On the Importance of Preserving Active Induced-Fit Conformations during Enzyme-Catalyzed Reactions: a Case Study of CARNMT1

Hao Deng[#,1,‡], Adua Rahman[#,2,] Ping Qian[*,1],[3] and Hong Guo[*,2]

[1]Chemistry and Material Science Faculty, Shandong Agricultural University, Taian 271018, P. R. China

[2]Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, United States

[3]Key Laboratory of Agricultural Film Application of Ministry of Agriculture and Rural Affairs, Taian 271018, P. R. China

[#]These authors make equal contributions

[‡]Current Address: School of Chemistry and Chemical Engineering, Huazhong University of Science and Technology, Wuhan 430074, P. R. China

Corresponding Authors:
*Ping Qian, **Email:** qianp@sdau.edu.cn
*Hong Guo, **Email:** hguo1@utk.edu

Keywords: Molecular Recognition; Induced Fit Conformation; Enzyme Catalysis; Catalytic Efficiency; Effects of Mutations; CARNMT1; QM/MM simulations.

Abstract

Considerable efforts have been made to understand how precise orientations of catalytic groups/substrate in enzyme-substrate complexes are generated for catalysis, leading to the proposals of lock-and-key, induced-fit and conformational selection models as well as their extensions. By contrast, little attention has been paid to the importance of preserving such active conformations during enzyme-catalyzed reactions to allow catalytic residues to exert their effects during catalysis, including transition state stabilization. Here MD and free energy simulations based on QM/MM potentials are carried out for CARNMT1, a histidine methyltransferase, to determine the importance and requirements for preservation of catalytic competent induced-fit conformations during methylation. We examine interplay of certain anchor residues that occupy similar positions before and after substrate binding (Asp316, Tyr386 and Tyr396) with the latch residue (His347) that is responsible for generating the active induced-fit conformation. For wild type, the catalytic competent orientations can be preserved well after the reaction passes TS allowing the active site residues to exert the effects. For the Tyr386Ala and Tyr396Asp mutants, the induced-fit conformation can be formed in the mutant-substrate complex but collapses in the early stage of reaction well before it can reach TS. Such pre-mature collapse of the active conformations as a result of the mutations is likely to be responsible for the significant reduction of catalytic efficiency. His347Phe also fails to preserve the active induced-fit conformation, and the active induced-fit conformation cannot be formed in the Asp316Ala-substrate complex. The results demonstrate the importance of preserving the active induced-fit conformation during enzyme-catalyzed reaction. They also show that generation of active induced-fit conformations of enzyme-substrate complexes may be a necessary but not sufficient condition for efficient catalysis. Interestingly, the catalytic efficiency seems to be correlated with the ability of the enzyme/mutant to preserve the active induced-fit conformation during catalysis.

Introduction

Molecular recognition is of fundamental importance in biology. Understanding its principles is crucial for designing drugs and studying diverse biological processes, including, but not limited to, enzyme catalysis, antibody–antigen recognition, protein-protein interactions, and transcriptional regulation. Therefore, this topic has attracted considerable attention, especially for protein-ligand interactions.[1-11] The early developments of Emil Fisher's lock-and-key hypothesis[12] and Koshland's induced-fit model[13] for molecular recognition were driven largely by the efforts to explain enzyme's activity and specificity. These textbook models were later extended to include original and extended conformational selection models.[3, 14, 15] The lock-and-key (LK) model hypothesizes that the way enzymes interact with substrates/non-substrates conceptually resembles that locks match keys (Figure 1a). Complimentary structural fit between substrate and enzyme is important in this model, and conformational changes of either enzyme or substrate are not considered during binding processes. The induced fit (IF) model, on the other hand, recognizes that the binding of a substrate may induce a change in the three-dimensional relationship of the active site residues, and precise orientation of catalytic groups required for enzyme's action can be generated through such induced fit process (Figure 1b, the upper pathway).[6, 13] It should be pointed out that catalytic competent conformations of enzyme-substrate complexes may also be generated through conformational changes of substrates induced by enzymes.[16, 17] Original conformational selection model (CS) recognizes that proteins are under dynamical fluctuations and that ligand may select and bind to the one that is compatible and shift the conformational ensemble towards this state (Figure 1b, the lower pathway).[14, 15] Extended conformational selection model allows conformational selection to be followed by conformational adjustment.[3]

Although the efforts to understand enzyme function and specificity have played a crucial role in shaping our current view for molecular recognition, for enzyme catalysis formation of enzyme-substrate complex is only the initial step and generation of catalytic competent orientations of residues relative to the substrate in enzyme-substrate complex through the lock-and-key (LK), induced fit (IF) and/or conformational selection (CS) processes may not be sufficient to lead to efficient catalysis. For an enzyme, the recognition may not only be essential in the substrate binding, but also for the reaction

that the enzyme catalyzes. The catalytic competent orientations of enzyme's groups relative to substrate may therefore need to be preserved, at least, for a part of the reaction process to allow catalytic residues to exert their effects, including transition state stabilization. Pre-mature collapse of such conformations (e.g., due to certain mutations) can lead to a significant reduction of catalytic efficiency. Moreover, it is known that, in addition to the residues that undergo induced-fit conformational changes upon substrate binding (e.g., latch residues), there may be some other active-site residues that occupy similar positions before and after substrate binding (e.g., anchor residues). Previous studies have shown that protein-protein interactions may go through a local lock-and-key mechanism from anchor residues followed by a induced-fit process involving latch residues,[18] and such step by step process may also exist for protein-ligand interactions.[3] However, it is not clear how the anchor and latch residues along with substrates would work together to preserve the active induced-fit conformations during enzyme catalyzed reactions and lead to efficient catalysis.

