Eliminating imaginary vibrational frequencies in quantum-chemical cluster models of enzymatic active sites

Paige E. Bowling,[1,2] Saswata Dasgupta,[2,3] and John M. Herbert[1,2*]
[1] *Biophysics Graduate Program, The Ohio State University, Columbus, Ohio 43210 USA*
[2] *Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210 USA*
[3] *Department of Chemistry and Biochemistry, University of California–San Diego, La Jolla, California 92093 USA*

Abstract

In constructing finite models of enzyme active sites for use in quantum-chemical calculations, atoms at the periphery of the model system are often constrained to prevent structural collapse during geometry relaxation. A simple fixed-atom or "coordinate lock" approach is commonly employed but leads to undesirable artifacts including the appearance of small imaginary frequencies. These preclude the evaluation of finite-temperature free energy corrections, limiting thermochemical calculations to enthalpies only. Full-dimensional vibrational frequency calculations are possible by replacing the fixed-atom constraints with harmonic confining potentials, and here we compare that approach to an alternative strategy in which fixed-atom contributions to the Hessian are simply omitted. While that approach does eliminate imaginary frequencies, it tends to underestimate both the zero-point energy and the vibrational entropy, in addition to the artificial rigidity already introduced by fixed-atom constraints. Harmonic confining potentials eliminate imaginary frequencies and provide a flexible means to construct models that can be used in unconstrained geometry relaxations.

1 Introduction

Efforts to unravel the principles of enzymatic catalysis have important practical implications. For instance, comprehensive insights into enzyme reaction mechanisms are invaluable in the strategic development of pharmaceuticals.[1] This involves developing inhibitors that imitate the structure of either the intermediate stages or the transition states of enzymatic reactions.[2] Biocatalysis has furthermore been revolutionized by enzyme engineering techniques,[3,4] and mechanistic information plays a crucial role in guiding that process. While classical molecular dynamics (MD) simulations play a central role in identifying binding sites and key enzyme–substrate interactions, and can inform site-directed mutagenesis studies designed to optimize enzymatic efficiency,[5] MD force fields cannot provide mechanistic information for enzyme-catalyzed reactions, for which a quantum-chemical approach is required.

The most common way to apply electronic structure theory to study enzymatic reactions is to use a hybrid quantum mechanics/molecular mechanics (QM/MM) formalism, coupled to all-atom MD simulations.[6–9] Setup of QM/MM calculations requires considerable care,[9–11] and an underappreciated aspect is just how slowly thermochemical predictions converge with respect to the size of the QM region, typically requiring hundreds of QM atoms.[12–24] Some progress has been made toward automated selection of QM model regions.[22–26]

An alternative to QM/MM-MD simulations to study mechanistic aspects of enzyme catalysis is to use limited "cluster" models of the active site.[27–31] This approach neglects any atomistic description of the larger protein environment, and is thus unable to describe chemical transformations that are driven by conformational changes of the protein, but for a limited set of problems the QM-cluster approach has an important advantage of simplicity. By eliminating the need for MD sampling (requiring a QM energy and gradient evaluation every 1–2 fs), the cluster approach becomes amenable to higher-level electronic structure calculations.[32,33]

That said, QM-cluster modeling faces its own challenges insofar as one must carefully select a model to mimic the active-site structure. Starting from a protein crystal structure, QM geometry relaxation is required in order to obtain bond lengths and bond angles that are consistent with the chosen level of electronic structure theory, but unconstrained relaxation in the absence of an extended protein scaffold typically results in a structural collapse. To avoid this, the C_α carbons at the periphery of the model system (where the crystal structure is truncated) are typically fixed in space, in what has been called a "coordinate-lock" approach.[28] We refer to these as fixed-atom constraints,[34] and their use may engender artificial rigidity in small model systems. In principle, this problem ought to become less severe in larger model systems, yet larger models are susceptible to the emergence of multiple minima.[35,36] This obviates some

*herbert@chemistry.ohio-state.edu

of the advantage of using an optimized geometry of a finite model system in lieu of MD simulations that might sample all the low-lying minima.

Use of fixed-atom constraints invariably results in a large number of imaginary vibrational frequencies upon structure relaxation.[28] These have been noted in numerous publications,[37–60] yet a satisfactory explanation has never been provided. The present work demonstrates unequivocally that imaginary frequencies arise from a discrepancy between the curvature of the potential energy surface and the Hessian that is diagonalized to obtain harmonic frequencies. Their magnitude is typically small, being variously reported (for the modes exhibiting negative curvature) as $|\nu| \sim 10$ cm^{-1},[37,38] $|\nu| \lesssim 20$ cm^{-1},[39–42] $|\nu| \lesssim 30$ cm^{-1},[43–46] $|\nu| \lesssim 40$ cm^{-1},[47–52] $|\nu| \leq 50$ cm^{-1},[55–58] $|\nu| \leq 60$ cm^{-1},[53] or $|\nu| < 100$ cm^{-1}.[59]

When the imaginary frequencies are small, a case can be made to obtain zero-point energy (ZPE) corrections to reaction enthalpies and barrier heights simply by ignoring them, including only the real-valued frequencies in the ZPE calculation. This may be a workable strategy to obtain $T = 0$ K thermochemistry and kinetics, but the harmonic partition function affording the finite-temperature vibrational entropy diverges in the presence of imaginary frequencies. Its numerical value is also sensitive to low-frequency modes; for example, a single 30 cm^{-1} frequency contributes 1.7 kcal/mol to ΔG°_{298}.[61] Alternatives to the harmonic approximation have been suggested as better approximations for the entropy associated with low-frequency motion, yet these alternatives still require harmonic frequencies as inputs.[62] In the context of quantum-chemical cluster modeling of enzymatic reactions, it has been argued that entropic effects are generally small (meaning $\Delta H^\circ \approx \Delta G^\circ$),[28] yet this assumption cannot be systematically tested so long as cluster models are beset by imaginary frequencies.

