICO) of substrate bound to the enzyme and the tetrahedral intermediate (TET). The goal was to see if the calculated ΔG_{tot} was close to the experimental $\Delta G^{\ddagger}$ for amide hydrolysis of a specific substrate by trypsin. We also carried out a similar calculation in which our quantum mechanical fragments (methanol to represent Ser195, methylimidazole to represent His57, and *N*-methylacetamide to represent the amide bond of the substrate), which are represented in Fig. 1, were solvated in water rather than embedded in the active site in order to compare ΔG_{tot} for the enzyme catalyzed and non-catalyzed reaction.

Fig. 2 presents the calculated free energies for amide hydrolysis both in trypsin and in aqueous solution. Encouragingly, with the cratic free energy included for the solution reaction,

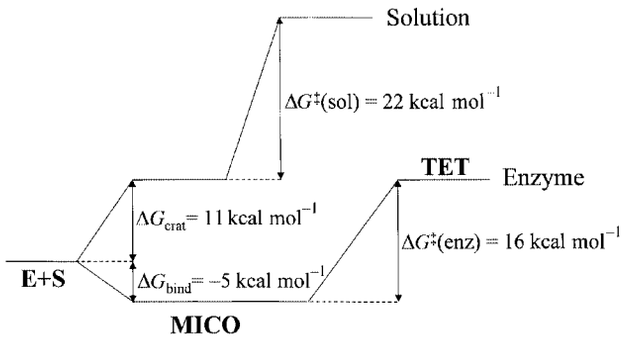
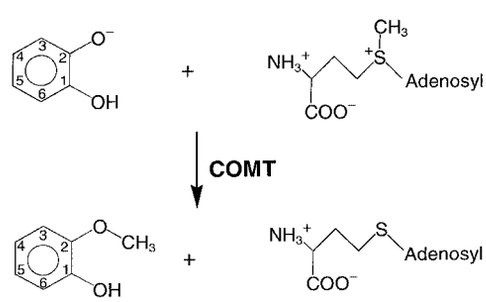


Fig. 2 Schematic free energy diagram for the amide hydrolysis reaction in trypsin and in aqueous solution. E + S denotes the separated enzyme and substrate, MICO is the Michaelis complex, and TET stands for the tetrahedral intermediate for acylation. The free energy for the transition state is assumed to be approximately that of the tetrahedral intermediate (see text).

both calculated barriers of activation, $\Delta G^{\ddagger}$ of 16 kcal mol^{-1} in the enzyme and 33 kcal mol^{-1} in solution are in excellent agreement with experimental values of 15 and 31 kcal mol^{-1}, respectively. Our results indicate that the major contribution to the enzyme catalysis in trypsin comes from the preorganization of the reacting groups and that smaller, but significant contributions are from enzymatic residues that stabilize the tetrahedral intermediate such as Asp102 and the oxyanion hole.

Catechol *O*-methyltransferase. Catechol *O*-methyltransferase is an enzyme which catalyzes the transfer of a CH_3^+ unit from the electrophilic donor *S*-adenosyl-L-methionine (SAM) to one of the two nucleophilic hydroxy groups of catechol. In our studies, we focused on the rate limiting step of this reaction, which is the direct S_N2 process depicted in Scheme 1,[27–30] in



Scheme 1

which two molecular ions of opposite charge, catecholate and SAM, react to form the neutral products. We used the X-ray structure of the enzyme complexed with the inhibitor 3,5-dinitrocatechol as starting structure for our simulations.[31] The X-ray structure shows a Mg^{2+} ion in the center of the active site which is octahedrally coordinated to three enzymatic residues, a water molecule, and the two oxygen atoms of catechol. The enzyme nicely orients the two reactants in a reactive geometry with the methyl group of SAM pointing towards one oxygen of catechol.

In our simulations we included the catecholate anion and a trimethylsulfonium cation, as a model for SAM, into the quantum mechanical part. The relative simplicity of the S_N2 methyltransfer reaction allowed us to define a reasonable approximation to the reaction path using linearly interpolated constraints, and hence to compute the reaction profile for the entire pathway. As can be seen from Fig. 3, the energy profile for the quantum mechanical model reaction reveals a relatively small transition state energy barrier of ~10 kcal mol^{-1} and a strongly negative reaction energy because of two opposite charges being annihilated *in vacuo*. Inclusion of the interactions with the protein or water environment in the calculation has