In this paper, QM/MM MD and potential of mean force (PMF) simulations were performed for the wild-type CARNMT1 and certain mutants (Asp316Ala, His347Phe, Tyr386Ala and Tyr396Asp) to determine the importance of preserving the active induced-fit conformations for efficient catalysis and understand the interplay of anchor and latch residues as well as the substrate in such process. CARNMT1 is a histidine $N_{\delta 1}$ position-specific methyltransferase and catalyzes the conversion of carnosine to anserine. Carnosine exerts numerous biological and physiological effects in multiple organs.[19] CARNMT1 is an ideal system for studying the role of the active site residues of enzymes in preserving the catalytic competent induced-fit conformations during catalysis. This is because the only active site residue that undergoes large induced-fit conformational changes upon substrate binding and product formation is His347. This makes it easier to study the stability of the active induced fit conformation during the enzyme catalyzed reaction with His347 and some other active-site residues being altered individually. Moreover, the induced-fit conformational change of His347 mainly involves the rotation around the C_α-C_β bond of His347 with the corresponding χ(C-C_α-C_β-C_γ) dihedral angle changing from about 55° in the apo enzyme to 157° in the enzyme-substrate complex (Figure 2); the r(Carnosine-$C_{\varepsilon 1}$⋯H347-$C_{\varepsilon 1}$) distance decreases from about 7.5Å to 3.5Å accordingly. This makes monitoring the formation and collapse of the

active induced-fit conformation in wild type and different mutants much easier. The kinetic data for several mutants of CARNMT1 have been determined along with the structure of the enzyme-product complex (5YF2).[20] The structures for the apo enzyme and enzyme-product complex are also highly similar, and the positions of His347 in 5YF0 and 5YF2 can be superposed well with similar $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ values (55° in 5YF0 and 51° in 5YF2). Thus, the substrate can lead to the induced-fit conformational change involving His347, while the product (a non-substrate) cannot. It has been proposed that such “hold-release” switching role of His347 can benefit substrate binding and product release.[20]

The results from the QM/MM MD and PMF simulations of this work showed that the wild-type enzyme could preserve the catalytic competent induced-fit conformation even after the reaction well passed the transition state (TS). Thus, the reasons that the wild-type enzyme has a high catalytic efficiency are not only because the formation of catalytic competent orientations of residues and substrate in the enzyme-substrate complexes but also due to its ability to preserve the catalytic competent conformations during the reaction process. For Asp316Ala, the active induced fit conformation could not be formed in the Asp316Ala-substrate complex due to missing certain key interactions involving Asp316. Although the active induced-fit conformation could be formed in Tyr386Ala-substrate, Tyr396Asp-substrate and His347Phe-substrate complexes, the active conformation collapsed in each of the cases in the early stage of reaction before it could reach TS, leading to an increase of the free energy barrier at the transition state and a lower catalytic efficiency. The results suggest that generating precise catalytic competent orientations of catalytic groups relative to substrates in enzyme-substrate complexes through the well-known models may be a necessary but not sufficient condition for efficient catalysis. It has also been shown that anchor residues may play a significant role in preserving catalytic competent induced-fit conformation for efficient catalysis. The principle discussed here may be useful in explanation of observed losses of activities due to certain mutations for some other enzymes and in the process of protein design.

Results and Discussion

Ability of Wild Type to Preserve Active Induced-Fit Conformation for Efficient Catalysis

The crystal structures of the apo enzyme (5YF0) and enzyme-substrate complex (5YF1, with SAM replaced by SFG) are superposed in Figure 2. It can be seen from Figure 2 that the carnosine histidine is deeply embedded into the active site pocket surrounded by some amino acid residues such as Phe313, Asp316, Leu345, His347, Tyr386, Tyr396, and Tyr398 with the substrate conformation stabilized by the corresponding interactions involving these residues. The structures of the apo enzyme (5YF0) and enzyme-substrate complex (5YF1) are highly similar with RMSD (based on C_α) of only 0.3Å. The RMSD value based on all atoms superposed for the active site residues without His347 is also quite small (0.3Å); the RMSD for His347 is 3.2Å. The site directed mutagenesis study[20] has shown that some of the active site residues are of important for efficient catalysis by the enzyme, including Asp316, Tyr386, Tyr396 as well as His347.

The average active site structure of the CARNMT1-SAM-Carnosine complex obtained from the QM/MM MD simulations is shown in Figure 3a. The interactions between carnosine and the key residues in the active site region (e.g., Asp316, His347, Tyr386, Tyr396, etc.) are likely to play a crucial role in the substrate binding and catalysis. The interactions obtained from the simulations are quite similar to those observed experimentally (Figure 2).[20] For instance, the two oxygen atoms of the Asp316 side chain interact with the $C_{\varepsilon 1}$ and $N_{\varepsilon 2}$ atoms of the substrate imidazole ring, respectively, in both crystal structure and reactant state from simulations; the average distance to $C_{\varepsilon 1}/N_{\varepsilon 2}$ is 3.43 Å/2.78 Å obtained from the simulations (Figure 3a), while it is 3.16 Å/2.94 Å in the crystal structure (Figure 2). Other interactions that are similar between the crystal structure and structure obtained from the simulations include the “π-π” stacking interaction between the Phe313 benzene ring and carnosine imidazole ring (e.g., ~3.95 Å and ~3.80 Å in Figure 3a and 2, respectively), the "π-π" stacking interaction between the His347 and carnosine imidazole rings (e.g., 3.67 Å in the simulated structure and 3.50 Å in the crystal structure), the "methyl-π" interaction between carnosine and the Tyr386 phenol ring (e.g., 3.73 Å in the simulated structure and 3.64 Å in the crystal structure), and the hydrogen bond between the hydroxyl group on Tyr398 and the substrate (3.13 Å in the simulated structure and 3.04 Å in the crystal structure). The positions of the residues mentioned above are fixed by the interactions with the substrate carnosine and their neighboring residues. For instance, the interaction

between the backbone carbonyl oxygen atom of Leu345 and $N_{\delta1}$−H of His347 stabilizes the His347 position, and His347 then forms the "π-π" interaction with the imidazole ring of the substrate. We have also observed some interactions involving the transferable methyl group in our simulated structure as a result of replacing SFG by SAM. For example, the hydroxyl group of Tyr386 and the backbone carbonyl oxygen of Phe313 form carbon hydrogen bonds with the transferable methyl group of SAM. In general, the average structure obtained from the QM/MM MD simulated and crystal structures are highly similar; the existence of some small differences is expected as SFG in the crystal structure was replaced by SAM for the simulations. The *r*-*θ* distribution map of the reactant state of wild type is also given in Figure 3a (see Methods below for the definitions). There are many structures with relatively short $r(C_M-N_{\delta1})$ distances and small values of the *θ* angle. The average distance between the target $N_{\delta1}$ atom and C_M atom of the transferable methyl group is approximately 3.0 Å, and the *θ* angle is mainly in the range of 0-30°. Therefore, when carnosine binds to the active site of the enzyme, it can form the near-attack conformations[21] to facilitate methyl transfer.