To eliminate imaginary frequencies, one approach is to set to zero each matrix element of the Hessian that is associated with a fixed atom. This method sidesteps imaginary frequencies but does so by reducing the number of nonzero eigenvalues of the Hessian, to a total of

$$N_{\text{vib}} = 3(N_{\text{atoms}} - N_{\text{fix}}) - 6 \tag{1}$$

for a system with N_{atoms} atoms in which N_{fix} of them are subject to fixed-atom constraints. As a result of this reduction in dimensionality, such a calculation cannot be expected to afford an accurate estimate of the vibrational entropy. In a previous study,[34] we demonstrated that the use of soft harmonic confining potentials for anchor atoms achieves the same objective as immobilizing those atoms but without producing imaginary frequencies in the first place. As such, no "zeroing-out" of the Hessian is required, and unconstrained geometry optimization algorithms can be used. In the present study, we compare the harmonic-confiner method to the aforementioned technique of zeroing out the Hessian matrix elements associated with fixed atoms. We assess the impact of either approach on the shape of potential energy surfaces for a variety of peptide and protein model systems.

2 Computational Methods

Fixed anchor-atom constraints can be implemented simply be setting to zero those components of the gradient vector ($\mathbf{g}$) that are associated with the anchor atoms,

$$g_i = \begin{cases} 0, & \text{if } x_i \text{ is constrained} \\ dE/dx_i & \text{otherwise} \end{cases} . \tag{2}$$

Note that $\mathbf{g}$ is computed in Cartesian coordinates before being transformed to any other coordinate system that might be used for geometry optimization. When vibrational frequency calculations are performed on a geometry that has been optimized using fixed-atom constraints, the result is often numerous (up to $3N_{\text{fix}}$) imaginary frequencies, due to an inconsistent treatment of the gradient and the Hessian, if no adjustments are made to the latter. In our previous work,[34] fixed-atom vibrational frequency calculations were performed in that way. However, problems with imaginary frequencies can be mitigated somewhat by means of a makeshift approach based on a Hessian matrix ($\mathbf{H}$) whose elements are

$$H_{ij} = \begin{cases} 0, & \text{if } x_i \text{ and } x_j \text{ are unconstrained} \\ d^2E/dx_i dx_j & \text{otherwise} \end{cases} . \tag{3}$$

This zeroed-out Hessian technique (as we will call it herein) typically avoids imaginary frequencies, but in a sense only hides the problem by removing the contributions of the constrained atoms to the normal mode analysis. The number of meaningful frequencies is reduced to N_{vib} in eq. 1

As an alternative to the coordinate-lock or fixed-atom approach, the use of harmonic confining potentials enables each restrained atom to oscillate around its anchor position, more accurately describing its vibrational motion in relation to the rest of the protein backbone. In this harmonic-confiner model,[34] an additional classical energy term is introduced, namely

$$V_{\text{conf}}(\mathbf{r}_1, \mathbf{r}_2, \ldots) = \frac{1}{2} \sum_{i=1}^{N_{\text{fix}}} k_i \|\mathbf{r}_i - \mathbf{r}_i^0\|^2 \tag{4}$$

where $\mathbf{r}_i$ indicates the Cartesian coordinates of the ith restrained atom, whose anchor position is $\mathbf{r}_i^0$. The same force constant is used for each restrained atom, and its value ($k = 450$ N/m) reflects a typical C–C single bond.[63] In the present work, all restrained atoms are C_α carbons of the peptide backbone. Analytic derivatives of V_{conf} are incorporated into the gradient and Hessian calculations, and no gradient or Hessian elements are set to zero or otherwise manipulated within this approach.

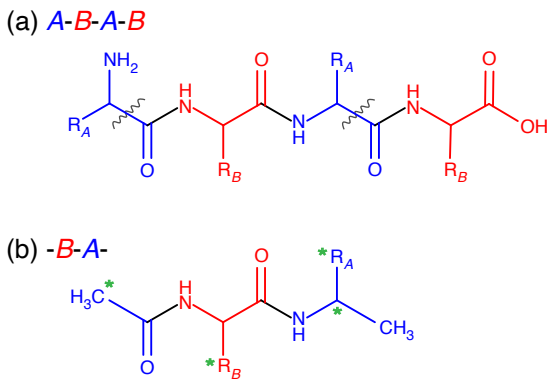

Fig. 1: (a) Tetrapeptide $ABAB$, optimized in the presence of constraints $\phi = 180^\circ = \psi$, and (b) its central dipeptide BA. Side chains are denoted with generic labels R_A and R_B for residues A and B. Green asterisks in (b) mark the locations of anchor atoms.

Electronic structure calculations were performed at the B3LYP+D3(BJ)/6-31G(d,p) level, where D3 indicates Grimme's empirical dispersion correction,[64] evaluated using the Becke-Johnson (BJ) damping function.[65,66] The SG-1 quadrature grid is used for the exchange-correlation integration.[67] The self-consistent field (SCF) convergence criterion was set to 10^{-8} Ha, with integral and shell-pair drop tolerances set to 10^{-11} a.u.. All calculations were performed using the Q-Chem program,[68] in which we have implemented both harmonic confining potentials and the zeroed-out Hessian approach. (Both methods are available starting from Q-Chem v. 6.0.) For some calculations, the conductor-like polarizable continuum model (C-PCM) was used to incorporate dielectric boundary conditions.[69–71] For these calculations, the solute cavity was constructed from atomic spheres with radii 1.2 times larger than those in the modified Bondi set.[71–73] This cavity was discretized using the switching/Gaussian algorithm,[71,74,75] with 110 Lebedev grid points for each hydrogen and 194 Lebedev points for other atomic spheres. Note that low-dielectric boundary conditions (with $\varepsilon = 2$–4) can be critically important for converging SCF on large protein models with ionic side chains.[76,77]

3 Results and Discussion

A. Constrained Dipeptides. To investigate how constraints impact vibrational frequencies and vibrational ZPE in proteins, a data set of ten tetrapeptides was prepared, each having the sequence $ABAB$ for distinct amino acids A and B with charge-neutral side chains. Geometry optimizations were performed using vacuum boundary conditions ($\varepsilon = 1$), with dihedral ψ and ϕ angles constrained at 180°, which maintains a virtually flat backbone structure. Next, the central BA dipeptide was extracted from the optimized geometry as indicated in Fig. 1, for use in subsequent restrained-atom optimization and vibrational frequency calculations. For those calculations, we defined four anchor atoms: two C_α atoms at the points of separation from the tetrapeptide, as well as two C_β atoms on side chains R_A and R_B; see Fig. 1b. Restrained-atom optimizations and harmonic frequency calculations were performed on these ten dipeptide models using either harmonic confining potentials or, alternatively, fixed-atom constraints. In the latter case we will compare frequencies obtained from an unmodified Hessian versus those computed from the zeroed-out Hessian that is defined in eq. 3.