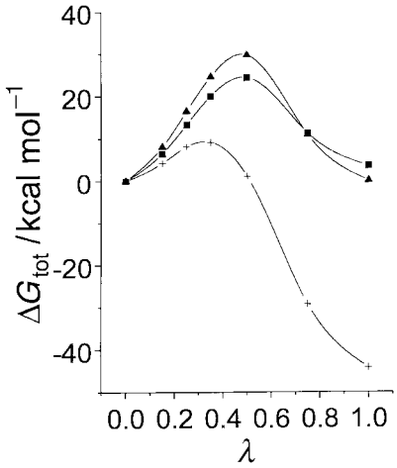


Fig. 3 Free energy profiles for the transmethylation reaction in vacuum (+), in catechol *O*-methyltransferase (squares), and in aqueous solution (triangles). The reaction path variable λ is defined as 0 for the reactant and 1 for the product state. The energy profile in vacuum was calculated at the HF/6-31+G*//MP2/AUG-cc-pVDZ level. The labeling of the ordinate axis should be replaced by ΔE_{QM} for the energy profile in vacuum.

a profound effect on the energy profile, leading to transition state energy barriers of 24.5 and 29.5 kcal mol^{-1} in the enzyme and in aqueous solution, respectively, and almost thermoneutral product formation in both systems. The increase in reaction free energy compared to the profile calculated for the model reaction by quantum mechanics reflects the fact that favorable interactions between the substrates and the surroundings vanish during the charge annihilation reaction. Since 95% of this comes from a change in electrostatic interactions; clearly, the enzyme has to bind the reactants effectively while not stabilizing them electrostatically too much. It does this by using a Mg^{2+} ion and several hydrophobic residues for binding of the substrates. As indicated by the 5 kcal mol^{-1} smaller transition state energy barrier in the enzyme compared to aqueous solution, the binding pattern in catechol *O*-methyltransferase leads to a less catalytically "unfavorable" stabilization of the ground state compared to water. Additionally, as revealed by our calculations, the enzyme uses the negatively charged sulfur atom of the residue Met40 to specifically stabilize the transferring methyl group in the transition state.

As was found in the study of trypsin, the major difference between the reaction in the enzyme and in aqueous solution lies in the cratic free energy contribution of orienting the substrates in a reactive geometry in solution. We calculated a ΔG_{cratic} of 9–13 kcal mol^{-1}, depending on the theoretical approach. Adding this to the difference between the transition state energy barriers for enzyme and water, we obtain an apparent difference in free energy of activation, $\Delta\Delta G^{\ddagger}$, of 14–18 kcal mol^{-1}. This corresponds to a rate enhancement of 10^{10}–10^{13} in the enzyme compared to the analogous reaction in solution, in good agreement with an experimental estimate of 10^{11}.

Quantum mechanical studies of reactions

β-Lactams, simple amides, and β-lactamases. We have recently completed an extensive quantum mechanical study of the reactivity of a wide range of β-lactams and simple amides.[32] We have investigated the effect of solvation on reaction barriers for hydrolysis and methanolysis of these compounds and the stability of their transition states and reaction intermediates. β-Lactams are widely used in medicine as potent antibacterials and in chemistry as synthetic intermediates. The activity of β-lactams is mediated by their ability to destroy the rigidity of the bacteria cell wall by inhibiting the cross-linking reactions of the major bacteria cell wall component, a peptidoglycan. This reaction is catalyzed by penicillin-binding proteins (PBPs)

which received their name for their high affinity for the typical β-lactams: penicillins and cephalosporins.

β-Lactamases are bacterial defensive enzymes which are evolutionarily related to penicillin-binding proteins. β-Lactamases can hydrolyze an amide bond of the β-lactam ring and convert β-lactam antibiotics into inactive compounds with no noticeable affinity to penicillin-binding proteins, thus, providing the resistance for a microorganism to β-lactam antibiotics. Most of the penicillin-binding proteins and three out of four classes of β-lactamases (A, C and D classes [33]) possess an active site serine residue which is acylated and deacylated by the substrates during enzymatic catalysis. Both catalytic-serine penicillin-binding proteins and β-lactamases have similar topology and are relatively easily acylated by β-lactams. However, their deacylation step is quite different: fast in β-lactamases and slow or absent in penicillin-binding proteins. Most of penicillin-binding proteins also show trans- and carboxypeptidase activity with the exception of high-molecular-weight class C penicillin-binding proteins,[34] whose known activity is a signal transduction for initiation of the synthesis of class A β-lactamases.[35] In contrast, β-lactamases are not peptidases.

