Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2023 November 16.

Published in final edited form as:

J Am Chem Soc. 2022 November 16; 144(45): 20739–20751. doi:10.1021/jacs.2c08515.

How To Stabilize Carbenes in Enzyme Active Sites Without Metal Ions

Rui Lai[†,‡], **Qiang Cui**[†,¶,§]

[†]Department of Chemistry, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, United States

[‡]Dalian Institute of Chemical Physics, Chinese Academy of Science, 457 Zhongshan Road, Dalian, China 116023

[¶]Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, United States

[§]Department of Biomedical Engineering, Boston University, 44 Cummington Mall, Boston, MA 02215, United States

Abstract

Carbenes are highly reactive compounds with unique values to synthetic chemistry. However, a small number of natural enzymes have been shown to utilize carbene chemistry, and artificial enzymes engineered with directed evolution required transition metal ions to stabilize the carbene intermediates. To facilitate the design of broader classes of enzymes that can take advantage of the rich carbene chemistry, it is thus important to better understand how to stabilize carbene species in enzyme active sites without metal ions. Motivated by our recent studies of the anaerobic ergothionine biosynthesis enzyme EanB, we examine carbene-protein interaction with both cluster models and QM/MM simulations. The cluster calculations find that an N-heterocyclic carbene interacts strongly with polar and positively charged protein motifs. In particular, the interaction between a guanidinium group and carbene is as strong as ~30 kcal/mol, making arginine a great choice for the preferential stabilization of carbenes. We also compare the WT EanB and its mutant in which the key tyrosine was replaced by a non-natural analog (F2Tyr) using DFTB3/MM simulations. The calculations suggest that the carbene intermediate in the F2Tyr mutant is more stable than that in the WT enzyme by ~3.5 kcal/mol, due to active site rearrangements that enable a nearby arginine to better stabilize the carbene in the mutant. Overall, the current work lies the foundation for the pursuit of enzyme designs that can take advantage of the unique chemistry offered by carbenes without the requirement of metal ions.

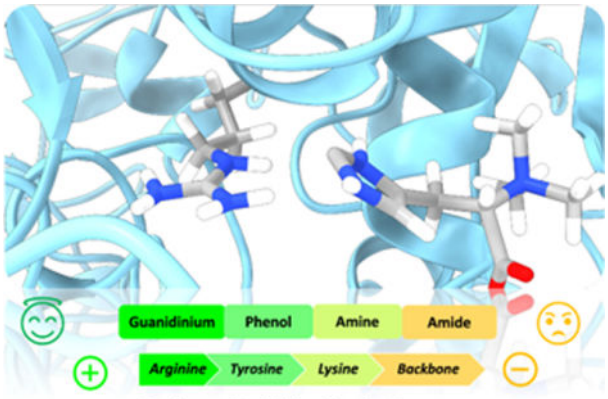
Graphical Abstract

qiangcui@bu.edu, Tel:(+1)-617-353-6189.

Supporting Information Available

The Cartesian coordinates for optimized model systems in Figs. 1 and 2 are included. Additional results for the model compounds calculations and QM/MM simulations are also included. This material is available free of charge via the Internet at http://pubs.acs.org/.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Carbene Stabilization in Enzyme

1. Introduction

Carbenes are a class of highly reactive compounds that feature rich chemistry.[1,2] Therefore, they are uniquely valuable in organic synthesis. Carbenes can exhibit diverse reactivities with different spin states. For example, singlet carbenes can have both electrophilic and nucleophilic characters and triplet carbenes have diradical reactivity.[3] Most carbene compounds are difficult to isolate due to their high reactivities, except for some relatively stable ones, such as N-heterocyclic carbenes, which have been used to conduct novel chemical reactions.[4] It appears that nature has not "discovered" carbene chemistry, as the number of enzymes that have been shown to involve a carbene intermediate is very limited (Scheme 1), including the thiamin diphosphate (ThDP) enzyme,[5] orotidine monophosphate (OMP) decarboxylase[6] and, more recently, the anaerobic ergothionine biosynthesis enzyme (EanB).[7,8] Such realization motivated the design of novel enzymes that may take advantage of carbene chemistry. For example, utilizing synthetic carbene precursors with repurposed heme-proteins and directed evolution, Arnold and co-workers successfully expanded the catalytic repertoire of enzymes toward novel chemical reactions not known in natural enzymes;[9] the selective carbene transfer reactions in those heme proteins involve stabilization of carbine by iron-porphyrins.[9–11]

To employ carbene chemistry more broadly in designed enzymes, it is essential to explore alternative strategies that stabilize carbenes without using transition metal ions. To this end, it is important to understand intermolecular interactions that may help stabilize carbenes in an enzyme active site; to the best of our knowledge, this has not been explored systematically in either experimental or computational studies. Indeed, since many carbenes are charge neutral, it is not intuitive to predict interactions that preferentially stabilize carbene over related species (*vide infra*). Establishing the physical nature of carbene interactions will also be instrumental to the development of proper computational models for treating the corresponding stabilization effects during the rational design process. For example, the fact that (singlet) carbenes feature both a lone pair and an empty valence orbital suggests that interactions involving them are likely beyond electrostatics and therefore require more sophisticated computational models than the standard empirical force fields.[12–15]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

As another potential strategy, one may incorporate non-natural amino acid(s) into the active site to provide additional stabilization of carbenes. The contribution from the non-natural amino acid substitution, however, is not straightforward to predict. For example, in a recent study of the anaerobic ergothionine biosynthesis enzyme (EanB),[7,8] to better understand the role of Tyr353 in catalysis, the authors replaced Tyr353 with a non-natural tyrosine analogue, 3,5-difluoro tyrosine (F2Tyr, see Scheme 2), whose hydroxyl pK_a is 2.8 units lower than that of Tyr. The rate constant (k_{cat}) for EanB with F2Tyr ($EanB_{Y353F2Tyr}$) was slightly lower than that of the wild-type (WT) EanB, but the substrate (hercynine) deuterium exchange is much more efficient in the $EanB_{Y353F2Tyr}$ variant relative to the WT enzyme,[8] suggesting that the carbene intermediate might be more stable in $EanB_{Y353F2Tyr}$. The origin of the stability difference was not clear, since the protonation state of the Tyr remains the same in the carbene intermediate and the reactant state (see Scheme 1). It is likely that the interaction patterns are altered in the $EanB_{Y353F2Tyr}$ active site, leading to different degrees of stabilization of the carbene intermediate relative to the WT enzyme, although a molecular level of understanding requires detailed computational analyses.

In this study, we conduct computational analysis of carbene stabilization by enzyme motifs using both cluster models and realistic systems. The cluster models are constructed based on key features of several enzymes that have been proposed to involve a carbine intermediate: EanB,[7,8] the ThDP enzyme[5] and the OMP decarboxylase;[6] stabilization of the carbene species is analyzed with multiple QM methods, ranging from semi-empirical density functional tight binding (DFTB) through various DFT methods to correlated *ab initio* (MP2, SCS-MP2 and CCSD(T)) approaches. The magnitude and nature of stabilization is further analyzed with an energy decomposition scheme.[16] Finally, we also conduct DFTB3/MM simulations to explicitly compare the properties of the carbene intermediate in the WT EanB and the $EanB_{Y353F2Tyr}$ variant; in particular, we aim to understand the underlying mechanism for the higher degree of carbene stabilization by F2Tyr.

We observe that carbenes feature strong intermolecular interactions with various enzyme groups, especially Arginine and Tyrosine sidechains; the degree of stabilization is lower with other charge-neutral groups, such as backbone amide and sidechain of Thr. Therefore, carbene intermediates can be substantially stabilized relative to the reactant state if the proper protein groups are well positioned. In fact, the higher stability of the carbene intermediate in $EanB_{Y353F2Tyr}$ relative to the WT EanB is due largely to the better positioning of an Arginine residue. Therefore, our study has provided valuable guidance to the design strategies that can be utilized to stabilize carbenes in enzyme active sites to take advantage of the unique stereospecificity/selectivity of carbene chemistry.

2 Computational Methods

2.1 Cluster calculations for carbene stabilization

To probe the interaction between carbene and typical enzyme motifs, several sets of cluster models are constructed based on the active sites of the three enzymes mentioned earlier: EanB (Fig. 1), ThDP enzyme and OMP decarboxylase (Fig. 2). For a comparative analysis of stabilization, these models are also studied with the carbene replaced by the reactant species for the corresponding enzymes. The geometries for these clusters are optimized with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the second-order Möller-Plesset perturbation theory (MP2)[17] and the aug-cc-pVTZ basis set (aTZ),[18–21] using the Gaussian 16 program.[22] For comparison, interaction energies with these optimized structures are also calculated using DFTB/3OB, DFT methods with different functionals (B3LYP,[23–25] CAM-B3LYP,[26] PBE0,[27] LC-ωHPBE,[28] ωB97XD,[29] MN15[30] and MN15L[31]), spin-component scaled MP2 (SCS-MP2), and coupled-cluster methods (CCSD(T)). Such comparisons are worthwhile due to the unique electronic structure of carbenes, which may feature a non-negligible degree of multi-reference character.[32] MP2, SCS-MP2 and CCSD(T) calculations are performed with the GAMESS package,[33–36] DFTB with CHARMM,[37] and the others with Gaussian16.[22] The complete basis set limit energies (E_{CBS}) are calculated based on the exponential extrapolation scheme;[38] the detailed equations and results are included in the Supporting Information.