Table 1 summarizes the results of relaxed-constrained vibrational frequency calculations, including the total ZPE and also the number of imaginary frequencies obtained, $n_{\rm im}$. In two cases, a single imaginary frequency remains when harmonic confining potentials are used, typically associated with the methyl group on residue B and which is an artifact of the dihedral constraints that are used in these dipeptide examples; these constraints will be lifted for the realistic enzyme models that are considered below. With fixed-atom constraints and an unmodified Hessian, 10 of 12 examples exhibit $n_{\rm im} = 1$ or 2, suggesting that most of these originate from the use of anchor atoms rather than the dihedral constraints. When the zeroed-out Hessian is used at the same fixed-atom relaxed geometry, however, the number of imaginary frequencies proliferates. Defining

$$n^0_{\rm im} = n_{\rm im} - 3N_{\rm fix} \ , \qquad (5)$$

which reduces the count by the $3N_{\rm fix}$ imaginary frequencies that we expect from fixing $N_{\rm fix}$ atoms in space, we find that $n^0_{\rm im} = 0$ or 1 for the zeroed-out Hessian technique. This suggests that one should expect $3N_{\rm fix}$ imaginary frequencies, in a fixed-atom calculation with $N_{\rm fix}$ anchor atoms, when the corresponding matrix elements of $\mathbf{H}$ are set to zero.

ZPE corrections in Table 1, which are computed using only the real-valued frequencies, are consistently lower for the zeroed-out Hessian than they are for the unmodified Hessian with fixed-atom constraints, by 8 kcal/mol on average. This is a direct result of the larger value of $n_{\rm im}$ for the zeroed-out Hessian and suggests that this approach likely underestimates the true ZPE. Using harmonic constraints, the ZPE is larger than it is for either fixed-atom calculation. The difference is greatest for the largest dipeptide considered, and in Fig. S1 we demonstrate that the difference scales approximately linearly with the number of anchor atoms, using several of the model systems that are considered below.

These dipeptide calculations emphasize shortcomings of the conventional fixed-atom method for vibrational frequency analysis. Although not explored in the present work, one might expect similar issues to arise in normal mode analysis based on fixed-atom constraints, as used for vibrational spectroscopy simulations of metalloenzymes.[78–80] The focus here is on calcula-

Table 1: Results of Relaxed-Constrained Vibrational Frequency Calculations on Dipeptides.

Peptide	Fixed-Atom					Harmonic Confiner	
	unmodified Hessian		zeroed-out Hessian				
	ZPE[a]	n_{im}	ZPE[a]	n_{im}	n^0_{im}	ZPE[a]	n_{im}
Ala–Arg	225.0	1	218.0	12	0	231.8	0
Asn–Asp	180.7	2	174.0	12	0	188.2	0
Cys–Glu	180.6	2	172.8	13	1	187.6	0
Gly–Gln	169.9	2	163.8	12	0	177.4	0
His–Ile	239.5	1	231.5	12	0	246.3	1
Leu–Lis	271.4	2	263.1	12	0	278.3	0
Met–Phe	240.6	1	232.8	12	0	248.0	1
Pro–Ser	179.4	0	170.3	13	1	185.6	0
Thr–Trp	243.6	0	234.9	12	0	250.5	0
Tyr–Val	242.6	1	234.7	12	0	245.9	0

[a]ZPE in kcal/mol, using only the real-valued frequencies. computed at the B3LYP+D3/6-31G(d,p) level.

tions of reaction barriers (to elucidate mechanistic information), and subsequent examples will explore realistic quantum-chemical cluster models of enzyme active sites from this point of view.

B. Structure Relaxation in an Enzyme Model. We next consider a realistic enzyme model and investigate the effect of harmonic restraining potentials on the shape of the potential surface, for a one-dimensional scan along a flexible torsion angle. These calculations employ a cluster model of the active site of threonyl-tRNA synthetase (ThrRS), in which *p*-biphenylalanine is found with the biphenyl moiety in a coplanar conformation (PDB: 4S03).[81] The charge-neutral cluster model is taken from Ref. 82 and which is shown in Fig. 2, highlighting the biphenyl moiety. We constructed a one-dimensional scan around the indicated dihedral angle (θ) between phenyl rings of the biphenyl moiety, relaxing other degrees of freedom at each fixed value of θ. These one-dimensional scans are plotted in Fig. 3 using either vacuum boundary conditions ($\varepsilon = 1$) or else dielectric boundary conditions with $\varepsilon = 4$, and carried out using either fixed-atom constraints or harmonic confining potentials for the anchor atoms.

A previous computational study of the same ThrRS model,[82] using B3LYP+D3/6-31G(d′) and C-PCM with $\varepsilon = 4$, reported a minimum-energy biphenyl angle $\theta = -2.5°$. Using fixed-atom constraints, and scanning in 5° increments, we obtain a minimum at $\theta = -5.0°$ using either $\varepsilon = 1$ or $\varepsilon = 4$ (see Fig. 3). Our procedure differs in minor details from that in Ref. 82, with a slightly different basis set [6-31G(d,p)] and a somewhat different solute cavity construction, but we regard our results as being entirely consistent with the previously reported minimum-energy structure. It is worth noting that neither of these basis sets is likely to afford a converged

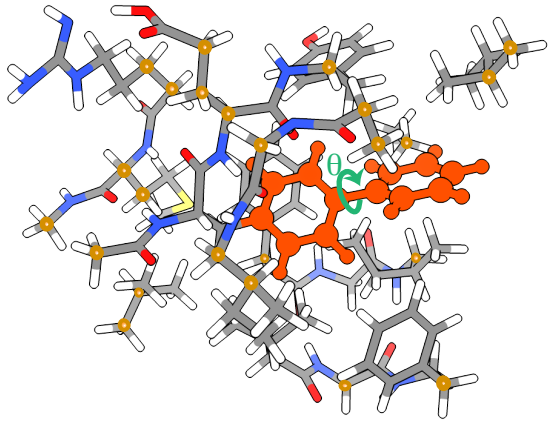

Fig. 2: Active-site model of ThrRS containing 275 atoms, from Ref. 82. The biphenyl unit is indicated in red and the green arrow indicates the torsion angle θ that is explored here. Anchor atoms are indicated by gold dots though not all of them are visible here.

energy profile,[83] but our purpose here is to make contact with previous studies that employ fixed-atom constraints.