Figure 3b shows the fluctuations of the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ dihedral angle of His347 collected from each of the PMF windows along the reaction coordinate from the free energy simulations (see Methods below). It can be seen that the wild-type enzyme can preserve the catalytic competent induced-fit conformation with His347 in the "hold" position[20] [i.e., $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ ~ 160°] well after the reaction passes TS [at R = $r(C_M-S_\delta) - r(C_M-N_{\delta1})$ ~ 0.2Å]. The induced-fit conformation collapses at R = $r(C_M-S_\delta) - r(C_M-N_{\delta1})$ ~ 1.3 Å with $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ of His347 changing back to the value of the "released" position in the apo enzyme and the product complex (~ 60°), consistent with experimental structure of the product complex.[20] The average structure of the active site near the transition state of methylation obtained from the QM/MM PMF simulations is given in Figure 3c. Comparison of the structures in Figure 3a and 3c shows that there are some small changes between the reactant state and transition state for the wild-type enzyme. Nevertheless, the "π-π" stacking interaction between the His347 side chain and the imidazole ring of the substrate remains intact at the transition state that helps to generate the well-positioned interactions at the transition state for TS stabilization and protect the imidazole ring of the substrate from solvent. To examine whether the apo enzyme can reach the induced-fit conformation as proposed in the CS model,[14, 15] we

performed the QM/MM MD simulations based on 5YF0. The results in Figure 3d (Left) shows that the induced-fit conformation could be formed as monitored by $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ (red circles), even though the frequency of formation of such structure is quite low. We have also examined what would happen if the substrate in the induced-fit structure of the enzyme-substrate complex was removed from the active site. Figure 3d (Right) shows that if the substrate was removed, His347 would move quickly from the "hold" to the "released" position (as observed in the apo enzyme).

Inability to Form or Early Collapse of Active Induced-Fit Conformation in Some Mutants

The structural results given in Figure 3a and 3c and previous experimental data[20] support the suggestion that Asp316 stabilizes the conformation of the $N_{\varepsilon 2}$-H tautomeric form of the substrate and makes it possible for the transferable methyl group of SAM to point towards the electron lone pair of $N_{\delta 1}$ of carnosine prior to the reaction. The Asp316 residue in the CARNMT1 seems to play a similar and presumably more important role as the Asn255 residue in SETD3[22]. If the hydrogen bonding interactions between Asp316 and carnosine are removed (e.g., through mutations), the binding of the correct tautomer and the formation of the near-attack reactive state may be compromised, leading to a loss of the activity.[20] The average active-site structure of the Asp316Ala-substrate complex obtained from the simulations is given in Figure 4a along with the distribution map. Although other active site residues (e.g., Phe313, His347, Leu345 and Tyr386) may also be important for anchoring the $N_{\varepsilon 2}$-H tautomeric form and for the formation of the near-attack conformation, they cannot achieve this by themselves without Asp316. Indeed, the distribution map in Figure 4a shows that the population of the structures with relatively short $r(C_M\text{−}N_{\delta 1})$ distances (about 3.0 Å) and small values of the θ angle (less than 30°) become much smaller in the Asp316Ala mutant compared to the case of wild type (Figure 3a). Moreover, comparison of the structures of Figure 3a and Figure 4a shows that the imidazole ring of the substrate underwent a rotation of ~70° in Asp316Ala, and this conformational change of the substrate also affects the correct alignment between the methyl donor and acceptor. Furthermore, the imidazole ring of His347 flipped away from its position in Figure 3a and is now far away from the substrate (Figure 4a). Interestingly, it occupies a similar position as in the apo enzyme and product complex of wild-type with a $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ value of about 60°. Thus, His347 is already in

the “released” position in the Asp316Ala-substrate complex, and the induced-fit conformation could not be formed. The fluctuations of the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ dihedral angle of His347 along the reaction coordinate obtained from the PMF simulations are given in Figure 4b. As is evident from Figure 4b, the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ values remained around 60° during the methyl transfer reaction, consistent with the structural data. The loss of such induced-fit conformation is expected to affect the efficiency of methylation process, contributing to the increase of free energy barrier (see below) and decrease of the activity.[20]

The average active site structure of the Tyr386Ala-substrate complex is given in Figure 4c along with the distribution map. Some interactions involving Tyr386 in wild type are lost in the mutant. For instance, the carbon hydrogen bond between the hydroxyl oxygen of Tyr386 and transferable methyl group and “methyl-π” interaction between the substrate and the ring of Tyr386 are lost. Nevertheless, the Tyr386Ala-substrate complex is still in the induced-fit conformation based on the results from the simulations with the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ value of His347 about 160° (Figure 4c). Moreover, there is a large population of structures with short $r(C_M{-}N_{\delta1})$ distances and small θ values, and the near attack conformations is therefore highly populated. Figure 4d shows the fluctuations of the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ dihedral angle of His347 along the reaction coordinate obtained from the PMF simulations for Tyr386Ala. As is evident from Figure 4d, the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ value changed from about 160° to 60° in the early stage of methyl transfer reaction [$R = r(C_M{-}S_\delta) - r(C_M{-}N_{\delta1}) \sim -1.5$ Å], well before the reaction reached the transition state [$R = r(C_M{-}S_\delta) - r(C_M{-}N_{\delta1}) \sim 0.2$Å]. Thus, the active induced-fit orientations of the catalytic groups relative to substrate could not be preserved to allow catalytic residues to exert their catalytic effects during the reaction. The pre-mature collapse of such conformation is expected to affect the efficiency of methylation process, contributing to an increase of free energy barrier (see below) and decrease of the activity.[20] The similar discussion can be for the Tyr396Asp mutant (Figure 4e and 4f) where the active induced-fit conformation collapsed at $R = r(C_M{-}S_\delta) - r(C_M{-}N_{\delta1}) \sim -0.8$Å.

It would be of interest to examine the effect of replacing the imidazole ring of His347 by the phenol ring on the maintenance of the induced-fit conformation. The average structure and distribution map for the His347Phe-substrate complex are given in Figure 4g. As in the cases of the Tyr386Ala and Tyr396Asp mutants, the His347Phe-

substrate complex also has a large population of structures with the relatively short $r(C_M-N_{\delta1})$ distances and small θ values. Nevertheless, the active induced-fit conformation also collapsed before the reaction reached the TS [at $R = r(C_M-S_\delta) - r(C_M-N_{\delta1}) \sim -0.2$ Å]; this is consistent with the experimental observation[20] that more than 80% of the activity was lost due to the His347Phe mutation.