The methanolysis/hydrolysis reactions of β-lactams are very similar to the enzymatic catalysis of serine-dependent β-lactamases and penicillin-binding proteins with methanol/hydroxy acting as a truncated serine residue with the difference that there is no analog to the Michaelis complex in the methanolysis reaction. Obviously, that intrinsic reactivity of the β-lactam could contribute to the activity of these compounds with the bacterial enzymes. Therefore, we have studied the contribution of the various structural elements to the activity of β-lactams. It has been suggested that the strain of the four-membered β-lactam ring and substituents at the carbon and nitrogen side can effect the activity as well as additional strain along the junction of the five- and six-membered rings to the β-lactam ring in penicillins and cephalosporins. We addressed these issues by reconstructing the pathways of the methanolysis/hydrolysis reactions of 18 β-lactams and simple amides in gas and solution. We used *ab initio* quantum mechanical methods at the HF/6-31+G* level combined with the state-of-the-art polarizable continuum model (PCM) and the UAHF approach, which utilizes a new definition of the solute cavity and significantly improves the calculated solvation free energies.[36] The PCM/UAHF method was recently incorporated into the GAUSSIAN98 software.[37] We studied the effects of solvation on the pathways of hydrolysis and methanolysis of the selected β-lactams and simple amides. The complexity of the β-lactams was gradually increased starting with the simple *N*-methyl-azetidin-2-one structure by adding various substituents at the carbon and nitrogen sides of the cleaved amide bond, fusing the additional ring and ultimately approaching the structure of penicillin G. The barriers of the reaction for the first step of the addition of the nucleophile to the carbonyl bond were predicted correctly. Experimental data for reported barriers were 20–30 kcal mol^{-1} for base catalyzed and neutral hydrolyses of simple amides [38,39] and 16.7 kcal mol^{-1} for hydrolysis of azetidin-2-one.[40] The calculations have predicted the barriers between 19 and 29 kcal mol^{-1} for hydroxide/methoxide-ion catalyzed hydrolyses, in excellent agreement with experiment. We have also studied the ground-state effect of the ring strain on the activation barriers. Experiments showed a 30–500-fold increase in rates of hydrolysis of β-lactams *vs*. corresponding acyclic amides that corresponds to a 2.0–3.7 kcal mol^{-1} decrease in the activation barrier.[41] We have found a 4.2 kcal mol^{-1} decrease in the barrier for hydrolysis and a 3.5 kcal mol^{-1} decrease in the barrier for methanolysis in water. It has been experimentally shown that an introduction of an electron-rich group at the nitrogen side increases the activation barriers and the opposite process, the esterification of the carboxylate group, results in a 130-fold increased rate which corresponds to a 2.9 kcal mol^{-1} decrease in an activation barrier.[42] Our calculations supported this finding showing the barrier increase of 69.8–72.3 kcal mol^{-1} in gas when the carboxylic group was introduced at the nitrogen side and of 1.5 to 5.5 kcal mol^{-1} in solution. The effect of the additional strain along the ring juncture when the second ring was introduced into β-lactams resulted in a 280–3900 fold increase in rates of hydrolysis which corresponded to lowering the barrier by 3.3–4.9 kcal mol^{-1} as a result of destabilization of the starting materials. Our calculations displayed the decrease of 8.7–13.8 kcal mol^{-1} in gas and 1.6–8.8 kcal mol^{-1} in solution, in accordance with the experiment. The inductive effect of the β-acylamino substituent at the carbonyl side of β-lactam was a little more difficult to reproduce, because the starting conformation for the substituent has been chosen as close as possible to the one in the acyl-enzyme intermediate of the complex of penicillin G and Glu166Asn mutant of TEM-1 β-lactamase from *Escherichia coli*,[43] and no additional conformational search has been performed. However, the consideration of all the structural conformers of a simpler structure of *N*-methyl-3β-hydroxyazetidin-2-one showed the decrease in the activation barrier for this compound, comparing to one unsubstituted at C3 β-lactam, thus supported the rationalization of the inconsistency with experiment in the study of the role of β-acylamino substituent effect on the reaction rate as due to limited conformational sampling.

The study of the directionality of the nucleophile approach in the hydrolysis/methanolysis reactions of β-lactams on example of *N*-methyl-3β-hydroxyazetidin-2-one has led to the broader conclusion that the predominant α-side approach of the base in nucleophilic reactions with penicillins in solution and under enzymatic catalysis is driven by the steric hindrance of the β-side and the fact that the amide nitrogen would form the hydrogen bond to the attacking nucleophile, thus stabilizing it and preventing such transition species from the collapsing to the tetrahedral intermediate.