In the present work, fixed-atom constraints result in nearly identical torsional potentials regardless of whether $\varepsilon = 1$ or $\varepsilon = 4$; see the overlay of the two potentials that is provided in Fig. S2. This system is relatively constrained, with 31 anchor atoms in a 275-atom model system, and clear differences in the dihedral potential arise when the fixed atoms are replaced with harmonic confining potentials. The latter alleviate the artificial rigidity introduced by the fixed atoms, leading to a softening of the potential. The differences are more pronounced for $\varepsilon = 4$, but even in the case of vacuum boundaries one can observe this lowering of the potential for $\theta \gtrsim 10°$. The minimum-energy structure is not affected because energy

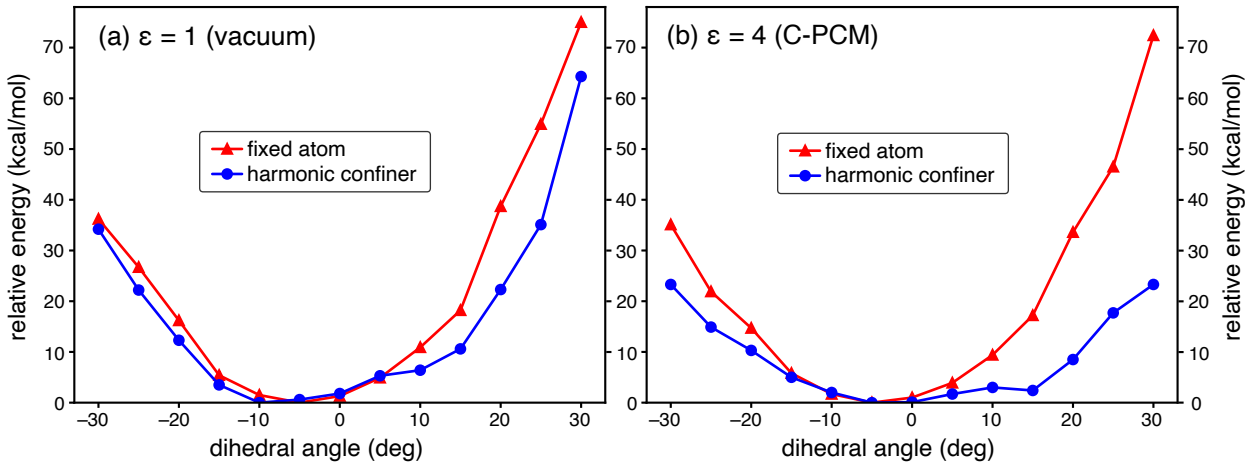

Fig. 3: Relaxed torsional scans of ThrRS around the biphenyl dihedral angle θ that is depicted in Fig. 2, using (a) vacuum boundary conditions or (b) dielectric boundary conditions with $\varepsilon = 4$. All calculations are performed at the B3LYP+D3/6-31G(d,p) level.

changes within about 5° of the minimum-energy torsion angle are no larger than ~ 0.1 kcal/mol for the calculations using harmonic confining potentials, but for larger values of θ there is significant energy lowering under dielectric ($\varepsilon = 4$) boundary conditions This suggests that the additional geometric relaxation that is possible when fixed-atom constraints are lifted leads to modest changes in the electronic structure that further polarize the environment. This may be an indication that larger model systems are required to converge thermochemical quantities when more flexible models with harmonic restraints are employed.

With that in mind, we note that it is often suggested that the the influence of dielectric boundary conditions wanes as the size of the QM-cluster model increases,[35,43,48] especially for models with $\gtrsim$ 150 atoms.[48] In our own work using sizable enzyme models,[34,77] including some with ionic side chains,[77] we observe that enthalpy changes and barrier heights computed using $\varepsilon = 2$ or $\varepsilon = 4$ are virtually indistinguishable from results obtained using much larger dielectric constants. However, results for $\varepsilon = 2$ are distinguishable from those obtained using vacuum boundary conditions ($\varepsilon = 1$). That said, ThrRS proves to be something of a counterexample, in which the added flexibility of harmonic restraints combines with the dielectric boundaries to afford a qualitatively different energy profile. These results caution against drawing blanket conclusions on the basis of fixed-atom structure relaxations, as such models may be overly constrained.

C. Reaction Energy Profiles for Enzyme Models.

We next examine several active-site models for an enzymatic reaction in which L-aspartate α-decarboxylase (AspDC) catalyzes conversion of L-aspartate to β-alanine. This is an essential process in biosynthesis of vitamin B5 that produces a precursor to 4′-phosphopantetheine and coenzyme A in bacteria.[84] Previous modeling by others explored the decarboxylation of L-aspartate using various cluster models of AspDC,[50] focusing on the C–C bond-breaking step leading to liberation of CO_2. One of these active-site models is depicted in Fig. 4, derived from the crystal structure of AspDC taken from *H. pylori* in association with iso-asparagine (PDB: 1UHE).[85] Here, we examine how various anchor-atom constraints impact the predicted barrier heights, ZPE, and vibrational entropy as the model size increases.