The Free Energy Profiles for Methylation in the Wild-Type and Mutants

The free-energy (PMF) profiles for the methylation reaction in the wild-type CARNMT1 and Asp316Ala, His347Phe, Tyr386Ala and Tyr396Asp mutants as a function of the reaction are given in Figure 5. As is shown in Figure 5, the free energy barriers increase by 3.72, 3.42, 1.71 and 1.14 $kcal\cdot mol^{-1}$ relative to that of wild type as a result of the Asp316→Ala, Tyr386→Ala, and Tyr396→Asp, His347→Phe mutations, respectively. The increases of the free energy barriers are consistent with the decreases of the activity observed in the previous experimental observations.[20] Interestingly, the increase of the free-energy barrier relative to that of wild type seems to be correlated with the inability of the mutant to preserve the active induced-fit conformation. For instance, the Tyr386Ala mutant had its induced-fit conformation collapse [at R = ~ -1.5 Å] earlier than that of Tyr396Asp [at R = ~ -0.8 Å] (indicating its ability to preserve the active conformation is poorer than Tyr396Asp), while its free energy barrier (~ 20 kcal/mol) is also higher than that of Tyr396Asp (18.3 kcal/mol). A similar discussion can also be made for the other mutants.

Conclusion

Previous studies have mainly focused on the question of how precise orientations of catalytic groups/substrate in enzyme-substrate complexes are generated, and these studies have led to the proposals of lock-and-key, induced-fit and conformational selection models as well as their extensions. However, the formation of enzyme-substrate complexes is only the initial step for enzyme catalysis, and generation of the active enzyme-substrate conformations with precise orientations of catalytic groups/substrate does not guarantee that such active orientations of catalytic groups/substrate would be maintained during the enzyme-catalyzed reactions. The ability of enzymes to preserve the active conformations may be crucial in allowing catalytic residues to exert their effects and making the reaction to pass the transition state (TS) easier. Thus, understanding how the active conformations generated through

the substrate binding can be preserved and what are the role of active-site residues in preventing the pre-mature collapse of such conformations is of fundamental importance.

The present study demonstrated the importance of preserving active induced-fit conformations for efficient catalysis. The crucial role of the anchor residues in generating and preserving such conformations has also been demonstrated. The key conclusions of this work are given in Figure 6. The reasons that the wild-type enzyme has a higher catalytic efficiency are not only because of the formation of the catalytic competent conformation in the enzyme-substrate complex but also due to its ability to preserve such active conformation during the reaction process (Figure 6a). For the mutants studied in this work, Type 1 mutant lost the ability to generate the catalytic competent conformation through substrate binding (Figure 6b) and Type 2 mutants could not properly preserve such conformations during the reaction (Figure 6c). Consistent with these observations, the free energy barriers for methyl transfer increases significantly because of the mutations. The results of the present study indicate that generating precise orientations of catalytic groups in enzyme-substrate complexes through the well-known models may be a necessary but not sufficient condition for efficient catalysis and that anchor residues may play an important role in generating and preserving catalytic competent induced-fit configurations during catalysis. The principle discussed here may be used for explanation of losses of activities due to certain mutations for some other enzymes and provide useful information in protein engineering.

Materials and Methods

The CHARMM program[23, 24] was used to perform the QM/MM MD and free energy simulations to study the formation and collapse of the active induced fit configuration in the enzyme-substrate complex and during the catalysis. The effects of the Asp316→Ala, His347→Phe, Tyr386→Ala and Tyr396→Asp mutations on the stability of the induced-fit conformation and the activity of the $N_{\delta 1}$ methylation were also determined. The crystal structure of CARNMT1 (PDB ID: 5YF1) containing carnosine and SFG[20] and was used to create the reaction state structure of the enzyme-substrate complex for methylation, and the N and C atoms of SFG was manually changed to C and S atoms, respectively, to generate SAM. The crystal structure of the apo enzyme (PDB ID: 5YF0) containing SAM was also used for the simulations. The simulations were also performed for the case where carnosine was removed from 5YF1 in the initial structure (with SFG

changed to SAM). This allows testing whether His347 could move back to its position in the apo enzyme when the substrate was removed from the enzyme-substrate complex. To perform the QM/MM MD simulations on the mutants, the Asp316→Ala, His347→Phe, Tyr386→Ala, or Tyr396→Asp mutation was performed manually in the crystal structure of the wild type (5YF1).

The $-CH_2-CH_2-S^+(Me)-CH_2-$ part of SAM and the whole substrate were treated by QM, and the rest of the system was treated by MM. The QM region and MM region were separated by using the link-atom approach[25] implemented in CHARMM; the QM region contains a total of 44 atoms, with the net charge of +1. The stochastic boundary molecular dynamics method[26] was used for the QM/MM MD and PMF simulations. A modified TIP3 water model[27] was used for the solvent. The DFTB3 method[28-31] was applied to the QM atoms, and the CHARMM potential function (PARAM27)[32] was adopted for the MM atoms. Hydrogen atoms of the system were built with the HBUILD module[33] in CHARMM. The reference center assigned by the system was chosen to be the $N_{\delta1}$ atom of target substrate carnosine. The system was separated into a reaction zone and a reservoir region, and the reaction zone was further divided into a reaction region (a sphere with radius r of 20 Å) and a buffer region (20 Å ≤ r ≤ 22 Å). The whole system contained around 5700 atoms, including about 460 water molecules. The initial structures were optimized for the entire stochastic boundary systems with the steepest descent (SD) and adopted-basis Newton−Raphson (ABNR) methods. The systems were gradually heated from 50.0 to 310.15 K in 50 ps. The time step for the integration of equation of motion was 1 fs, and the coordinates were saved every 50 fs for analyses. For each of the systems, the QM/MM MD simulations of 2 ns were carried out. The distribution maps of $r(C_M\text{-}N_{\delta1})$ and θ (Figure 7) were generated in each case and the fluctuations of the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ dihedral angle and $r(\text{carnosine-}C_{\varepsilon1}\cdots\text{H347-}C_{\varepsilon1})$ distance of His347 were monitored.

The umbrella sampling method[34] was applied along with the weighted histogram analysis method (WHAM)[35] in the determination of the free energy change along the reaction coordinate for the methyl transfer from S-adenosyl-L-methionine (SAM) to the $N_{\delta1}$ atom of carnosine. The reaction coordinate was defined as a linear combination of $r(C_M-N_{\delta1})$ and $r(C_M-S_\delta)$ [$R = r(C_M-S_\delta) - r(C_M-N_{\delta1})$]. For the methyl transfer process, 28 windows were used, and for each window, the production runs of 100 ps were performed

after 50 ps equilibration. The force constants of the harmonic biasing potentials used in the PMF simulations were 50 to 400 kcal·mol^{-1}·Å^{-2}. In each case, four independent PMF simulations were performed. The free energies and statistical errors were taken as the average values and standard deviations from the four runs, respectively. The fluctuations of the χ(C-C_α-C_β-C_γ) dihedral angle as well as r(carnosine-$C_{\varepsilon1}$···H347-$C_{\varepsilon1}$) distance of His347 were monitored in each PMF window along the reaction coordinate to detect when the induced fit configuration was formed or collapsed during the methyl transfer reaction.