The interaction energies are further analyzed using the canonical molecular orbital energy decomposition analysis (CMO-EDA) with MP2 using the aug-cc-pVDZ (aDZ), aug-cc-pVTZ (aTZ) and aug-cc-pVQZ (aQZ) basis sets.[18–21] The results with the aQz basis set and correction of the basis set superposition error (BSSE) are shown in Table 2. Results at other levels of theory are included in the Supporting Information for comparison (see Table S1 and S3, S4). The CMO-EDA calculations are performed with the GAMESS package.[16,36,39]

In CMD-EDA, the total interaction energy is decomposed into five terms, including electrostatics (ΔE^{ele}), exchange (ΔE^{ex}), repulsion (ΔE^{rep}), polarization (ΔE^{pol}) and dispersion energies (ΔE^{disp}):

$$\Delta E^{MP2} = \Delta E^{ele} + \Delta E^{ex} + \Delta E^{rep} + \Delta E^{pol} + \Delta E^{disp}. \quad (1)$$

It is common to group ΔE^{rep} and ΔE^{ex} together as the exchange-repulsion (ΔE^{ex-rep}) term, as shown in Table 2. Basis set superposition error (BSSE) is considered in the intermolecular interaction calculations using the counterpoise method originally proposed by Boys and Bernadi,[40] which is similar to the posteriori BSSE correction method defined by Xantheas.[41]

2.2 QM/MM simulations of WT EanB and EanB$_{Y353F2Tyr}$

To compare carbene stabilization in the WT EanB and EanB$_{Y353F2Tyr}$, QM/MM simulations are conducted using CHARMM.[37] The crystal structure of EanB with Cys412 polysulfide in complex with hercynine (PDB: 6KTV)[7] is used as the starting structure. The QM/MM setup is illustrated in Fig. 3, in which a water droplet of 25 Å radius centered on the terminal sulfur atom of Cys412 persulfide is used to solvate the enzyme. Non-crystallographic waters within 2.5 Å of crystallographic atoms are deleted. The system is partitioned into three regions within the framework of the generalized solvent boundary potential (GSBP):[43–45] a QM region that consists of 134 atoms, an inner MM region that contains all non-QM atoms within a 28 Å sphere surrounding the active site (7961 atoms), and the frozen outer MM region (1971 atoms). The QM region contains key reacting residues, the hercynin substrate and some residues of α-helix 18 (i.e. Glu345, Tyr353/Y353F2Tyr, Tyr411, Cys412, Gly413, Thr414, Gly415, Trp416, Arg417 and Gly418), and QM link atoms, using the DIV scheme,[46] are placed between the α and β carbons of the QM residues. The QM region is treated with the DFTB3 level of theory[47–49] with the 3OB-3–1 parameter set;[50,51]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

as discussed in Ref.[7] and the Supporting Information, the C-H repulsive potential has been modified to better balance the C-H proton affinity of the substrate and the O-H proton affinity of Tyr353 for the carbene generation. Compared to our previous work,[7] the calculations here include the Grimme's third version semi-empirical dispersion correction (D3[52,53]) for the DFTB atoms.

The inner region is propagated with classical Newtonian dynamics. A buffer region of 3 Å from the edge of the inner region is treated with Langevin dynamics, and protein atoms in the buffer region are harmonically restrained with force constants determined from the average of crystallographic B factors.[54] Relevant free energy profiles are computed using metadynamics simulations[55] with the PLUMED package[56] interfaced to CHARMM. The collective variables (CVs) for metadynamics are the antisymmetric stretch describing the proton transfers between Tyr353/Y353F2Tyr and the hercynine substrate. SHAKE[57] is applied to all bonds involving hydrogen atoms that do not participate in proton transfer reactions; due to the involvement of proton transfers, an integration time step of 0.5 fs is adopted. The well-tempered multiple walker metadynamics method is employed with 50 walkers; each walker samples ~2.3 ns, leading to a cumulative simulation time of ~115 ns for each enzyme system. In terms of wall time, using two Intel Xeon Gold 6132 processor, it takes approximately 48 hours to complete 100 ps (with an integration time of 0.5 fs) of sampling for each walker. More details about the QM/MM metadynamics simulations and benchmark calculations are included in the Supporting Information.

3 Results and Discussions

In this section, we first examine the cluster models inspired by EanB to compare the interactions between the carbene and nearby enzyme groups with those for the reactant (using an imidazole to mimic hercynin); results for cluster models relevant to two other enzymes (Fig. 2) are also briefly discussed although the magnitude of stabilization due to neutral amine is much more modest. The interaction energies for all these cluster models are then decomposed into various components to gain further insights into the physical nature of intermolecular interactions involving carbene. Finally, we analyze the roles of a non-natural amino acid (F2Y) in stabilizing the carbene intermediate of EanB using QM/MM simulations.

3.1 Interactions Between Carbene and Active Site Residues in EanB Enzyme

In our previous work,[7,8] using both DFTB3/MM metadynamics simulations and DFT cluster models, we examined three possible reaction pathways for EanB catalysis. Both types of calculations found that the S-S bond cleavage is the rate-determining step, while Tyr353 participates in the substrate activation step that is a prerequisite for the subsequent S-S bond cleavage. Once the imidazole moiety in the substrate hercynin is activated (protonated) by Tyr353, the deprotonated Tyr353 may extract the proton of ε-carbon to produce a carbene intermediate (Scheme 1C). Both DFTB3/MM and DFT cluster calculations suggested that a carbene-based mechanism is energetically feasible, which motivated subsequent deuterium exchange experiment that further supported the involvement of a carbene species.[8] Inspired by the observation of a carbene intermediate in the absence of metal ions in the EanB active

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

site, we construct a set of cluster models (see Fig. 1) to probe the interactions between the carbene intermediate and nearby enzyme residues, in comparison to those computed for the reactant (modeled by imidazole). The interaction energies are shown in Tables 1 and S1.

Hydroxyl Group—Following its formation, the carbene intermediate in EanB is likely to interact directly with the Tyr (or F2Tyr), which can provide a substantial amount of stabilization. As references, the phenol groups in Tyr and F2Tyr form hydrogen bonds with the neutral N atom in imidazole with distances of 1.79 Å and 1.71 Å, respectively (Fig. 1A–B); the shorter distance suggests that F2Tyr forms a stronger interaction, which is consistent with its lower sidechain pK_a. The calculated interaction energies are −7.5 kcal/mol and −8.2 kcal/mol, respectively, for Tyr and F2Tyr, at the SCS-MP2 level of theory in the estimated complete basis set (CBS) limit. For the carbene, the hydroxyl group prefers interactions with the ε-carbon that has a lone pair of electrons. The H-ε-C distances are 1.83 Å and 1.70 Å, respectively, for Tyr and F2Tyr; the corresponding interaction energies are −9.9 kcal/mol and −13.7 kcal/mol, at the SCS-MP2/CBS level of theory. Similarly, at the CCSD(T)/CBS levels of theory, the interaction between the carbene and F2Tyr (−15.4 kcal/mol) is stronger than that of carbene-Tyr (−11.2 kcal/mol). Therefore, it is clear that the interaction with Tyr/F2Tyr is stronger for the carbene than with the substrate imidazole, especially with F2Tyr. Indeed, the magnitude of ~15 kcal/mol is substantially stronger than a typical hydrogen bonding interaction that involves charge-neutral species.

In the previous study,[7] we found that Thr414 also plays an important role in stabilizing the substrate through a hydrogen bonding interaction with N_δ atom of the imidazole. Therefore, we also probe the interaction between a Thr sidechain and the carbene/imidazole using the cluster models shown in Fig. 1C. The optimized geometries suggest that the hydroxyl group in Thr prefers to interact with the imidazole N atom or accepts a hydrogen bond from the NH group in the carbene, instead of donating a hydrogen-bond to the carbene ε-carbon; these trends are also observed in the DFTB3/MM simulations (see below). The magnitude of interactions with Thr is comparable for the carbene and imidazole, suggesting that the Thr sidechain has limited ability to preferentially stabilize carbenes.

Guanidinium Group—In the active site of EanB, Arg417 is in close contact with the substrate and thus may provide a significant stabilization. With the cluster models (Fig. 1D), the interaction energies for Arg-carbene and Arg-imidazole are −24.4 kcal/mol and −20.3 kcal/mol, respectively, at the SCS-MP2/CBS level of theory, and both are stronger by ~2 kcal/mol at the CCSD(T)/CBS level of theory. The ΔE^{inter} of Arg-carbene is 4.1 kcal/mol larger in magnitude than that of Arg-imidazole, suggesting that, in EanB, Arg417 may provide a significant preferential stabilization of the carbene intermediate if it is well positioned (see below). Compared to the interactions with Tyr and F2Tyr, the remarkable magnitude of interaction between carbene and Arginine suggests that the latter can potentially provide much more effective stabilization of carbenes.