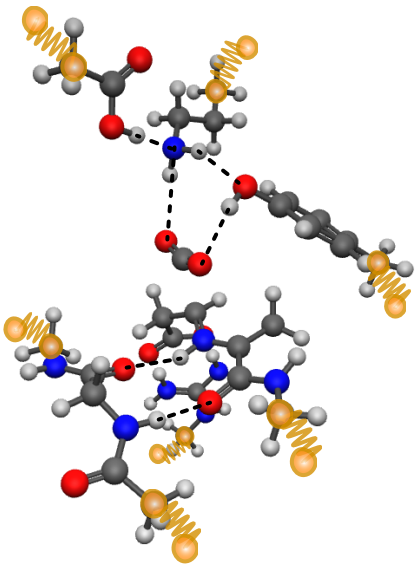

Fig. 4: Transition-state structure for model II of AspDC (from Ref. 50), leading to CO_2 release. Anchor atoms are indicated by connections to cartoon springs.

Four different model systems, designated I–IV and con-

Table 2: Barrier height ($\Delta^{\ddagger}E$) and Reaction Energy ($\Delta_{\rm rxn}E$) for Models of AspDC Decarboxylation.[a]

Model	$N_{\rm atoms}$	$N_{\rm fix}$	$\Delta^{\ddagger}E$ (kcal/mol)			$\Delta_{\rm rxn}E$ (kcal/mol)		
			Fixed-Atom		Harmonic	Fixed-Atom		Harmonic
			unmodified[b]	zeroed-out[b]	Confiner	unmodified[b]	zeroed-out[b]	Confiner
I	76	5	13.4 (−2.7)	13.6 (−2.6)	9.4 (−2.4)	8.5 (−3.1)	8.7 (−2.9)	5.4 (−2.5)
II	95	7	9.6 (−2.5)	9.3 (−2.4)	8.9 (−2.2)	−0.0 (−2.2)	0.2 (−1.9)	1.8 (−2.7)
III	135	9	8.3 (−2.5)	8.4 (−2.4)	8.2 (−1.5)	−0.2 (−1.2)	0.3 (−1.2)	3.3 (−1.3)
IV	189	14	14.2 (−4.0)	14.4 (−3.9)	10.2 (−1.7)	3.5 (−0.9)	3.4 (−1.1)	3.3 (−1.0)

[a] B3LYP+D3/6-311+G(2df,2p)//B3LYP+D3/6-31G(d,p) level with the ZPE correction included. (This correction is also indicated in parenthesis.) [b] Describes how the Hessian is treated.

taining 76–189 atoms, were taken from Ref. 50. The largest of these (model IV, which is called model IV.2 in Ref. 50) consists of 15 amino acid residues, some of which are truncated, and a crystallographic water molecule. The total number of anchor atoms ranges from $N_{\rm fix} = 5$–14 as listed in Table 2. These model structures were relaxed at the B3LYP+D3/6-31G(d,p) level, using Becke–Johnson damping for the D3 correction.[66] Only minor changes are observed relative to the structures provided in Ref. 50. These are documented for model I in Fig. S3 and consist primarily of some movement of a 4-methylphenol ring, which is a truncated model of tyrosine and which turns slightly about the fixed anchor atom in our optimizations.

Energetics are reported here at the B3LYP+D3/6-311+G(2df,2p)//B3LYP/6-31G(d,p) level, where the basis set used for the single-point energy calculations afford error statistics comparable to much larger basis sets, in benchmark tests.[83] In Ref. 50, energetics were reported at the B3LYP/6-311+G(2d,2p)//B3LYP/6-31G(d,p) level but we consider that the use of dispersion-corrected B3LYP is important for conformational energies, thermochemistry, and barrier heights.[86] In Ref. 50, vibrational frequency calculations were omitted for model IV but they are reported here.

Table 2 presents computed values for the forward barrier height

$$\Delta^{\ddagger}E = E_{\rm TS} - E_{\rm reactants} \tag{6}$$

and for the reaction energy

$$\Delta_{\rm rxn}E = E_{\rm products} - E_{\rm reactants} \tag{7}$$

for models I–IV, using both fixed-atom and harmonic restraints. All values include ZPE corrections, which are also listed separately in Table 2. In the case of fixed-atom constraints, results are compared for two different Hessians, one that is unmodified and one in which the fixed-atom contributions are deleted, as in eq. 3. Small imaginary frequencies are observed with the unmodified Hessian, which are absent when eq. 3 is used instead, although neither $\Delta^{\ddagger}E$ nor $\Delta_{\rm rxn}E$ is significantly affected by the presence or absence of these imaginary frequencies. The use of harmonic restraints affords structures that are strictly free of imaginary frequencies save for the over-the-barrier mode at the transition state.

Although models I–III appear to be approaching converged values for both $\Delta^{\ddagger}E$ and $\Delta_{\rm rxn}E$, using either harmonic or fixed-atom restraints, results for model IV are rather different when fixed-atom constraints are used. In our calculations, the ZPE-corrected barrier height increases from 8 kcal/mol (model III) to 14 kcal/mol (model IV), such that $\Delta^{\ddagger}E$(IV) is closer to $\Delta^{\ddagger}E$(I), after having decreased in the order $\Delta^{\ddagger}E$(I) > $\Delta^{\ddagger}E$(II) > $\Delta^{\ddagger}E$(III). (This is consistent with a sizable jump in the energetics between model III and model IV.2 that is reported in Ref. 50, from 9 kcal/mol to 13 kcal/mol under vacuum boundary conditions.) Although the same trend is observed when harmonic confining potentials are used, the magnitude of the effect is significantly suppressed. Similar trends are observed for $\Delta_{\rm rxn}E$. Results for model II thus appear to be converged to within < 2 kcal/mol of those for model IV, when harmonic confining potentials are employed. The same cannot be said when fixed-atom constraints are used.

Dielectric boundary conditions are omitted in these calculations. In Ref. 50, boundary conditions with $\varepsilon = 4$ increase $\Delta^{\ddagger}E$ for model IV by 4.0 kcal/mol and increase $\Delta_{\rm rxn}E$ by 5.6 kcal/mol. These corrections are similar for the smaller models, *e.g.*, 3.9 kcal/mol ($\Delta^{\ddagger}E$) and 2.7 kcal/mol ($\Delta_{\rm rxn}E$) in the case of model III, and are only reduced below 1 kcal/mol for a larger model V with 220 atoms.[50] In that case, values for both $\Delta^{\ddagger}E$ and $\Delta_{\rm rxn}E$ do converge (within 1 kcal/mol) for model V using fixed-atom constraints, but the present work suggests that convergence is more rapid when harmonic confining potentials are used instead.