Acknowledgments

We thank Professor Martin Karplus for a gift of the CHARMM program. This work was partly supported by the Natural Science Foundation of China (No. 22177064 to P.Q.) and the Natural Science Foundation of Shandong Province (No. ZR2021MB050 to P.Q.).

References

1. Boehr, D. D.; Nussinov, R.; Wright, P. E., The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chemical Biology* **2009,** *5* (11), 789-796.
2. Boehr, D. D.; Wright, P. E., How do proteins interact? *Science* **2008,** *320* (5882), 1429-1430.
3. Csermely, P.; Palotai, R.; Nussinov, R., Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends in Biochemical Sciences* **2010,** *35* (10), 539-546.
4. Pang, X. D.; Zhou, H. X., Rate Constants and Mechanisms of Protein-Ligand Binding. In *Annual Review of Biophysics, Vol 46*, Dill, K. A., Ed. 2017; Vol. 46, pp 105-130.
5. Redhair, M.; Atkins, W. M., Analytical and functional aspects of protein-ligand interactions: Beyond induced fit and conformational selection. *Archives of Biochemistry and Biophysics* **2021,** *714*.
6. Koshland, D. E., THE KEY-LOCK THEORY AND THE INDUCED FIT THEORY. *Angewandte Chemie-International Edition* **1994,** *33* (23-24), 2375-2378.
7. Du, X.; Li, Y.; Xia, Y. L.; Ai, S. M.; Liang, J.; Sang, P.; Ji, X. L.; Liu, S. Q., Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *International Journal of Molecular Sciences* **2016,** *17* (2).
8. Johnson, K. A., Role of induced fit in enzyme specificity: A molecular forward/reverse switch. *Journal of Biological Chemistry* **2008,** *283* (39), 26297-26301.
9. Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. Y., Binding affinities of host-guest, protein-ligand, and protein-transition-state complexes. *Angewandte Chemie-International Edition* **2003,** *42* (40), 4872-4897.
10. Herschlag, D., THE ROLE OF INDUCED FIT AND CONFORMATIONAL-CHANGES OF ENZYMES IN SPECIFICITY AND CATALYSIS. *Bioorganic Chemistry* **1988,** *16* (1), 62-96.
11. Stank, A.; Kokh, D. B.; Fuller, J. C.; Wade, R. C., Protein Binding Pocket Dynamics. *Accounts of Chemical Research* **2016,** *49* (5), 809-815.
12. E, F., The influence of configuration on enzyme activity. (Translated from German). *Dtsch Chem Ges* **1894,** *27*, 2984 –2993.
13. Koshland, D. E., APPLICATION OF A THEORY OF ENZYME SPECIFICITY TO PROTEIN SYNTHESIS. *Proceedings of the National Academy of Sciences of the United States of America* **1958,** *44* (2), 98-104.
14. Tsai, C. J.; Kumar, S.; Ma, B. Y.; Nussinov, R., Folding funnels, binding funnels, and protein function. *Protein Science* **1999,** *8* (6), 1181-1190.
15. Tsai, C. J.; Ma, B. Y.; Nussinov, R., Folding and binding cascades: Shifts in energy landscapes. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** *96* (18), 9970-9972.
16. Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M., Substrate conformational transitions in the active site of chorismate mutase: Their role in the catalytic mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** *98* (16), 9032-9037.
17. Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M., Understanding the role of active-site residues in chorismate mutase catalysis from molecular-dynamics simulations. *Angewandte Chemie-International Edition* **2003,** *42* (13), 1508-1511.
18. Rajamani, D.; Thiel, S.; Vajda, S.; Camacho, C. J., Anchor residues in protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (31), 11287-11292.
19. Boldyrev, A. A.; Aldini, G.; Derave, W., Physiology and pathophysiology of carnosine. *Physiological reviews* **2013**.
20. Cao, R.; Zhang, X.; Liu, X.; Li, Y.; Li, H., Molecular basis for histidine N1 position-specific methylation by CARNMT1. *Cell Research* **2018,** *28* (4), 494-496.
21. Bruice, T. C., Computational approaches: Reaction trajectories, structures, and atomic motions. Enzyme reactions and proficiency. *Chemical Reviews* **2006,** *106* (8), 3119-3139.
22. Deng, H.; Ma, Y.; Ren, W. S.; Vuong, V. Q.; Qian, P.; Guo, H., Structure and Dynamics of the Reactive State for the Histidine Methylation Process and Catalytic Mechanism of SETD3: Insights from Quantum

Mechanics/Molecular Mechanics Investigation. *Acs Catalysis* **2020,** *10* (22), 13314-13322.
23. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: The Biomolecular Simulation Program. *Journal of Computational Chemistry* **2009,** *30* (10), 1545-1614.
24. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M., CHARMM - a program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* **1983,** *4* (2), 187-217.
25. Field, M. J.; Bash, P. A.; Karplus, M., A COMBINED QUANTUM-MECHANICAL AND MOLECULAR MECHANICAL POTENTIAL FOR MOLECULAR-DYNAMICS SIMULATIONS. *Journal of Computational Chemistry* **1990,** *11* (6), 700-733.
26. Brooks, C. L.; Brunger, A.; Karplus, M., Active-site dynamics in protein molecules - a stochastic boundary molecular-dynamics approach. *Biopolymers* **1985,** *24* (5), 843-865.
27. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., COMPARISON OF SIMPLE POTENTIAL FUNCTIONS FOR SIMULATING LIQUID WATER. *Journal of Chemical Physics* **1983,** *79* (2), 926-935.
28. Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G., Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Physical Review B* **1998,** *58* (11), 7260-7268.
29. Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M., A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *Journal of Physical Chemistry B* **2001,** *105* (2), 569-585.
30. Christensen, A. S.; Kubař, T.; Cui, Q.; Elstner, M., Semiempirical quantum mechanical methods for noncovalent interactions for chemical and biochemical applications. *Chem. Rev.* **2016,** *116* (9), 5301-37.
31. Lu, X.; Gaus, M.; Elstner, M.; Cui, Q., Parametrization of DFTB3/3ob for magnesium and zinc for chemical and biological applications. *Journal of Physical Chemistry B* **2015,** *119* (3), 1062-1082.
32. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B* **1998,** *102* (18), 3586-3616.
33. Brünger, A. T.; Karplus, M., Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins* **1988,** *4* (2), 148-156.
34. Torrie, G. M.; Valleau, J. P., Monte-Carlo free-energy estimates using Non-Boltzmann sampling - application to subcritical Lennard-Jones fluid. *Chemical Physics Letters* **1974,** *28* (4), 578-581.
35. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The weighted histogram analysis method for free-energy calculations on biomolecules .1. the method. *Journal of Computational Chemistry* **1992,** *13* (8), 1011-1021.