Amide Group—In addition to interactions with the sidechains discussed so far, the substrate in EanB also interacts with the backbone amide groups of α-helix 18, including Cys412, Gly413, Thr414, Gly415, Trp416, Arg417 and Gly418. Therefore, it is important to analyze the interaction of imidazole and carbene with amide groups. With the model systems

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

shown in Fig. 1E, the calculated ΔE^{inter} at the CCSD(T)/CBS level of theory for backbone-carbene and backbone-imidazole are −6.2 kcal/mol and −5.3 kcal/mol, respectively. Although the difference of ~1 kcal/mol is modest, there are multiple backbone amide groups that interact with the substrate in EanB. Therefore, the accumulative contribution from backbone amide interactions toward carbene stabilization can be substantial.

Amine Group—Besides backbone amide groups, the amine groups (such as lysine side chain groups) may also be key factors that stabilize the carbene intermediate. For example, among the several proposed mechanisms for the proficient decarboxylation of OMP,[58] Lee and Houk[6] proposed a mechanism in which a stable carbene is involved for orotidine 5'-monophosphate decarboxylase (OMPD) by using quantum mechanical models. The proposed carbene intermediate is shown in Scheme 1B. In the active site of OMPD, there are two lysine (Lys42 and Lys72) residues near the substrate, participating in the reaction, as shown in Fig. 2A. These lysine residues may either act as proton donors or acceptors during the enzymatic reaction. If carbene is indeed involved in the reaction, there is a high possibility that lysine is the key residue that contributes to the stabilization of the carbene intermediate. Moreover, Mayer et. al.[5] demonstrated that the accumulation of a stable carbene in thiamin diphosphate enzyme (ThDPE) is a major resonance contributor to the deprotonated thiamin according to near-UV CD spectroscopic data. As indicated by the crystal structure[5] (see Fig. 2B), the stable carbene intermediate in ThDPE is likely the result of the strong intramoleular interaction between the $-NH_2$ group and the carbene in thiamin. Bearing these previous studies in mind, we employ cluster models to probe the effects of lysine side chain on the stability of carbenes, including the thiazole carbene, orotidine carbene and imdidazole carbene, as shown in Fig. 1F and Fig. 2; to be consistent with the putative mechanisms, the lysine here is taken to be deprotonated (see below).

As shown in Fig. 1F, the calculated ΔE^{inter} at the CCSD(T)/CBS level of theory for lysine-carbene and lysine-imidazole are −5.4 kcal/mol and −1.4 kcal/mol, respectively. The difference is about 4 kcal/mol, larger than the effects of backbone amide groups (~1 kcal/mol), though the lysine-carbene interaction (−5.4 kcal/mol) is similar to the backbone-carbene interaction (−6.2 kcal/mol).

For thiazole carbene, as shown in Fig. 2B, the calculated interaction energy (ΔE^{inter}) with methylamine is −4.6 kcal/mol and −3.4 kcal/mol, respectively, at the MP2/aQz and SCS-MP2/aQz levels of theory; the corresponding ΔE^{inter} for thiazole and methylamine is −5.1 kcal/mol and −4.1 kcal/mol, respectively. The similar interaction energies suggest that the neutral Lys can stabilize the thiazole carbene but not more than the reactant state. Similarly, for orotidine carbene, the calculated interaction energy (ΔE^{inter}) at the MP2/aQz and SCS-MP2/aQz levels of theory for Lys-carbene is −5.5 and −4.5 kcal/mol, respectively. The corresponding ΔE^{inter} values for Lys-orotidine are −6.4 kcal/mol and −4.7 kcal/mol, indicating that the interaction between Lys and orotidine carbene is slightly weaker than Lys and orotidine. Therefore, the neutral Lys side chain is expected to offer limited stabilization of carbene intermediates. One may wonder if the positively charged Lys side chain may better interact with the carbene intermediate. In our QM cluster model, at the MP2/aTz level of theory, the charged Lys side chain ($-NH_3^+$) spontaneously transfers its proton to the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

carbene C atom during optimization for all studied carbene cases (imidazole, thiazole and orotidine), suggesting that carbene does not co-exit with a positively charged Lys side chain.

3.2 Nature of intermolecular interactions involving carbene

To provide guidance into design strategies that potentially stabilize carbene relative to enzyme substrates, it is valuable to better understand the physical nature of intermolecular interactions involving carbene; such understanding also helps inform the development of force field models appropriate for describing interactions involving carbene in computational design. Indeed, the cluster model calculations show that the magnitude of interactions between the carbene intermediate in EanB and nearby enzyme motifs can be very large. With charge neutral groups such as Tyr and F2Tyr, the interaction energy is about 13–20 kcal/mol; with a charged group like Arg, the interaction is as strong as ~26 kcal/mol. These values are substantially larger than the typical hydrogen bonding interactions involving a charge neutral species (carbene), thus it is of interest to better understand the physical origin of such interactions with an energy decomposition analysis; while there is not a unique way to decompose intermolecular interactions,[61] a self-consistent energy decomposition scheme will provide valuable insights and help develop physical intuitions that can guide the design of favorable stabilizations. We have conducted such analysis here at the MP2/aQz level of theory using the decomposition scheme of Li et al.[16] and the results are summarized in Table 2.

In terms of magnitude, electrostatic interactions tend to dominate for all cases studied here; moreover, compared to imidazole, the carbene intermediate features stronger electrostatic interactions. For example, ΔE^{ele} for F2Tyr is −34.5 kcal/mol and −21.4 kcal/mol, with carbene and imidazole, respectively; with Tyr, the corresponding values are −25.3 and −19.0 kcal/mol. The stronger electrostatic interactions are likely correlated with the presence of a lone pair on the carbene carbon, enabling it to be a hydrogen-bond acceptor. On the other hand, the lone pair is also expected to experience stronger exchange repulsion with the hydrogen bonding partner. Indeed, ΔE^{ex-rep} is substantially more positive for carbene than for the imidazole substrate. For example, ΔE^{ex-rep} for F2Tyr is 43.3 and 25.5 kcal/mol, with carbene and imidazole, respectively.

Both polarization and dispersion make considerable contributions to the total interactions, again favoring carbene over imidazole in many cases. For example, ΔE^{pol} for F2Tyr is −20.2 and −10.9 kcal/mol, with carbene and imidazole, respectively; the corresponding values for ΔE^{disp} are −9.8 and −6.6 kcal/mol, respectively. These observations are consistent with the consideration that the carbene also features an empty valence orbital, suggesting that it exhibits a larger polarizability than imidazole and therefore stronger ΔE^{pol} and ΔE^{disp}.[14,62]

It is also of interest to analyze the correlation between ΔE^{inter} and its components for the various cluster models. A plot of ΔE^{ele} vs. ΔE^{inter} shows (see Fig. 4A) that they are strongly correlated with an R^2 value of 0.91, again highlighting the importance of electrostatics. By comparison, the correlations between ΔE^{inter} and other components are substantially weaker; the R^2 values are 0.78 for ΔE^{pol}, 0.62 for ΔE^{ex-rep} and 0.57 for ΔE^{disp}. Further distinguishing interactions with carbenes and with reactant substrates, the correlations between ΔE^{inter} and ΔE^{pol} are comparable (Fig. 4B), with R^2 values of 0.74 for

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

carbenes and 0.85 for reactant substrates. Larger differences are observed for the correlation between ΔE^{inter} and ΔE^{disp} (Fig. 4C); while there is hardly any correlation for the reactant substrates, the correlation is remarkably high for the carbenes with an R^2 value of 0.92. This difference highlights that intermolecular interactions with carbenes are distinct in nature as compared to typical hydrogen bonding interactions common in enzyme active sites; indeed, instead of N…H or O…H pairs, interactions with carbenes often involve C…H pairs in which the carbon has both a lone pair and an empty valence orbital. Therefore, for the parameterization of classical models to realize preferential stabilization of carbene species in enzyme design, a more sophisticated model that reflects the charge anisotropy on the carbene carbon is likely required.

Finally, we note that in the specific energy decomposition scheme used here,[16] the electron (charge) transfer interaction is not explicitly separated from polarization interactions (POL) as an independent term. The consideration is that the charge transfer contribution is likely sensitive to the size of the basis set used in the calculations. For example, in the limit of a complete basis set on two monomers, the charge transfer contribution is expected to be very small, making the polarization term equivalent to that defined in the CMO-EDA scheme used here.[63] In our study, we have included basis set superposition error correction, which makes potential charge transfer contribution small in magnitude. An indirect support along this line is the observation in Figure 4 that the correlations between ΔE_{pol} and ΔE_{inter} for carbene and substrate are similar, with R^2 values being 0.74 and 0.85, respectively; if charge transfer contributions were significant, the correlation is expected to be much weaker. Finally, we have qualitatively examined the charge distributions using natural charges.[64] As shown in Table S5 in the Supporting Information, change in the natural charge upon complex formation is modest (typically <0.03 e) for the carbene carbon. Nevertheless, the total amount of natural charge for the carbene is reduced more significantly upon complex formation as compared to the substrate; for example, for the interaction with F2Tyr, the total natural charge for carbene in the complex is +0.163 e, which is substantially larger than the value of +0.072 e for the substrate. Therefore, there is a larger but modest amount of charge transfer associated with carbene interactions when compared to standard hydrogen-bonding complexes.