We have also studied the relationship between various structural, physicochemical and kinetic parameters for the investigated β-lactams and amides and their intermediates and products. We have found an excellent correlation of 99% between the β-lactam C=O bond length and the observed second-order rate constants for hydroxide ion catalyzed hydrolysis of β-lactams. This finding relates the stability of the starting materials to the activation barrier height for the first step of the reaction, supporting the finding that the first step of the anionic hydrolysis is rate limiting for β-lactam reactions.[41] A 77% correlation has been found between the β-lactam C=O bond length and the calculated barrier heights for the first step of the methanolysis reaction. This finding supports the reliability of the continuum solvent PCM/UAHF model used in this work.[32] Therefore, the β-lactam C=O bond length could be used as criteria in assessment of the relative reactivity of the β-lactams and can be useful in prediction of the activity of novel β-lactam antibiotics.

Pseudouridine synthase. Pseudouridine (ψ) is the most common modified nucleotide present in 93 modified bases identified in various RNAs.[44] Pseudouridine synthases (ψS) catalyze the conversion of specific uridine residues in RNA into pseudouridine. However, not much is known about its mechanism of action. A commonly proposed mechanism is as follows: the C6 carbon is first attacked by a cysteine residue, which serves as a nucleophile, followed by the cleavage of the carbon–nitrogen glycosyl bond. The uracil ring then processes a 180° flip (or 120° rotation), followed by the formation of the C5–C1′ bond to form the final product.[45,46] This mechanism is consistent with most experimental evidence. Kammen showed that tRNA Pseudouridine Synthase I (ψSI) activity was inhibited by sulfhydryl reagents.[47] 5-FUra-RNA, which can form stable 5,6-dihydropyrimidine adducts with enzymes involving a methyltransferase mechanism, has been shown to be an inhibitor of pseudouridine synthase.[48] However, no covalent intermediates

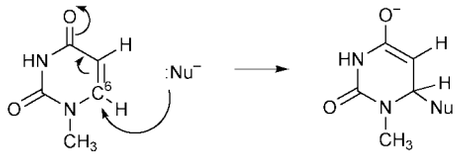


Fig. 4 The reaction and the model molecules used in attack at the C6 position in pseudouridine synthase reaction. "Nu" is the nucleophile which can be a Cys (modeled by CH_3S–) or an Asp (modeled by CH_3COO–).

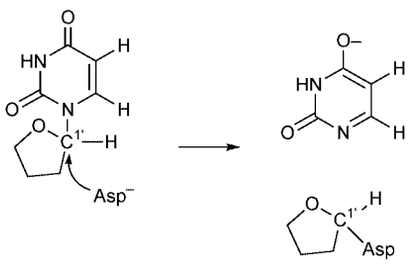


Fig. 5 The reaction and the model molecules used in attack at the C1′ position in pseudouridine synthase reaction.

have yet been detected as conclusive evidence. Recently, it also has been shown that a Cys residue is not conserved and, moreover, Cys is even not required for catalytic activity in the pseudouridine synthase reaction.[46,49] The last two pieces of evidence clearly argue against the above sulfhydryl mechanism.

Huang and co-workers have proposed an alternative mechanism in which a conserved Aspartate serves as the nucleophile and the nucleophilic attack occurs at either the C6 or the C1′ position as the first step in catalysis.[49] Although COO– is known as a weak nucleophile, it has been found that aspartate or glutamate may serve as the catalytic nucleophile in glycosidases.[50–54] In tRNA guanine transglycosylase, a covalent reaction intermediate has been isolated and supports the fact that an aspartate residue can serve as the catalytic nucleophile to attack the C1′ carbon.[55,56] This mechanism is very similar to the mechanism proposed by Huang.[49]

To examine the enzyme mechanism, we have performed *ab initio* quantum mechanical calculations on model systems at the MP2 level and have tried to analyze aspects of both the original mechanism of ψS and the new mechanism proposed by Huang. The calculations we performed were to determine an energy profile along the reaction coordinate. From those data, we examined the energy barrier for the nucleophilic attack and the relative stabilities of reaction intermediates for different mechanisms. We first modeled the nucleophilic attack by Asp and by Cys at the C6 position. We also calculated the nucleophilic attack by Asp at the C1′ position, as in the new proposed mechanism.

For the nucleophilic attack on the C6 position, we chose model molecules as shown in Fig. 4. The energy profile along the reaction coordinate is obtained by constraining the distance between the nucleophile and the C6 atom of the uracil ring and calculating the MP2/6-31+G* single point energy using the HF/6-31+G* optimized geometry. All geometric parameters were optimized except the constrained distance. For the nucleophilic attack on the C1′ position, the same calculation protocol is used and the model molecules are shown in Fig. 5.