Figures Captions

Figure 1. (a) lock-and-key (LK) model. (b) Induced fit model (IF) and original conformational selection model (CS). Blue: enzyme; Yellow: substrate.

Figure 2. Superposition of apo enzyme (5YF0, orange) and enzyme-substrate complex (5YF1 with SAM replaced by SFG, blue). The χ(C-Cα-Cβ-Cγ) dihedral angle changes from 55° in 5YF0 to 157° in 5YF1, and r(Carnosine-Cε1⋯H347-Cε1) decreases from about 7.5Å to 3.5Å accordingly. Part of the SFG/SAM, carnosine and water molecules are shown in ball and stick, and the rest of the active sites are shown in stick. Waters from 5YF1 are colored in red, and those from 5YF0 in orange. Some distances are also given (in angstroms).

Figure 3. (a) The average structure for the active site of the reactant complex containing SAM and carnosine in wild-type CARNMT1 (left) along with the distribution map of $r(C_M\text{-}N_{\delta 1})$ and θ (right) obtained from the QM/MM MD simulations. The $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ dihedral angle of His347 is about 160°, and r(carnosine-$C_{\varepsilon 1}$⋯H347-$C_{\varepsilon 1}$) is 3.7Å. (b) The fluctuations of the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ dihedral angle of His347 in the PMF windows along the reaction coordinate. The red vertical bar designates the location for the transition state of the methyl transfer. (c) The average structure for the active site near the transition state obtained from the QM/MM PMF simulations for methylation in wild-type CARNMT1. (d) Left: The $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ values for the apo enzyme during the QM/MM simulations based on the crystal structure of 5YF0. The red circles indicate the system reached the induced fit configuration even without the substrate. Right: The $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ values from the simulations after the substrate was first removed from the active site in the enzyme-substrate complex. substrate complex.

Figure 4. (a), (c), (e) and **(g)** The average structures for the active sites of the Asp316Ala-substrate, Tyr386Ala-substrate, Tyr396Asp-substrate and His347Phe-substrate complexes, respectively, along with the corresponding distribution maps. **(b), (d), (f)** and **(h)** The fluctuations of the $\chi(C\text{-}C_\alpha\text{-}C_\beta\text{-}C_\gamma)$ dihedral angle of His347 in the PMF windows along the reaction coordinate for the Asp316Ala, Tyr386Ala, Tyr396Asp and His347Phe complexes, respectively.

Figure 5. Free energy (PMF) profiles for the methylation reaction in wild-type CARNMT1, Asp316Ala, His347Phe, Tyr386Ala and Tyr396Asp, respectively, as a function of the reaction coordinate [$R = r(C_M{-}S_\delta) - r(C_M{-}N_{\delta 1})$]. Wild type: blue line with the free energy barrier of 16.59 kcal/mol; Asp316Ala: yellow line with the free energy barrier of 20.31 kcal/mol; Tyr386Ala: grey line with the free energy barrier of 20.01 kcal/mol; Tyr396Asp: green line with the free energy barrier of 18.30 kcal/mol; His347Phe: orange line with the free energy barrier of 17.73 kcal/mol.

Figure 6. Wild-type enzyme can preserve catalytic competent induced-fit configuration at TS. Type 1 mutation of key anchor residue (e.g, Asp316Ala) destroys the enzyme's ability to form the induced-fit conformation with substrate. Type 2 mutation (e.g., Tyr386Ala) leads to the collapse of the induced-fit conformation before the reaction reaches TS, even though such conformation can be formed in the enzyme-substrate complex.

Figure 7. Relative orientation of SAM and carnosine in the reactant complex. Here, θ is defined as the angle between two vectors r_1 and r_2, which represent the direction of the C_M−S_δ bond and the direction of the lone pair of electrons on $N_{\delta 1}$, respectively, which can determine the connection between the distributions and reactant structure properties.

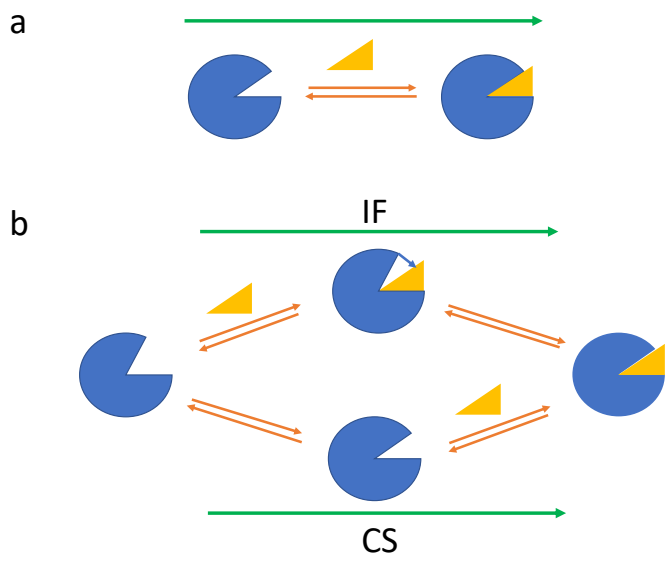


Figure 1

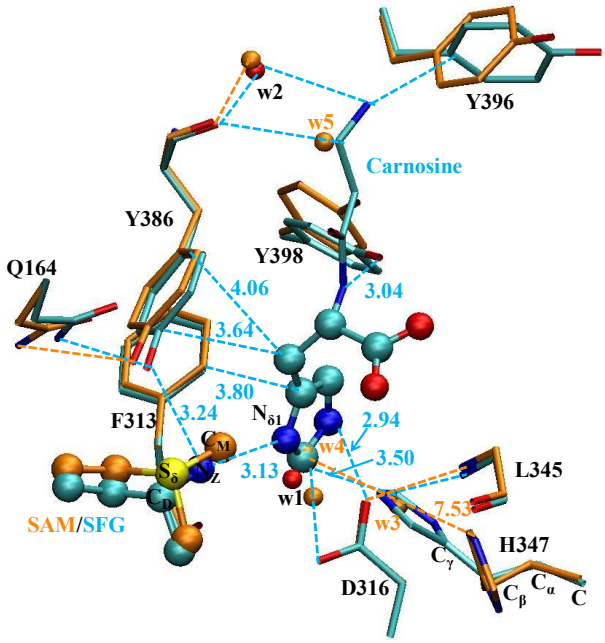

w2
Y396
w5
Carnosine
Y386
Y398
Q164
4.06
3.04
3.64
3.80
F313
3.24
$N_{\delta 1}$
2.94
C_M
w4
3.50
L345
S_δ
N_Z
3.13
w1
C_D
7.53
SAM/SFG
w3
H347
C_γ
D316
C_β
C_α
C