Harmonic restraints also lead to more rapid convergence of the ZPE correction, which is converged even for model III. This is a useful observation because DFT analytic frequency calculations incur a cubic-scaling memory bottleneck for solution of the coupled-perturbed equations.[87–89] With harmonic restraints, it may be possible to use smaller models to estimate the thermal vibrational contribution to $\Delta_{\rm rxn}G$.

Reaction-profile diagrams for all four models are shown in Fig. 5, with vibrational free energy corrections ($-T\Delta S_{\rm vib}$) computed at $T = 298.15$ K. Entropic and ZPE effects both generally stabilize the transition state and the product state relative to the reactant state, and they do so to a much greater degree than the 0.1 kcal/mol

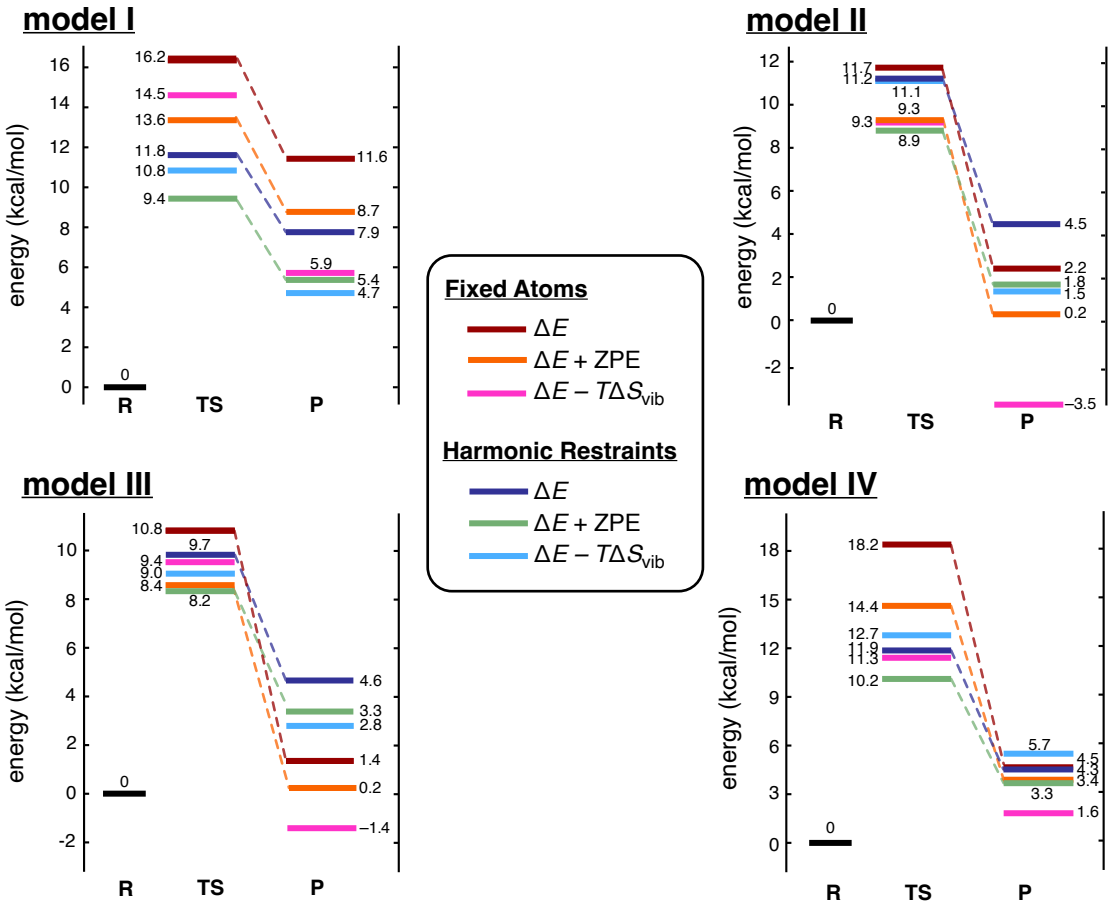

Fig. 5: Energy profiles for AspDC-catalyzed decarboxylation, computed at the B3LYP+D3/6-311+G(2df,2p)//B3LYP+D3/6-31G(d,p) level using model systems of increasing size, I–IV. All energies are relative to the reactant state (**R**). Some of the transition state (**TS**) and product state (**P**) energies, computed with the same corrections, are connected by dashed lines as guides for the eye.

correction (at $T = 298$ K) that was estimated in Ref. 50 based on a "model 0" containing just 27 atoms, which is likely to be over-constrained. A larger model size does not always equate to a larger vibrational correction, however, as one might have anticipated based on dimensionality or flexibility arguments. With harmonic restraints, the largest corrections are found for the product state in model I, although this is not true in the case of fixed-atom constraints. Numerical values for the entropy and ZPE corrections are provided in Table 3, and one can see that the fixed-atom values sometimes oscillate with model size, and may differ in sign as compared to fully harmonic values. The latter are consistent in sign and appear to converge as the model size increases. Furthermore, for this particular system there is really no justification to argue that ZPE is more important than finite-temperature contributions to the free energy at $T = 298$ K.