3.3 Roles of a non-natural amino acid in carbene stabilization in EanB

As discussed above, the rate of deuterium exchange for the substrate εC-H was observed to be more than tenfold faster in EanB$_{Y353F2Tyr}$ than in the WT enzyme. Assuming that the deuterium exchange is limited by the carbene formation, this observation suggested that the carbene intermediate is a few kcal/mol more stable in EanB$_{Y353F2Tyr}$ than in the WT enzyme. This is puzzling since the protonation state of Tyr (or F2Tyr) remains the same in the carbene intermediate and in the reactant (see Scheme 1); the only net change is the proton exchange between εC and εN in the substrate. To gain insights into this difference, we have conducted DFTB3-D3/MM simulations for the various chemical steps in the WT EanB and EanB$_{Y353F2Tyr}$. As discussed in the Supporting Information, based on the imidazole model, the DFTB3/3OB approach systematically overestimates the stability of the carbene intermediate relative to imidazole by a significant amount. Therefore, we will not focus on the quantitative reaction energetics from the DFTB3-D3/MM simulations

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in the current discussions. However, we note from Table 1 and Table S6 that the DFTB3-D3 approach gives reliable intermolecular interactions in comparison to various DFT and *ab initio* methods, especially in terms of relative interaction energies of carbene and the imidazole substrate. Therefore, we conclude that the DFTB3-D3/MM calculations are adequate for probing the different interaction patterns of the carbene intermediate in EanB$_{Y353F2Tyr}$ and the WT enzyme.

According to the QM/MM simulations, the stabilities of the carbene intermediate in EanB$_{Y353F2Tyr}$ and EanB$_{WT}$ differ by 3.5 kcal/mol, suggesting that, in EanB$_{Y353F2Tyr}$, the carbene intermediate is indeed more stable than that in EanB$_{WT}$. A close inspection of the structures indicates that the carbene intermediate in EanB$_{Y353F2Tyr}$ is stabilized by the charged Arg417 side chain (as shown in Fig. 5A–C) as well as the phenol group of the non-natural amino acid (F2Tyr). In the WT EanB, the carbene intermediate is stabilized by Tyr353 and nearby backbone amide groups, including Arg417 and Trp416, as shown in Fig. 5D and E.

The structural differences are rather local and limited to the active site. Examination of structures from QM/MM metadynamics simulations suggests that the different positionings of Arg417 in EanB$_{Y353F2Tyr}$ and EanB$_{WT}$ are likely due to the favorable interactions between Arg417 and the fluorine atoms of F2Tyr, which orients the sidechain of Arg417 to stabilize the carbene intermediate more effectively than Tyr353 and backbone amide groups in the WT enzyme. These observations support the idea of employing well-positioned cationic residues and non-natural amino acids to stabilize carbene intermediates when designing enzymes for novel catalytic activities.

3.4 Additional discussions

For the WT EanB, the catalytic reactivity is rather low[7] with a k_{cat} of 0.68±0.01 min^{-1}. Therefore, there is significant interest in engineering the enzyme to further enhance the enzymatic activity. Although this is beyond the scope of the current work and will be explored separately, we can make brief comments based on the insights from the interaction energy calculations and QM/MM simulations of EanB. In the WT enzyme, as shown in Fig. 6, the carbene intermediate is largely stabilized by the main chain interactions between the NH groups in Gly413, Gly415 and Trp416 and the carbene carbon. As discussed above, arginine and tyrosine residues can stabilize the carbene intermediate more effectively than the backbone amide group. Therefore, positioning such sidechains in the active site to interact with the carbene carbon will be beneficial. A careful inspection of the active site suggests that a potential site is Ile262, which is located in a loop near the active site and therefore has the conformational flexibility to accommodate mutations. Replacing it with an arginine or tyrosine residue may stabilize the carbene intermediate by forming the strong carbene C…H interaction discussed above. If the mutation is done in the background of the Y353F2Tyr mutant, in which Arg417 is already displaced to stabilize the carbene intermediate (Fig. 5), mutation of Ile262 to a neutral group like Tyr is expected to be more productive to avoid electrostatic repulsion with Arg417. Since conformational rearrangements are likely to occur upon active site mutations, as illustrated by the above discussion concerning the comparison between the WT enzyme and the Y353F2Tyr mutant,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

these suggestions need to be explored explicitly with classical and QM/MM simulations, which will be reported separately along with experimental characterizations.

4 Conclusions

Considering the rich chemistry of carbenes, it is somewhat surprising that nature has not "discovered" carbene chemistry in the sense that only a very limited number of enzymes have been shown to implicate carbene intermediates. One might argue that this is because most carbene species are charge neutral and therefore it may not be straightforward to preferentially stabilize them relative to their precursors. Accordingly, transition metal ions were used to stabilize carbene species in artificial enzymes designed via the help of directed evolution.[9–11]

The current computational analysis, motivated by our recent studies of the enzyme EanB,[7,8] finds that an N-heterocyclic carbene interacts strongly with polar and positively charged protein motifs. For example, its interactions with the tyrosine and arginine sidechains are in the range of 10–16 and 26–34 kcal/mol (depending on the level of calculations), respectively, which are substantially higher than the typical hydrogen-bonding interactions involving a charge-neutral species. Indeed, these interactions are stronger by ~2–5 kcal/mol than those calculated for the imidazole molecule, which is a model for the reactant in EanB. An energy decomposition analysis at the MP2 level of theory suggests that stronger polarization and dispersion contributions are likely key to the stronger intermolecular interactions for carbenes, highlighting the importance of the lone pair and empty valence orbital on the carbene carbon.

These results suggest that it is possible to preferentially stabilize carbenes in an enzyme active site by strategically positioning specific protein groups. As an example, we compare the WT EanB and its mutant in which the key tyrosine was replaced by a non-natural analog (F2Tyr) using DFTB3/MM simulations. We observe that the carbene intermediate in the F2Tyr mutant is more stable than that in the WT enzyme by ~3.5 kcal/mol, due to structural rearrangements in the active site that enable a nearby arginine sidechain to better stabilize the carbene in the mutant. The higher degree of stability of the carbene species is qualitatively consistent with the experimental observation[8] that H/D exchange of the hercynine in the mutant is substantially faster than the WT.

In conclusion, the observations from the current work provide strong support to the notion that carbenes can be preferentially stabilized by well-positioned protein groups in an enzyme active site without resorting to transition metal ions. The incorporation of non-natural amino acids can provide further controls through either direct interactions or indirect positioning of amino acid sidechains. We hope our study helps stimulate experimental pursuit of enzyme designs that take advantage of the rich chemistry offered by carbenes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Acknowledgement

This work was supported by R01-GM106443 and R35-GM141930. Computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE[65]), which is supported by NSF grant number ACI-1548562, are greatly appreciated; part of the computational work was performed on the Shared Computing Cluster, which is administered by Boston University's Research Computing Services (URL: www.bu.edu/tech/support/research/).

References

(1). Hahn FE Introduction: Carbene Chemistry. Chem. Rev 2018, 118, 9455–9456. [PubMed: 30301355]

(2). Hopkinson MN; Richter C; Schedler M; Glorius F An overview of N-heterocyclic carbenes. Nature 2014, 510, 485–496. [PubMed: 24965649]

(3). De Fremont P; Marion N; Nolan SP Carbenes: Synthesis, properties, and organometallic chemistry. Coord. Chem. Rev 2009, 253, 862–892.

(4). Moerdyk JP; Schilter D; Bielawski CWN, N-diamidocarbenes: isolable divalent carbons with bona fide carbene reactivity. Acc. Chem. Res 2016, 49, 1458–1468. [PubMed: 27409520]

(5). Meyer D; Neumann P; Ficner R; Tittmann K Observation of a stable carbene at the active site of a thiamin enzyme. Nat. Chem. Biol 2013, 9, 488–490. [PubMed: 23748673]

(6). Lee JK; Houk K A proficient enzyme revisited: the predicted mechanism for orotidine monophosphate decarboxylase. Science 1997, 276, 942–945. [PubMed: 9139656]

(7). Cheng R; Wu L; Lai R; Peng C; Naowarojna N; Hu W; Li X; Whelan SA; Lee N; Lopez J; Zhao C; Yong Y; Xue J; Jiang X; Grinstaff MW; Deng Z; Chen J; Cui Q; Zhou J; Liu P Single-Step Replacement of an Unreactive C–H Bond by a C–S Bond Using Polysulfide as the Direct Sulfur Source in the Anaerobic Ergothioneine Biosynthesis. ACS Cata. 2020, 10, 8981–8994.

(8). Cheng R; Lai R; Peng C; Lopez J; Li Z; Naowarojna N; Li K; Wong C; Lee N; Whelan SA; Qiao L; Grinstaff MW; Wang J; Cui Q; Liu P Implications for an Imidazole-2-yl Carbene Intermediate in the Rhodanase-Catalyzed C–S Bond Formation Reaction of Anaerobic Ergothioneine Biosynthesis. ACS Cata. 2021, 11, 3319–3334.