Figure 2

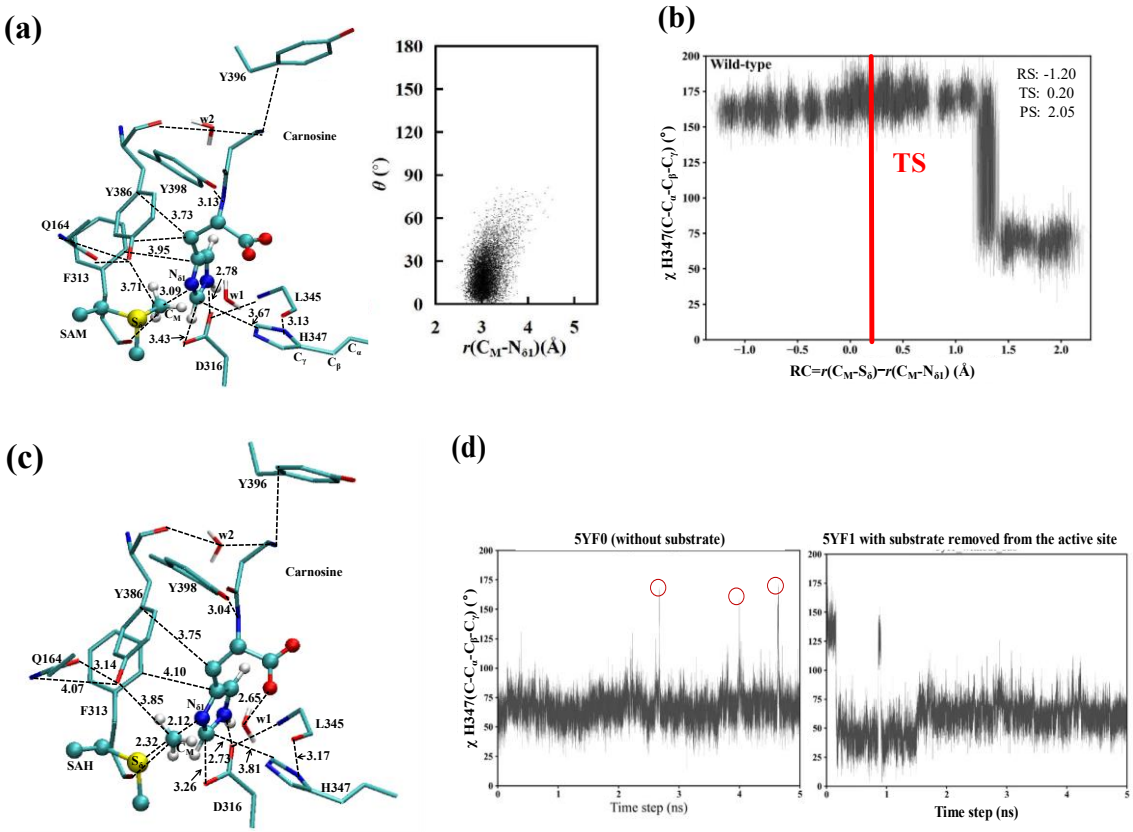

(a)
Y396
w2
Carnosine
Y386
Y398
3.13
3.73
Q164
3.95
F313
3.71
3.09
2.78
w1
L345
3.67
3.13
SAM
3.43
H347
D316
θ (°)
r(C_M-N_δ1)(Å)
(b)
Wild-type
RS: -1.20
TS: 0.20
PS: 2.05
TS
χ H347(C-C_α-C_β-C_γ) (°)
RC=r(C_M-S_δ)−r(C_M-N_δ1) (Å)
(c)
Y396
w2
Carnosine
Y386
Y398
3.04
3.75
Q164
3.14
4.10
4.07
3.85
2.65
F313
2.12
w1
L345
2.32
3.17
SAH
2.73
3.81
3.26
H347
D316
(d)
5YF0 (without substrate)
5YF1 with substrate removed from the active site
χ H347(C-C_α-C_β-C_γ) (°)
Time step (ns)
Time step (ns)