All calculations were done in the gas phase. Because there is no charge cancellation in the reaction, it should be a reasonable first step to model the reaction in the gas phase. The energy profile along the reaction coordinate is shown in Fig. 6 for the C6 position attack by Cys. The results show the attack by Cys has a local energy minimum at $r = 1.9$ Å, which corresponds to the covalent reaction intermediate. The attack by Asp shown in Fig. 7, however, does not have any minima along the reaction coordinate. Thus, Asp is a much weaker nucleophile than Cys

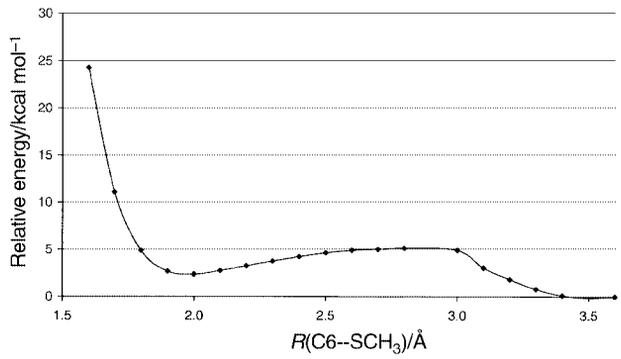


Fig. 6 The energy profile for the C6 position attack by Cys (modeled by CH_3S–) in pseudouridine synthase reaction. The *x*-axis is the distance between the nucleophile and the C6 atom (in Angstrom). The *y*-axis is the energy relative to $r = 3.6$ Å. The unit for energy is kcal mol^{-1}.

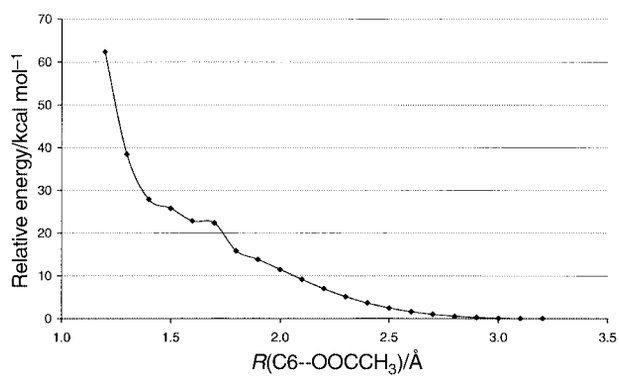


Fig. 7 The energy profile for the C6 position attack by Asp (modeled by CH_3COO–) in pseudouridine synthase reaction. The *x*-axis is the distance between the nucleophile and the C6 atom (in Angstrom). The *y*-axis is the energy relative to $r = 3.2$ Å. The unit for energy is kcal mol^{-1}.

and cannot form a stable intermediate during attack at the C6 position. Combined with the fact that Cys is not conserved and not required for the pseudouridine synthase activities, we can conclude that attack by an Asp of the enzyme on the C6 position of the uridine is unlikely.

We also tested Asp attack at the C6 position of 5-F Uracil. We found that the 5-F substitution slightly stabilized the adduct, leading to an energy increase relative to separate reactants of only ~20 kcal mol^{-1} at an O···C6 distance of 1.5 Å, compared to ~25 kcal mol^{-1} in Fig. 7 at this distance. However, the basic shape and energies in the potential surface were quite similar to Fig. 7 and very different from Fig. 6.

The C1′ position attack by Asp in the gas phase is shown in Fig. 8. The results show that the attack by Asp has a covalent intermediate at $r = 1.5$ Å and the reaction has an energy barrier about 31 kcal mol^{-1}. The energy of the covalent intermediate is higher than the energy of the separate reactants by about 5 kcal mol^{-1}.

A rough estimate of the solvent effect was obtained using the COSMO model,[36,57] a conductor-like continuum model, in the GAUSSIAN98 package.[37] We calculated solvation free energies at the MP2/6-31+G* level for the energy profiles of C1′ position attack by Asp using low ($\varepsilon = 4.9$) and high ($\varepsilon = 78.4$) relative permittivities. The results are shown in Fig. 8 and reflect the fact that the charge distribution is more delocalized in the reaction intermediate than in the reactants. As a result, the energy barrier and the relative energy of reaction intermediate increase when the solvation effect are considered. The barrier to C1′ attack on uracil is raised to ~35 kcal mol^{-1} and the reaction intermediate has an energy of ~12 kcal mol^{-1} higher than reactants in the low dielectric environment ($\varepsilon = 4.9$) and they become ~37 and ~17 kcal mol^{-1} in the high dielectric environment ($\varepsilon = 78.4$).