As model size increases from I to III, the fixed-atom models go from having the highest product energy to having the lowest, although results for model IV are similar in this respect. For the fixed-atom models, difference between transition state and product energies become more pronounced in models II and III as compared to model I but then contract significantly in model IV. In all cases the difference between transition state and products is smaller when harmonic confining potentials were used. This suggests that the harmonic restraints ef-

Table 3: Vibrational Entropy and ZPE Corrections (in kcal/mol) for Models of AspDC.[a]

Model		$-T\Delta S_{\text{vib}}$		ZPE	
		fixed	harmonic	fixed	harmonic
I	$\Delta^{\ddagger}E$	1.67	1.07	−2.56	−2.41
II	$\Delta^{\ddagger}E$	2.42	0.04	−2.39	−2.23
III	$\Delta^{\ddagger}E$	1.35	0.75	−2.40	−1.51
IV	$\Delta^{\ddagger}E$	6.92	−0.73	−3.85	−1.74
I	$\Delta_{\text{rxn}}E$	5.75	3.14	−2.89	−2.46
II	$\Delta_{\text{rxn}}E$	5.64	3.00	−1.93	−2.68
III	$\Delta_{\text{rxn}}E$	2.82	1.14	−1.19	−1.31
IV	$\Delta_{\text{rxn}}E$	2.92	−1.33	−1.07	−1.08

[a] B3LYP+D3/6-31G(d,p)

fectively mitigate rigidity, and imparting physical meaning to the anchor atoms, particularly when dealing with smaller models.

4 Conclusions

Quantum-chemical cluster models of enzyme active sites have emerged as a valuable and practical approach for understanding enzymatic reaction mechanisms. However, the requisite trimming of the active-site model can exert a significant influence on the resulting energetics, as can the details of how the model is constrained for geometry optimization and vibrational frequency calculations. Fixed-atom constraints are popular but their use is marred by the emergence of numerous imaginary vibrational frequencies. These can be eliminated via *ad hoc* deletion of Hessian matrix elements associated with the fixed (anchor) atoms, yet the rigidity that is introduced by this approach may afford barrier heights and reaction energies that differ significantly from those obtained when soft harmonic confining potentials replace the fixed-atom constraints. The latter approach eliminates imaginary frequencies while offering a convenient means to account for both ZPE and finite-temperature vibrational entropy.

The present work demonstrates that the crude approach of simply deleting Hessian matrix elements can be an effective way to eliminate imaginary vibrational frequencies in the presence of fixed-atom constraints, yet this simple approach is not without limitations. It tends to underestimate the ZPE, which can hinder the prediction of reaction thermochemistry and kinetics, convergence with respect to model size is sometimes erratic, and finally the fixed-atom technique introduces artificial rigidity that is evident in steeper torsional potentials for conformational changes within the active site. In contrast, the use of harmonic confining potentials is equally simple yet preserves the number of vibrational modes, resulting in a more realistic representation of vibrational contributions to thermochemistry and barrier heights. Application of harmonic restraints to an enzyme-catalyzed decarboxylation reaction, using various active-site models containing 76–189 atoms and necessitating numerous anchor atoms, reveals that convergence with respect to model size is generally better than what is observed using fixed-atom restraints. The simplicity of the harmonic approach suggests that this should be the default paradigm for quantum-chemical cluster modeling of enzyme active sites.

Notes

The authors declare the following competing financial interest(s): J.M.H. is part owner of Q-Chem and serves on its board of directors.

Acknowledgments

This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences under Award No. DE-SC0008550, Calculations were performed at the Ohio Supercomputer Center.[90]

References

[1] Schramm, V. L. Transition states, analogues, and drug development. *ACS Chem. Biol.* **2013,** *8,* 71–81.

[2] Turner, N. J.; Humphreys, L. *Biocatalysis in Organic Synthesis: The Retrosynthetic Approach;* Royal Society of Chemistry: London, 2018.

[3] Bornscheuer, U.; Huisman, G.; Kazlauskas, R.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the third wave of biocatalysis. *Nature* **2012,** *485,* 185–194.

[4] Li, C.; Zhang, R.; Wang, J.; Wilson, L. M.; Yan, Y. Protein engineering for improving and diversifying natural product biosynthesis. *Trends Biotechnol.* **2020,** *38,* 729–744.

[5] Childers, M. C.; Daggett, V. Insights from molecular dynamics simulations for computational protein design. *Mol. Syst. Des. Eng.* **2017,** *2,* 9–33.

[6] Senn, H. M.; Thiel, W. QM/MM methods for biological systems. *Top. Curr. Chem.* **2007,** *268,* 173–290.

[7] Groenhof, G. Introduction to QM/MM simulations. In *Biomolecular Simulations: Methods and Protocols*, Vol. 924; Monticelli, L.; Salonen, E., Eds.; Springer Science+Business Media: New York, 2013; Chapter 3, pages 43–66.

[8] Brunk, E.; Rothlisberger, U. Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited states. *Chem. Rev.* **2015,** *115,* 6217–6263.

[9] Clemente, C. M.; Capece, L.; Marti, M. A. Best practices on QM/MM simulations of biological systems. *J. Chem. Inf. Model.* **2023,** *63,* 2609–2627.

[10] Giudetti, G.; Polyakov, I.; Grigorenko, B. L.; Faraji, S.; Nemukhin, A. V.; Krylov, A. I. How reproducible are QM/MM simulations? Lessons from computational studies of the covalent inhibition of the SARS-CoV-2 mainn protease by carmofur. *J. Chem. Theory Comput.* **2022,** *18,* 5056–5067.

[11] K.-S.Csizi,; Reiher, M. Universal QM/MM approaches for general nanoscale applications. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2023,** *13,* e1656.

[12] Hu, L.; Eliasson, J.; Heimdal, J.; Ryde, U. Do quantum mechanical energies calculated for small models of protein-active sites converge? *J. Phys. Chem. A* **2009,** *113,* 11793–11800.

[13] Hu, L.; Söderhjelm, P.; Ryde, U. On the convergence of QM/MM energies. *J. Chem. Theory Comput.* **2011,** *7,* 761–777.

[14] Hu, L.; Söderhjelm, P.; Ryde, U. Accurate reaction energies in proteins obtained by combining QM/MM and large QM calculations. *J. Chem. Theory Comput.* **2013,** *9,* 640–649.