(9). Yang Y; Arnold FH Navigating the unnatural reaction space: directed evolution of heme proteins for selective carbene and nitrene transfer. Acc. Chem. Res 2021, 54, 1209–1225. [PubMed: 33491448]

(10). Lewis RD; Garcia-Borràs M; Chalkley MJ; Buller AR; Houk K; Kan SJ; Arnold FH Catalytic iron-carbene intermediate revealed in a cytochrome c carbene transferase. Proc. Natl. Acad. Sci. U.S.A 2018, 115, 7308–7313. [PubMed: 29946033]

(11). Garcia-Borràs M; Kan SJ; Lewis RD; Tang A; Jimenez-Osés G; Arnold FH; Houk KN Origin and Control of Chemoselectivity in Cytochrome c Catalyzed Carbene Transfer into Si–H and N–H bonds. J. Am. Chem. Soc 2021, 143, 7114–7123. [PubMed: 33909977]

(12). Köppe R; Schnöckel H The boron–boron triple bond? A thermodynamic and force field based interpretation of the N-heterocyclic carbene (NHC) stabilization procedure. Chem. Sci 2015, 6, 1199–1205. [PubMed: 29560205]

(13). Stoppelman JP; McDaniel JG Physics-based, neural network force fields for reactive molecular dynamics: Investigation of carbene formation from [EMIM+][OAc-]. J. Chem. Phys 2021, 155, 104112. [PubMed: 34525833]

(14). Hollóczki O Unveiling the peculiar hydrogen bonding behavior of solvated N-heterocyclic carbenes. Phys. Chem. Chem. Phys 2016, 18, 126–140. [PubMed: 26592182]

(15). Gehrke S; Hollóczki O A molecular mechanical model for N-heterocyclic carbenes. Phys. Chem. Chem. Phys 2016, 18, 22070–22080. [PubMed: 27426687]

(16). Su P; Li H Energy decomposition analysis of covalent bonds and intermolecular interactions. J. Chem. Phys 2009, 131, 014102. [PubMed: 19586091]

(17). Møller C; Plesset MS Note on an approximation treatment for many-electron systems. Phys. Rev 1934, 46, 618.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(18). Dunning TH Jr Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J. Chem. Phys 1989, 90, 1007–1023.

(19). Woon DE; Dunning TH Jr Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. J. Chem. Phys 1993, 98, 1358–1371.

(20). Dunning TH Jr; Peterson KA; Wilson AK Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited. J. Chem. Phys 2001, 114, 9244–9253.

(21). Wilson AK; Woon DE; Peterson KA; Dunning TH Jr Gaussian basis sets for use in correlated molecular calculations. IX. The atoms gallium through krypton. J. Chem. Phys 1999, 110, 7667–7676.

(22). Frisch MJ; Trucks GW; Schlegel HB; Scuseria GE; Robb MA; Cheeseman JR; Scalmani G; Barone V; Petersson GA; Nakatsuji et al., H. Gaussian 16 Revision C.01. Gaussian Inc.: Wallingford CT.

(23). Becke AD Density-functional exchange-energy approximation with correct asymptotic behavior. Phys. Rev. A 1988, 38, 3098–3100.

(24). Becke AD Density-functional thermochemistry. III. The role of exact exchange. J. Chem. Phys 1993, 98, 5648–5652.

(25). Lee C; Yang W; Parr RG Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B 1988, 37, 785–789.

(26). Yanai T; Tew DP; Handy NC A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). Chem. Phys. Lett 2004, 393, 51–57.

(27). Adamo C; Barone V Toward reliable density functional methods without adjustable parameters: The PBE0 model. J. Chem. Phys 1999, 110, 6158–6170.

(28). Henderson TM; Izmaylov AF; Scalmani G; Scuseria GE Can short-range hybrids describe long-range-dependent properties? J. Chem. Phys 2009, 131, 044108. [PubMed: 19655838]

(29). Chai J-D; Head-Gordon M Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. Phys. Chem. Chem. Phys 2008, 10, 6615–6620. [PubMed: 18989472]

(30). Haoyu SY; He X; Li SL; Truhlar DG MN15: A Kohn–Sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. Chem. Sci 2016, 7, 5032–5051. [PubMed: 30155154]

(31). Yu HS; He X; Truhlar DG MN15-L: A new local exchange-correlation functional for Kohn–Sham density functional theory with broad accuracy for atoms, molecules, and solids. J. Chem. Theory Comput 2016, 12, 1280–1293. [PubMed: 26722866]

(32). Schwilk M; Tahchieva DN; von Lilienfeld OA Large yet bounded: Spin gap ranges in carbenes. arXiv:2004.10600 2020, 10.48550/arXiv.2004.10600, Accessed 2022-10-19.

(33). Piecuch P; Kucharski SA; Kowalski K; Musia l M Efficient computer implementation of the renormalized coupled-cluster methods: the r-ccsd [t], r-ccsd (t), cr-ccsd [t], and cr-ccsd (t) approaches. Comput. Phys. Comm 2002, 149, 71–96.

(34). Bentz JL; Olson RM; Gordon MS; Schmidt MW; Kendall RA Coupled cluster algorithms for networks of shared memory parallel processors. Comput. Phys. Comm 2007, 176, 589–600.

(35). Olson RM; Bentz JL; Kendall RA; Schmidt MW; Gordon MS A novel approach to parallel coupled cluster calculations: Combining distributed and shared memory techniques for modern cluster based systems. J. Chem. Theory Comput 2007, 3, 1312–1328.

(36). Barca GMJ; Bertoni C; Carrington L; Datta D; De Silva N; Deustua JE; Fedorov DG; Gour JR; Gunina AO; Guidez E and others, Recent developments in the general atomic and molecular electronic structure system. J. Chem. Phys 2020, 152, 154102. [PubMed: 32321259]

(37). Brooks BR; III CLB; Mackerell AD; Nilsson L; Petrella RJ; Roux B; Won Y; Archontis G; Bartels C; Boresch S; Caflisch A; Caves L; Cui Q; Dinner AR; Feig M; Fischer S; Gao J; Hodoscek M; Im W; Kuczera K; Lazaridis T; Ma J; Ovchinnikov V; Paci E; Pastor RW; Post CB; Pu JZ; Schaefer M; Tidor B; Venable RM; Woodcock HL; Wu X; Yang W; York DM; Karplus M CHARMM: The Biomolecular Simulation Program. J. Comput. Chem 2009, 30, 1545–1614. [PubMed: 19444816]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(38). Halkier A; Helgaker T; Jørgensen P; Klopper W; Olsen J Basis-set convergence of the energy in molecular Hartree–Fock calculations. Chem. Phys. Lett 1999, 302, 437–446.

(39). Schmidt MW; Baldridge KK; Boatz JA; Elbert ST; Gordon MS; Jensen JH; Koseki S; Matsunaga N; Nguyen KA; Su S and others, General atomic and molecular electronic structure system. J. Comput. Chem 1993, 14, 1347–1363.

(40). Boys SF; Bernardi F The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. Mol. Phys 1970, 19, 553–566.

(41). Xantheas SS On the importance of the fragment relaxation energy terms in the estimation of the basis set superposition error correction to the intermolecular interaction energy. J. Chem. Phys 1996, 104, 8821–8824.

(42). Pettersen EF; Goddard TD; Huang CC; Couch GS; Greenblatt DM; Meng EC; Ferrin TE UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem 2004, 25, 1605–1612. [PubMed: 15264254]

(43). Im W; Berneche S; Roux B Generalized solvent boundary potential for computer simulations. J. Chem. Phys 2001, 114, 2924–2937.

(44). Im W; Lee M; Brooks CL Generalized Born model with a simple smoothing function. J. Comput. Chem 2003, 24, 1691–1702. [PubMed: 12964188]

(45). Schaefer P; Riccardi D; Cui Q Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules. J. Chem. Phys 2005, 123, Art. No. 014905.

(46). König PH; Hoffmann M; Frauenheim T; Cui Q A critical evaluation of different QM/MM frontier treatments with SCC-DFTB as the QM method. J. Phys. Chem. B 2005, 109, 9082–9095. [PubMed: 16852081]

(47). Elstner M SCC-DFTB: what is the proper degree of self-consistency? J. Phys. Chem. A 2007, 111, 5614–5621. [PubMed: 17564420]

(48). Gaus M; Cui Q; Elstner M DFTB-3rd: Extension of the self-consistent-charge density-functional tight-binding method SCC-DFTB. J. Chem. Theory Comput 2011, 7, 931–948.

(49). Cui Q; Elstner M Density Functional Tight Binding: values of semi-empirical methods in an *ab initio* era,. Phys. Chem. Chem. Phys 2014, 16, 14368–14377. [PubMed: 24850383]

(50). Gaus M; Lu X; Elstner M; Cui Q Parameterization of DFTB3/3OB for Sulfur and Phosphorus for chemical and biological applications. J. Chem. Theory Comput 2014, 10, 1518–1537. [PubMed: 24803865]

(51). Lu X; Gaus M; Elstner M; Cui Q Parameterization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications. J. Phys. Chem. B 2015, 119, 1062–1082. [PubMed: 25178644]

(52). Grimme S; Antony J; Ehrlich S; Krieg H A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J. Chem. Phys 2010, 132, 154104. [PubMed: 20423165]

(53). Grimme S; Hansen A; Brandenburg JG; Bannwarth C Dispersion-Corrected Mean-Field Electronic Structure Methods. Chem. Rev 2016, 116, 5105–5154. [PubMed: 27077966]

(54). Brooks CL III; Karplus M Solvent effects on protein motion and protein effects on solvent motion: Dynamics of the active site region of lysozyme. J. Mol. Biol 1989, 208, 159–181. [PubMed: 2769750]

(55). Laio A; Gervasio FL Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. Rep. Prog. Phys 2008, 71, 126601.