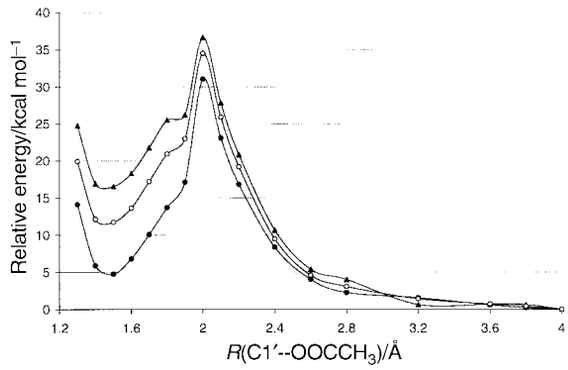


Fig. 8 The energy profile for the C1′ position attack by Asp (modeled by CH_3COO–) in vacuum and different dielectric environment using COSMO model: filled circle: vacuum; filled triangle: high relative permittivity ($\varepsilon = 78.4$); unfilled circle: low relative permittivity ($\varepsilon = 4.9$). The *x*-axis is the distance between the nucleophile and the C1′ atom (in Angstrom). The *y*-axis is the energy relative to $r = 4.0$ Å. The unit for energy is kcal mol^{-1}.

From our calculations, we found that for C6 position attack, Cys forms a much more stable adduct than Asp. Thus, Asp is an intrinsically unfavorable nucleophile for the attack at this position. Combined with the fact that the Cys residue is not conserved and is not required for catalytic activity, we suggest that the nucleophilic attack in pseudouridine synthase is not likely to occur at the C6 position. Hence all our results support the mechanism proposed by Huang[49] which suggested that the C1′ attack by Asp could be the first step of the mechanism.

Our calculations showed that the energy difference between the reaction intermediate and the reactants is ~5 kcal mol^{-1} and the energy barrier is ~31 kcal mol^{-1} for Asp attack at the C1′ position. The energy barrier is still high compared to typical enzymatic reactions. If the solvent effect is included, the energy barrier is suggested to be further raised. However, the influence from the protein environment has not been included because there is no X-ray structure available. In other cases, it has been found that the electrostatic interactions can certainly stabilize ionic reaction intermediates by amounts appropriate to reduce the energy barriers to the level of typical enzymatic reactions.[58–62]

Our results suggest that the first step of the catalytic reaction of pseudouridine synthase does not involve the C6 attack by a cysteine residue as previously thought. C1′ attack by Asp is found to be a possible pathway. There are certainly other plausible routes for this enzymatic reaction. For example, Schröder *et al*. have shown that in the NADH–glycohydrolase system, the O4′ atom of the sugar ring can be protonated by a nearby neutral aspartic acid or glutamic acid residue, followed by an S_N1 reaction occurring at the C1′ atom.[63] In this case, the aspartic acid or glutamic acid residues serve as a general acid rather than a nucleophile. A similar pathway could exist in the enzymatic reaction. Until the X-ray structure is available for this enzyme, one cannot carry out the detailed theoretical studies required to evaluate how Asp is functioning as a nucleophile or, in its protonated form, donating a proton to the sugar oxygen to enable glycosyl bond cleavage. Also, the availability of the X-ray structure is necessary to model the subsequent steps of this enzymatic reaction, including ring rotation and C–C bond formation.

Conclusions

With the advances in computers and software it became possible to use a combination of quantum mechanics and molecular mechanics methods to study enzymatic reactions. We applied different techniques to study the mechanisms of four different enzymes: trypsin, catechol *O*-methyltransferase, β-lactamase and pseudouridine synthase.

The main challenge which still remains in this area is that only part of the ligand (substrate or inhibitor) and the active site are treated with the high level of quantum mechanical theory, while the rest of the system has to be treated on the molecular mechanical level due to its size and complexity. This causes difficulty related to the accurate description of the interactions between regions considered quantum and molecular mechanically, the so-called link-atom problem. One of the solutions to this problem has been demonstrated in this paper. Another uncertainty is often associated with the evaluation of the "cratic" energy and the continuum solvent models used for such estimates. Nevertheless, quantum mechanics/molecular mechanics and quantum mechanics–free energy hybrid methods seem to be extremely promising quantitative approaches to study enzymatic reactions and have been recently applied to several systems.[3,4,15,22,23]

The approach taken for studying enzymatic reactions depends on the questions being asked. There are typically two types of questions asked: First, what is the cause of the rate enhancement of enzymatic reactions? Our approach has been the first to answer this question *ab initio* with no explicit parameterization of the energetics of the solution reaction and the critical role of the cratic term has been noted.[15,22]

The second question is: what is the detailed mechanism of the enzymatic reaction? In the case of trypsin and catechol *O*-methyltransferase, this was clear but in the case of many enzymes, *e.g*., β-lactamase and pseudouridine synthase, where the detailed mechanism has not been unequivocally established, QM–FE can be carried out on alternative pathways and, provided the estimated free energy is sufficiently accurate, some proposed mechanisms can at least be ruled out.