Figure 3

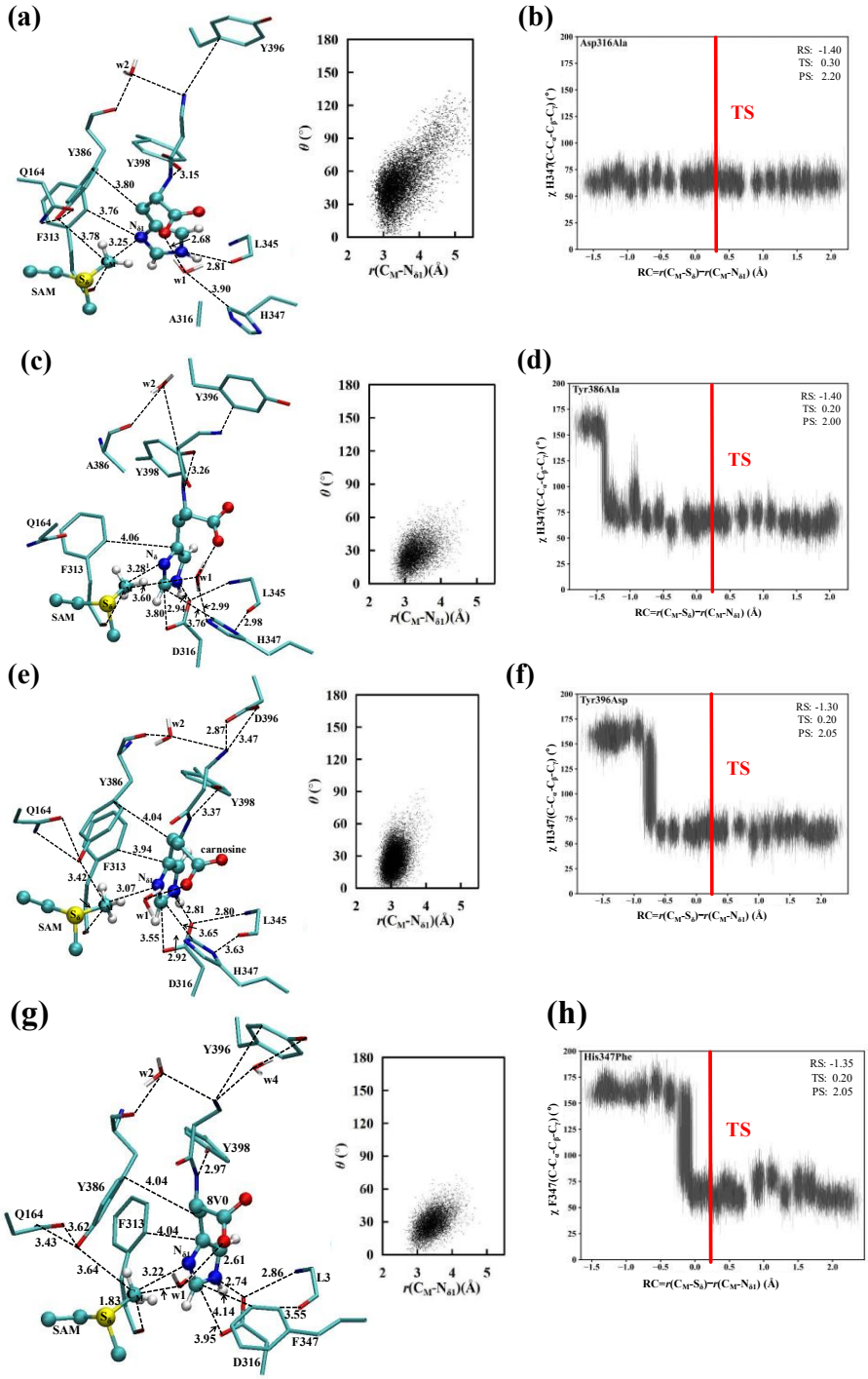

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
Asp316Ala
RS: -1.40
TS: 0.30
PS: 2.20
TS
Tyr386Ala
RS: -1.40
TS: 0.20
PS: 2.00
TS
Tyr396Asp
RS: -1.30
TS: 0.20
PS: 2.05
TS
His347Phe
RS: -1.35
TS: 0.20
PS: 2.05
TS
SAM
carnosine
8V0
θ (°)
r(C_M-N_δ1)(Å)
RC=r(C_M-S_δ)−r(C_M-N_δ1) (Å)

Figure 4

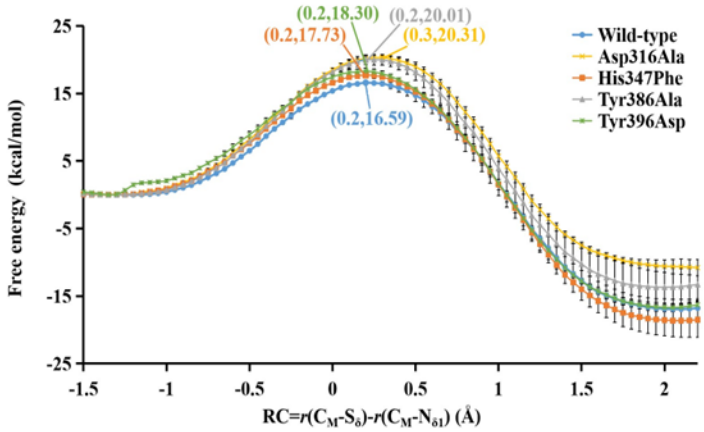

(0.2,18.30)
(0.2,20.01)
(0.2,17.73)
(0.3,20.31)
(0.2,16.59)
Wild-type
Asp316Ala
His347Phe
Tyr386Ala
Tyr396Asp
Free energy (kcal/mol)
25
15
5
-5
-15
-25
-1.5
-1
-0.5
0
0.5
1
1.5
2
RC=r(C_M-S_δ)-r(C_M-N_δ1) (Å)

Figure 5

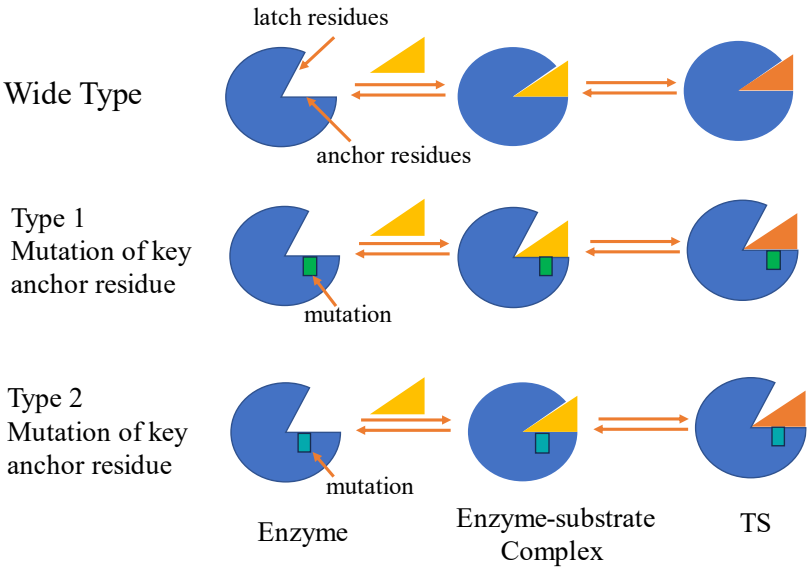

latch residues
Wide Type
anchor residues
Type 1
Mutation of key
anchor residue
mutation
Type 2
Mutation of key
anchor residue
mutation
Enzyme
Enzyme-substrate
Complex
TS

Figure 6

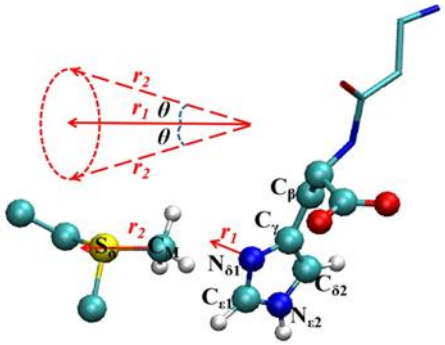
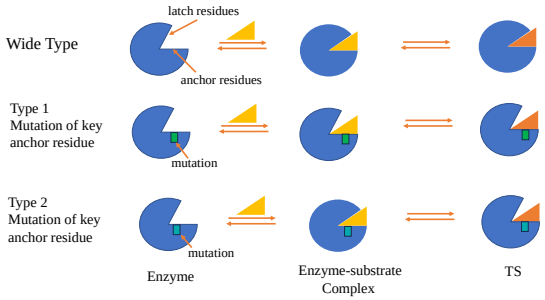


Figure 7

TOC


latch residues
Wide Type
anchor residues
Type 1
Mutation of key
anchor residue
mutation
Type 2
Mutation of key
anchor residue
mutation
Enzyme
Enzyme-substrate
Complex
TS