(56). Bonomi M; Branduardi D; Bussi G; Camilloni C; Provasi D; Raiteri P; Donadio D; Marinelli F; Pietrucci F; Broglia RA; Parrinello M PLUMED: A portable plugin for free-energy calculations with molecular dynamics. Comp. Phys. Comm 2009, 180, 1961–1972.

(57). Ryckaert J-P; Ciccotti G; Berendsen HJC Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys 1977, 23, 327–341.

(58). Miller BG; Wolfenden R Catalytic proficiency: the unusual case of OMP decarboxylase. Annu. Rev. Biochem 2002, 71, 847–885. [PubMed: 12045113]

(59). Gaus M; Goez A; Elstner M, Parametrization and Benchmark of DFTB3 for Organic Molecules. J. Chem. Theory Comput 2012, 9, 338–354. [PubMed: 26589037]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(60). Christensen AS; Kubar T; Cui Q; Elstner M, Semi-empirical Quantum Mechanical Methods for Non-covalent Interactions for Chemical and Biochemical Applications. Chem. Rev 2016, 116, 5301–5337. [PubMed: 27074247]

(61). Mao YZ; Loipersberger M; Horn PR; Das A; Demerdash O; Levine DS; Veccham SP; Head-Gordon T; Head-Gordon M From Intermolecular Interaction Energies and Observable Shifts to Component Contributions and Back Again: A Tale of Variational Energy Decomposition Analysis. Annu. Rev. Phys. Chem 2021, 72, 641–666. [PubMed: 33636998]

(62). Cowan JA; Clyburne JA; Davidson MG; Harris RLW; Howard JA; Küpper P; Leech MA; Richards SP On the Interaction between N-Heterocyclic Carbenes and Organic Acids: Structural Authentication of the First N-H…C Hydrogen Bond and Remarkably Short C-H…O Interactions. Angew. Chem. Int. Ed 2002, 41, 1432–1434.

(63). Ghanty TK; Staroverov VN; Koren PR; Davidson ER Is the Hydrogen Bond in Water Dimer and Ice Covalent? J. Am. Chem. Soc 2000, 122, 1210–1214.

(64). Reed AE; Weinstock RB; Weinhold F Natural population analysis. The Journal of Chemical Physics 1985, 83, 735–746.

(65). Towns J; Cockerill T; Dahan M; Foster I; Gaither K; Grimshaw A; Hazelwood V; Lathrop S; Lifka D; Peterson GD; Roskies R; Scott JR; WilkinsDiehr N XSEDE: Accelerating Scientific Discovery. Comput Sci. & Engn 2014, 16, 62–74.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

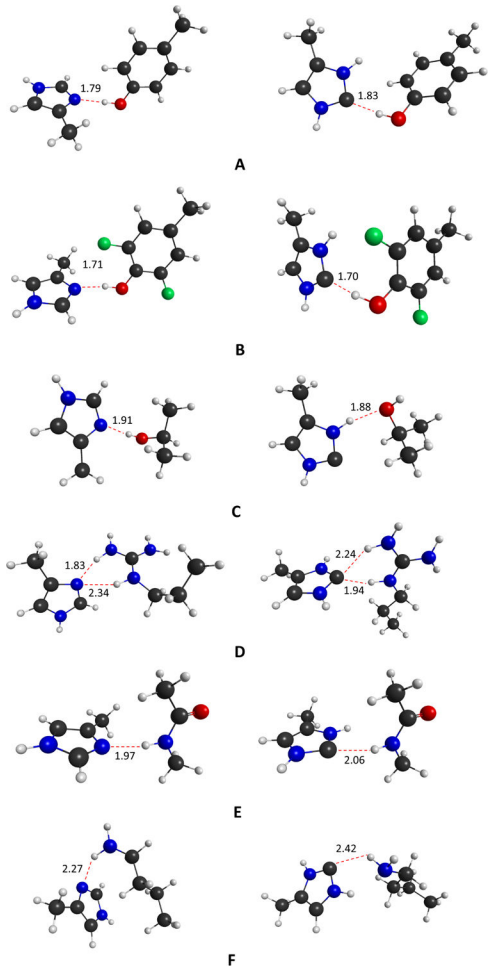


Figure 1:

Model systems representing the interaction between substrate (imidazole, left column) or carbene (right column) with protein motifs in the active site of EanB: (A) Tyrosine (Tyr); (B) Fluorine substituted tyrosine (F2Tyr); (C) Threonine (Thr); (D) Arginine (Arg); (E) Backbone; (F) Neutral Lysine (Lys). Geometries were optimized in vacuo at the MP2/aTz level of theory. The distances shown are in the unit of Å.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

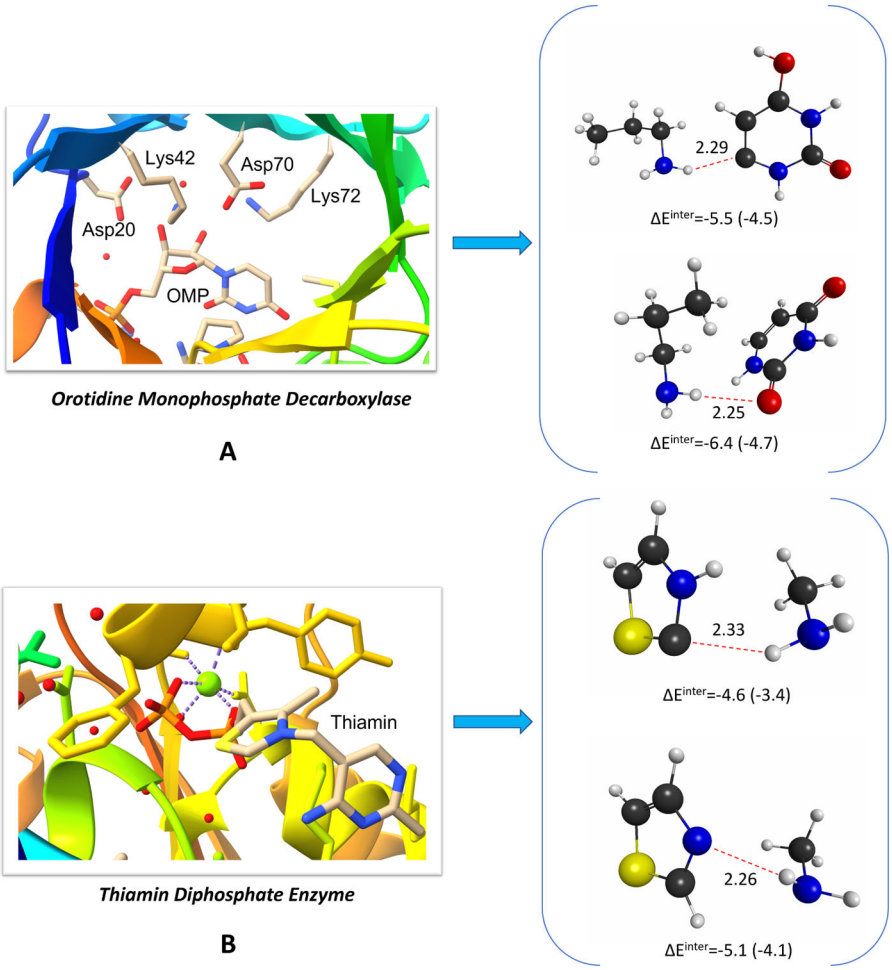


Figure 2:

The model systems used to probe carbene stabilization in the active site of A. Orotidine Monophosphate Decarboxylase, and B. Thiamin Diphosphate Enzyme. ΔE^{inter} values are the interaction energies between the two moieties calculated at the MP2/aQz and SCS-MP2/aQz (values in parentheses) levels of theory.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

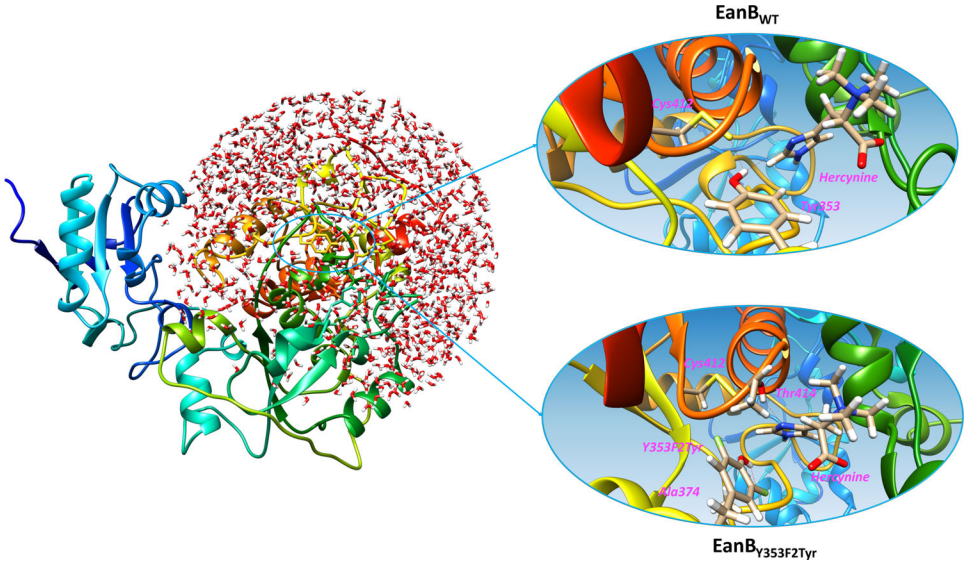


Figure 3:

The computational set up of the EanB enzyme in QM/MM-GSBP metadynamics simulations. The reactant-substrate (RS) complexes for the wild-type (EanB$_{WT}$) and Y353F2Tyr mutant (EanB$_{Y353F2Tyr}$) feature different active site interactions: F2Tyr353 forms a hydrogen bond with the backbone carbonyl group of Ala374, instead of being hydrogen bonded with the persulfide and hercynine substrate as observed for Tyr353 in the wild-type enzyme. Figures are generated using Chimera.[42]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

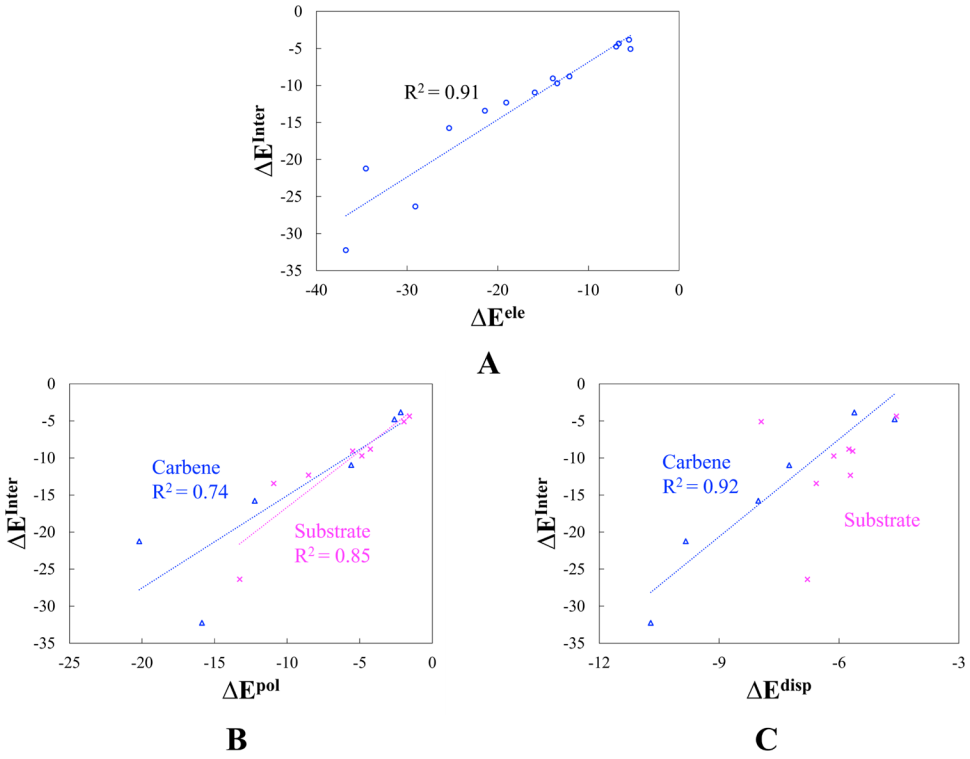


Figure 4:

The relation between total interaction energy and energy components obtained by the canonical molecular orbital energy decomposition analysis (CMO-EDA) at the MP2/aQz level of theory with basis set superposition error (BSSE) correction for model systems shown in Figs. 1 and 2: A. the total interaction energy (ΔE^{inter}) is linearly correlated with the electrostatic energy (ΔE^{ele}); B. correlation between polarization energy (ΔE^{pol}) and ΔE^{inter} for carbenes (in blue) is less linear than that for the substrate molecules (imidazole, thiazole or orotidine, in magenta); C. correlation between dispersion energy (ΔE^{disp}) and ΔE^{inter} for carbenes (in blue) is close to be linear; for the substrate molecules (imidazole, thiazole or orotidine, in magenta), the correlation is much weaker.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

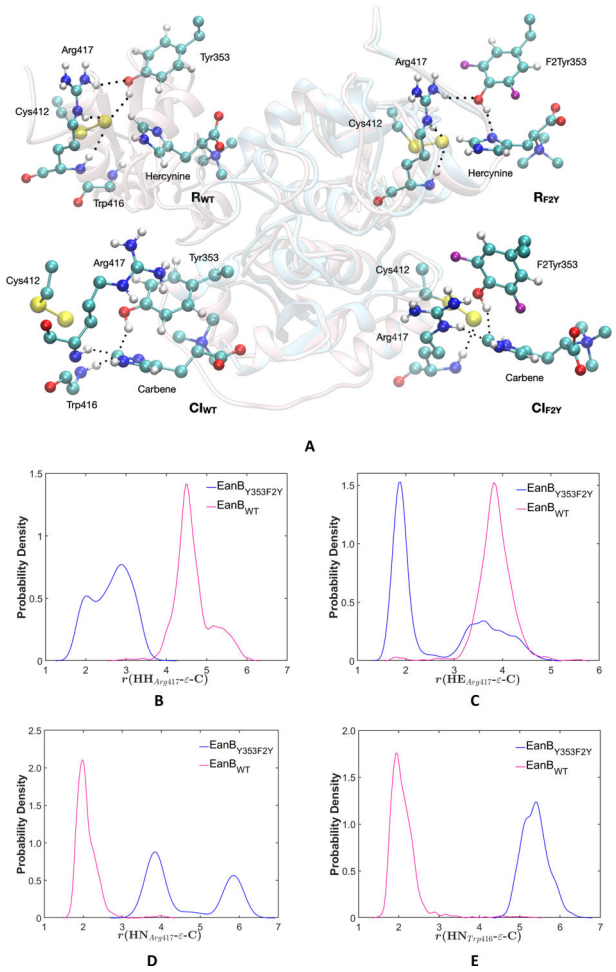


Figure 5:

Structural differences between the WT and Y353F2Tyr mutant EanB enzymes from QM/MM simulations. (A) Snapshots that illustrate the comparison of reactant states (R_{WT} vs. R_{F2Y}) and the carbene intermediates (CI_{WT} vs. CI_{F2Y}): though the reactant states are reasonably similar, the carbene intermediates differ in terms of key active site interactions, especially those involving Arg417 and (F2)Tyr353. (B-E) Probability density distributions for distances between the carbene intermediate (ε-C atom) and nearby hydrogen atoms during QM/MM metadynamics simulations: (B) Arg417 side chain HH_{Arg417} and (C) HE_{Arg417}; (D) backbone −NH of Arg417; (E) backbone −NH of Trp416. Magenta: $EanB_{WT}$; Blue: $EanB_{Y353F2Tyr}$.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

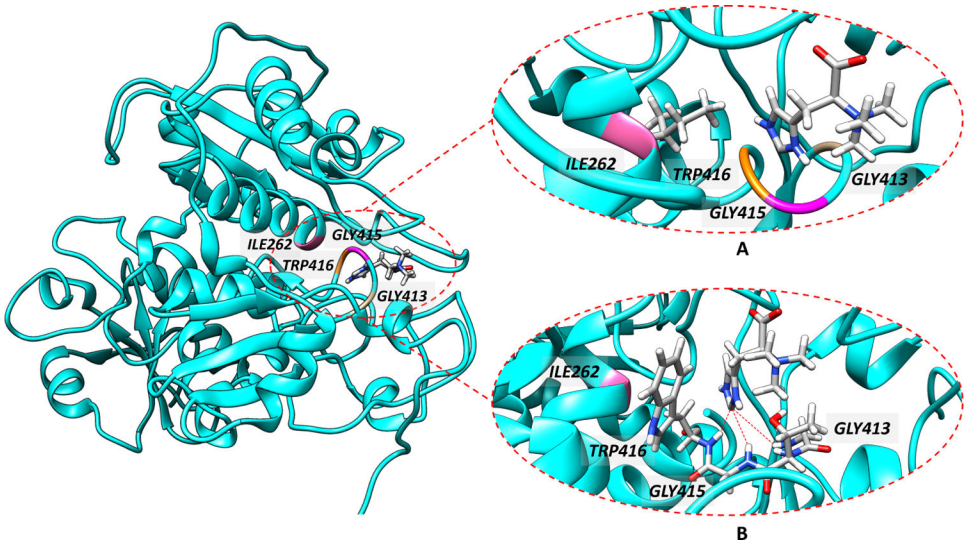
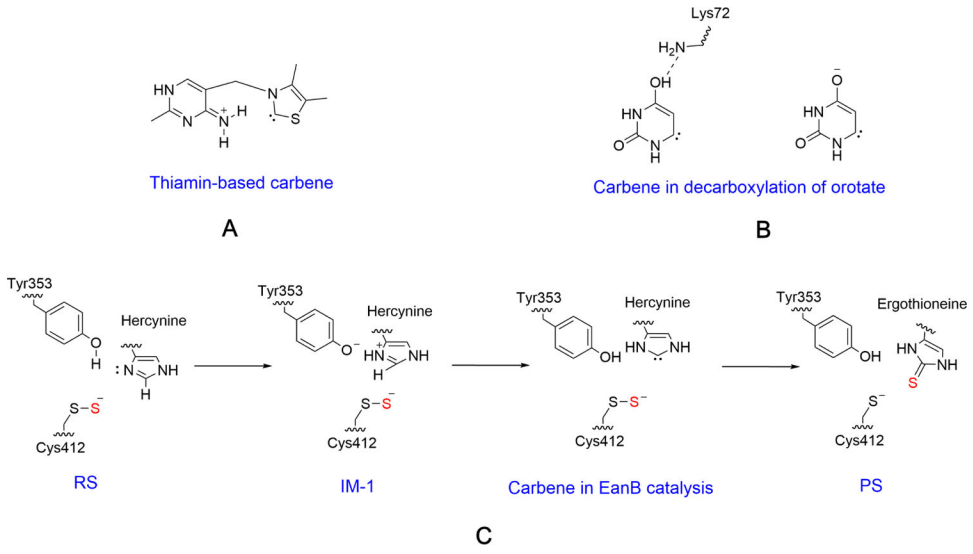