In the case of radical reaction applied to answer the second question, in contrast to the ionic reactions considered here,[64–66] a pure *ab initio* approach without considering solvent effects may be adequate, since the second term on the right hand side of eqn. (1) will not change much for radical reactions. The key necessity is simply realistic geometries of the reaction groups. Just like in the cases of trypsin and catechol *O*-methyltransferase, ΔG_{cratic} will be small because the three dimensional structure of the enzyme has "pre-organized" the reacting groups.

The knowledge disclosed in this paper sheds light on enzyme mechanisms in general and can be useful in the design of new potent pharmaceuticals targeting various biological systems.

Acknowledgements

The authors would like to thank the National Institutes of Health for research support (GM-29072 to PAK, AI-09998 to I. M., and GM-19410 to T. L.). B. K. gratefully acknowledges support from the German Academic Exchange Service (DAAD) through a research scholarship.

References

1 A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227.
2 E. Clementi, in *Computational Aspects for Large Chemical Systems*, Springer, New York, 1980.
3 A. Warshel, in *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, Wiley, New York, 1991.
4 J. Gao, in *Reviews in Computational Chemistry*, K. B. Lipkowitz and D. B. Boyd, eds., vol. 7, 119, VCH, New York, 1992.
5 U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, 1986, **7**, 718.
6 M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.*, 1990, **11**, 700.
7 V. Théry, D. Rinaldi, J.-L. Rivail, B. Maigret and G. G. Ferency, *J. Comput. Chem.*, 1994, **15**, 269.
8 J. Gao, P. Amara, C. Alhambra and J. J. Field, *J. Phys. Chem. A*, 1998, **102**, 4714.
9 Y. K. Zhang, T.-S. Lee and W. Yang, *J. Chem. Phys.*, 1999, **110**, 46.
10 F. Maseras and K. Morokuma, *J. Comput. Chem.*, 1995, **16**, 1170.

11 S. Humbel, S. Sieber and K. Morokuma, *J. Chem. Phys.*, 1996, **105**, 1959.
12 M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma, *J. Phys. Chem.*, 1996, **100**, 19357.
13 I. B. Bersuker, M. K. Leong, J. E. Boggs and R. S. Pearlman, *Int. J. Quantum Chem.*, 1997, **63**, 1051.
14 D. S. Hartsough and K. M. Merz, *J. Phys. Chem.*, 1995, **99**, 11266.
15 R. V. Stanton, M. Peräkylä, D. Bakowies and P. A. Kollman, *J. Am. Chem. Soc.*, 1998, **120**, 3448.
16 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269.
17 J. Janin, *Proteins: Struct., Funct., Genet.*, 1996, **24**, 1.
18 J. Hermans and L. Wang, *J. Am. Chem. Soc.*, 1997, **119**, 2707.
19 S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.
20 S. Miertus and J. Tomasi, *Chem. Phys.*, 1982, **65**, 239.
21 J. Hermans and L. Wang, *J. Am. Chem. Soc.*, 1997, **119**, 2707.
22 B. Kuhn and P. A. Kollman, *J. Am. Chem. Soc.*, submitted.
23 A. Warshel and S. J. Russell, *J. Am. Chem. Soc.*, 1986, **108**, 6569.
24 J. J. P. Stewart, *J. Comp. Chem.*, 1989, **10**, 209.
25 J. J. P. Stewart, *J. Comp. Chem.*, 1991, **12**, 320.
26 V. Daggett, S. Schröder and P. A. Kollman, *J. Am. Chem. Soc.*, 1991, **113**, 8926.
27 M. F. Hegazi, R. T. Borchardt and R. L. Schowen, *J. Am. Chem. Soc.*, 1979, **101**, 4359.
28 J. O. Knipe, P. J. Vasquez and J. K. Coward, *J. Am. Chem. Soc.*, 1982, **104**, 3202.
29 R. W. Woodard, M. D. Tsai, H. G. Floss, P. A. Crooks and J. K. Coward, *J. Biol. Chem.*, 1980, **255**, 9124.
30 J. Rodgers, D. A. Femec and R. L. Schowen, *J. Am. Chem. Soc.*, 1982, **104**, 3263.
31 J. Vidgren, L. A. Svensson and A. Liljas, *Nature*, 1994, **368**, 354.
32 I. Massova and P. A. Kollman, *J. Phys. Chem.*, 1999, **103**, 8628.
33 K. Bush and G. Jacoby, *Antimicrob. Agents Chemother.*, 1995, **39**, 103.
34 J. M. Ghuysen, *Annu. Rev. Microbiol.*, 1991, **45**, 37.
35 K. Hardt, B. Joris, S. Lepage, R. Brasseur, J. O. Lampen, J. M. Frère, A. L. Fink and J. M. Ghuysen, *Mol. Microbiol.*, 1997, **23**, 935.
36 V. Barone, M. Cossi and J. Tomasi, *J. Chem. Phys.*, 1997, **107**, 3210.
37 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, *Revision A.6*, Gaussian 98, Gaussian, Inc., Pittsburgh, PA, 1998.
38 J. P. Guthrie, *J. Am. Chem. Soc.*, 1974, **96**, 3608.
39 H. Slebocka-Tilk, A. J. Bennet, H. J. Hogg and R. S. Brown, *J. Am. Chem. Soc.*, 1991, **113**, 1288.
40 K. Bowden and K. Bromley, *J. Chem. Soc.*, *Perkin Trans. 2*, 1990, **12**, 2103.
41 M. I. Page, *The Chemistry of β-Lactams*, Blackie Academic & Professional, London, 1992, p. 351.
42 A. P. Laws and M. I. Page, *J. Chem. Soc.*, *Perkin Trans. 2*, 1989, **10**, 1577.
43 N. C. J. Strynadka, H. Adachi, S. E. Jensen, K. Johns, A. Sielecki, C. Betzel, K. Sutoh and M. N. G. James, *Nature*, 1992, **359**, 700.
44 P. A. Limbach, P. F. Crain and J. A. McCloskey, *Nucleic Acids Res.*, 1994, **22**, 2183.
45 D. V. Santi, Y. Wataya and A. Matsuda, *Proceedings of the International Symposium on Substrate-Induced Irreversible Inhibition of Enzymes*, New York, Elsevier-North Holland, 1978, 291.
46 X. Zhao and D. A. Horne, *J. Biol. Chem.*, 1997, **17**, 1950.
47 H. O. Kammen, C. C. Marvel, L. Hardy and E. E. Penhoet, *J. Biol. Chem.*, 1988, **263**, 2255.
48 T. Samuelsson, *Nucleic Acids Res.*, 1991, **19**, 6139.
49 L. Huang, M. Pookanjanatavip, X. Gu and D. V. Santi, *Biochemistry*, 1998, **37**, 344.
50 J. D. McCarter and S. G. Withers, *Curr. Opin. Struct. Biol.*, 1994, **4**, 885.
51 A. White, D. Tull, K. Johns, S. G. Withers and D. R. Rose, *Nature Struct. Biol.*, 1996, **3**, 149.
52 J. D. McCarter and S. G. Withers, *J. Biol. Chem.*, 1996, **271**, 6889.
53 J. D. McCarter, D. L. Burgoyne, S. C. Miao, S. Q. Zhang, J. W. Callahan and S. G. Withers, *J. Biol. Chem.*, 1997, **272**, 396.
54 C. Braun, T. Lindhorst, N. B. Madsen and S. G. Withers, *Biochemistry*, 1996, **35**, 5458.
55 C. Romier, K. Reuter, D. Suck and R. Ficner, *Biochemistry*, 1996, **35**, 15734.
56 C. Romier, K. Reuter, D. Suck and R. Ficner, *EMBO J.*, 1996, **15**, 2850.
57 V. Barone, M. Cossi and J. Tomasi, *J. Comput. Chem.*, 1998, **19**, 404.
58 A. Warshel, J. Åqvist and S. Creighton, *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 5820.
59 J. P. Guthrie and R. Kluger, *J. Am. Chem. Soc.*, 1993, **115**, 11569.
60 G. Alagona, C. Ghio and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 9855.
61 A. Warshel, A. Papazyan and P. A. Kollman, *Science*, 1995, **269**, 102.
62 A. Warshel, *J. Biol. Chem.*, 1998, **273**, 27035.
63 S. Schröder, N. Buckley, N. Oppenheimer and P. Kollman, *J. Am. Chem. Soc.*, 1992, **114**, 8232.
64 F. Himo and L. A. Eriksson, *J. Am. Chem. Soc.*, 1998, **120**, 11449.
65 P. E. M. Siegbahn and R. H. Crabtree, *J. Am. Chem. Soc.*, 1999, **121**, 117.
66 D. M. Smith, B. T. Golding and L. Radom, *J. Am. Chem. Soc.*, 1999, **121**, 9388.

Paper a907793f