Figure 6:

Inspection of the WT EanB enzyme suggests that Ile262 is a site of potential interest for further engineering studies aimed at enhancing the catalytic activity of EanB. Ile262 is located on a loop near the active site and therefore likely has the conformational flexibility to accommodate mutations into Arginine or Tyrosine, which can better stabilize the carbene intermediate compared to the main-chain interactions from Gly413, Gly415 and Trp416.

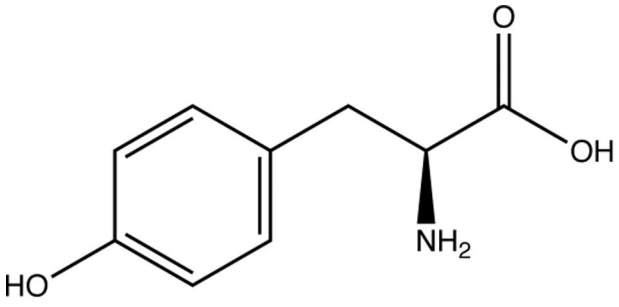
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



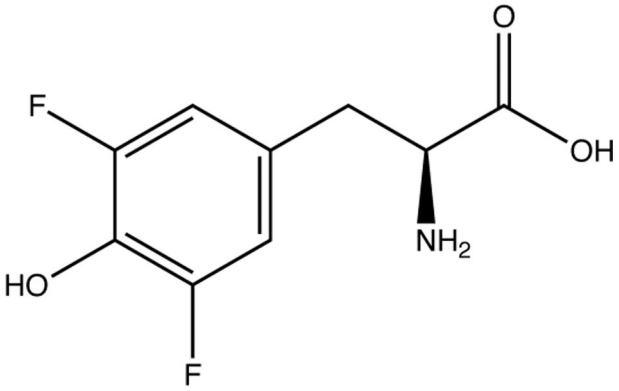
Scheme 1:

The carbene intermediate proposed in natural enzymes: A. Thiamin-based carbene in pyruvate oxidase.[5] B. Proposed carbene intermediate in Orotidine monophosphate decarboxylase.[6] C. The proposed mechanism for EanB,[7,8] where a carbene intermediate is involved in the catalytic reaction. RS: reactant; IM-1: first intermediate with a deprotonated tyrosine (Tyr353); PS: product state with ergothininene formation.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Tyrosine



3, 5-difluoro tyrosine

Scheme 2:

The structure of tyrosine and 3,5-difluoro tyrosine (F2Tyr). F2Tyr has a sidechain pK_a about 2.8 units lower than that of tyrosine in water.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1:

Comparison of the interaction energies for substrate(Subs)/carbene(Carb) with residues in the active site of EanB enzyme at different QM level of theories.[a]

Methods	Arg			Backbone			Tyr			F2Tyr			Thr		
	Carb	Subs	Δ	Carb	Subs	Δ	Carb	Subs	Δ	Carb	Subs	Δ	Carb	Subs	Δ
DFTB3[b]	−16.9	−17.1	−0.1	−2.4	−4.4	−2.0	−4.2	−6.1	−1.9	−8.7	−6.5	2.1	−2.6	−4.4	−1.8
DFTB3-D3[b]	−19.4	−19.7	−0.4	−5.0	−7.4	−2.4	−6.5	−8.5	−2.0	−11.1	−8.9	2.2	−4.8	−7.1	−2.3
DFTB3m[b]	−26.9	−20.4	6.5	−8.7	−7.2	1.5	−15.2	−8.7	6.5	−18.2	−8.6	9.6	−6.6	−6.4	0.2
DFTB3m-D3[b]	−29.4	−23.1	6.3	−11.2	−10.2	1.0	−17.5	−11.1	6.4	−20.6	−10.9	9.7	−8.8	−9.0	−0.3
B3LYP-D3/aQz	−31.2	−26.3	4.9	−10.3	−9.7	0.6	−15.5	−12.4	3.1	−20.8	−13.6	7.3	−8.3	−9.2	−0.9
PBE0-D3/aQz	−32.4	−26.7	5.6	−10.9	−9.8	1.1	−16.4	−12.7	3.6	−21.7	−13.9	7.8	−8.5	−9.4	−0.8
MP2/aQz	−33.5	−27.6	5.9	−12.0	−10.9	1.1	−16.9	−13.5	3.4	−22.6	−14.7	7.9	−9.7	−10.1	−0.4
MP2/CBS	−27.9	−22.9	5.0	−7.0	−5.9	1.1	−12.8	−9.7	3.0	−17.4	−10.7	6.7	−5.0	−6.4	−1.4
SCS-MP2/aQz	−31.2	−26.0	5.3	−10.3	−9.5	0.8	−15.0	−12.1	2.9	−20.1	−13.1	7.0	−8.3	−8.8	−0.5
SCS-MP2/CBS	−24.4	−20.3	4.1	−4.2	−3.3	0.9	−9.9	−7.5	2.5	−13.7	−8.2	5.5	−2.6	−4.3	−1.7
CCSD(T)/aDz	−33.0	−27.8	5.3	−12.3	−11.7	0.6	−16.3	−13.6	2.8	−21.5	−14.7	6.8	−10.3	−10.7	−0.4
CCSD(T)/CBS[c]	−26.2	−22.1	4.1	−6.2	−5.3	0.9	−11.2	−8.7	2.5	−15.4	−9.7	5.7	−4.5	−6.0	−1.5

[a] All energies are in the unit of kcal/mol. For the results of additional QM methods, see Supporting Information.

[b] The 3OB-3–1 parameter set[59] was used for DFTB3 and DFTB3-D3 calculations; DFTB3m indicates that the modified C-H repulsive potential developed in Ref. 7 is used. The D3 is the empirical dispersion model developed for DFTB3.[53,60]

[c] The complete basis set limits were estimated by using exponential extrapolations obtained at the MP2 level of theory.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2:

Energy decomposition analysis of intermolecular interaction energies between carbene/substrate with possible interacting residues in enzyme active sites at the MP2/aQz level of theory with basis set superposition error (BSSE) correction.[a]

	ΔE^{e1e}	$\Delta E^{ex\text{-}rep}$	ΔE^{pol}	ΔE^{disp}	ΔE^{inter}
Tyr-Carbene	−25.3	29.8	−12.2	−8.0	−15.8
Tyr-Substrate	−19.0	20.9	−8.5	−5.7	−12.3
F2Tyr-Carbene	−34.5	43.3	−20.2	−9.8	−21.2
F2Tyr-Substrate	−21.4	25.5	−10.9	−6.6	−13.4
Thr-Carbene	−12.1	13.3	−4.3	−5.8	−8.8
Thr-Substrate	−13.9	16.0	−5.5	−5.7	−9.1
Arg-Carbene	−36.7	31.1	−15.9	−10.7	−32.3
Arg-Substrate	−29.1	22.8	−13.3	−6.8	−26.4
Backbone-Carbene	−15.9	17.8	−5.6	−7.3	−11.0
Backbone-Substrate	−13.4	14.7	−4.9	−6.1	−9.7
Lys-Carbene	−15.1	17.2	−4.9	−7.1	−10.0
Lys-Substrate	−7.8	13.1	−2.4	−10.2	−7.3
Lys-OMP$_{Carbene}$	−6.9	9.4	−2.6	−4.6	−4.8
Lys-OMP$_{Substrate}$	−5.3	10.2	−2.0	−8.0	−5.1
Lys-Thiamin$_{Carbene}$	−5.5	9.5	−2.2	−5.6	−3.8
Lys-Thiamin$_{Substrate}$	−6.6	8.4	−1.6	−4.6	−4.4

[a] All energies are in the unit of kcal/mol. As indicated in Figs.1–2 and discussed in the text, the Lys sidechain is taken to be charge neutral in these calculations.