[15] Sumner, S.; Söderhjelm, P.; Ryde, U. Effect of geometry optimizations on QM-cluster and QM/MM studies of reac-

tion energies in proteins. *J. Chem. Theory Comput.* **2013,** *9,* 4205–4214.
[16] Liao, R.-Z.; Thiel, W. Comparison of QM-only and QM/MM models for the mechanism of tungsten-dependent acetylene hydratase. *J. Chem. Theory Comput.* **2012,** *8,* 3793–3803.
[17] Liao, R.-Z.; Thiel, W. Convergence in the QM-only and QM/MM modeling of enzymatic reactions: A case study for acetylene hydratase. *J. Comput. Chem.* **2013,** *34,* 2389–2397.
[18] Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martínez, T. J. How large should the QM region be in QM/MM calculations? The case of catechol *o*-methyltransferase. *J. Phys. Chem. B* **2016,** *120,* 11381–11394.
[19] Karelina, M.; Kulik, H. J. Systematic quantum mechanical region determination in QM/MM simulation. *J. Chem. Theory Comput.* **2017,** *13,* 563–576.
[20] Kulik, H. J. Large-scale QM/MM free energy simulations of enzyme catalysis reveal the influence of charge transfer. *Phys. Chem. Chem. Phys.* **2018,** *20,* 20650–20660.
[21] Yang, Z.; Mehmood, R.; Wang, M.; Qi, H. W.; Steeves, A. H.; Kulik, H. J. Revealing quantum mechanical effects in enzyme catalysis with large-scale electronic structure simulation. *React. Chem. Eng.* **2019,** *4,* 298–315.
[22] Summers, T. J.; Cheng, Q.; Palma, M. A.; Pham, D.-T.; Kelso III, D. K.; Webster, C. E.; DeYonker, N. J. Cheminformatic quantum mechanical enzyme model design: A catechol-O-methyltransferase case study. *Biophys. J.* **2021,** *120,* 3577–3587.
[23] Cheng, Q.; DeYonker, N. J. The glycine *N*-methyltransferase case study: Another challenge for QM-cluster models? *J. Phys. Chem. B* **2023,** *127,* 9282–9294.
[24] Agbaglo, D. A.; Summers, T. J.; Cheng, Q.; DeYonker, N. J. The influence of model building schemes and molecular dynamics on QM-cluster models: The chorismate mutase case study. *ChemRxiv* **2023,** DOI: 10.26434/chemrxiv-2023-fvzp6-v2.
[25] Summers, T. J.; Daniel, B. P.; Cheng, Q.; DeYonker, N. J. Quantifying inter-residue contact through interaction energies. *J. Chem. Inf. Model.* **2019,** *59,* 5034–5044.
[26] Cheng, Q.; DeYonker, N. J. A case study of the glycoside hydrolase enzyme mechanism using an automated QM-cluster model building toolkit. *Front. Chem.* **2022,** *10,* 854318.
[27] Siegbahn, P. E. M.; Himo, F. Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. *J. Biol. Inorg. Chem.* **2009,** *14,* 643–651.
[28] Siegbahn, P. E. M.; Himo, F. The quantum chemical cluster approach for modeling enzyme reactions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011,** *1,* 323–336.
[29] Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M. Quantum chemical studies of mechanisms for metalloenzymes. *Chem. Rev.* **2014,** *114,* 3601–3658.
[30] Himo, F. Recent trends in quantum chemical modeling of enzymatic reactions. *J. Am. Chem. Soc.* **2017,** *139,* 6780–6786.
[31] Sheng, X.; Himo, F. The quantum chemical cluster approach in biocatalysis. *Acc. Chem. Res.* **2023,** *56,* 938–947.
[32] Wappett, D. A.; Goerigk, L. Toward a quantum-chemical benchmark set for enzymatically catalyzed reactions: Important steps and insights. *J. Phys. Chem. A* **2019,** *123,* 7057–7074. Erratum: *J. Phys. Chem. A* **2020,** *124,* 1062.
[33] Wappett, D. A.; Goerigk, L. Exploring CPS-extrapolated DLPNO–CCSD(T_1) reference values for benchmarking DFT methods on enzymatically catalyzed reactions. *J. Phys. Chem. A* **2024,** *128,* 62–72.
[34] Dasgupta, S.; Herbert, J. M. Using atomic confining potentials for geometry optimization and vibrational frequency calculations in quantum-chemical models of enzyme active sites. *J. Phys. Chem. B* **2020,** *124,* 1137–1147.
[35] Himo, F. Quantum chemical modeling of enzyme active sites and reaction mechanisms. *Theor. Chem. Acc.* **2006,** *116,* 232–240.
[36] Lind, M. E. S.; Himo, F. Quantum chemistry as a tool in asymmetric biocatalysis: Limonene epoxide hydrolase test case. *Angew. Chem. Int. Ed. Engl.* **2013,** *52,* 4563–4567.
[37] Liao, R.-Z.; Yu, J.-G.; Raushel, F. M.; Himo, F. Theoretical investigation of the reaction mechanism of the dinuclear zinc enzyme dihydroorotase. *Chem. Eur. J.* **2008,** *14,* 4287–4292.
[38] Chen, S.-L.; Marino, T.; Fang, W.-H.; Russo, N.; Himo, F. Peptide hydrolysis by the binuclear zinc enzyme aminopeptidase from *Aeromonas proteolytica*: A density functional theory study. *J. Phys. Chem. B* **2008,** *112,* 2494–2500.
[39] Georgieva, P.; Himo, F. Density functional theory study of the reaction mechanism of the DNA repairing enzyme alkylguanine alkyltransferase. *Chem. Phys. Lett.* **2008,** *463,* 214–218.
[40] Chen, S.-L.; Fang, W.-H.; Himo, F. Reaction mechanism of the binuclear zinc enzyme glyoxalase II—A theoretical study. *J. Inorg. Biochem.* **2009,** *103,* 272–281.
[41] Georgieva, P.; Wu, Q.; McLeish, M. J.; Himo, F. The reaction mechanism of phenylethanolamine *N*-methyltransferase: A density functional theory study. *Biochim. Biophys. Acta* **2009,** *1794,* 1831–1837.
[42] Georgieva, P.; Himo, F. Quantum chemical modeling of enzymatic reactions: The case of histone lysine methyltransferase. *J. Comput. Chem.* **2010,** *31,* 1707–1714.
[43] Sevastik, R.; Himo, F. Quantum chemical modeling on enzymatic reactions: The case of 4-oxalocrotonate tautomerase. *Bioorg. Chem.* **2007,** *35,* 444–457.
[44] Liao, R.-Z.; Himo, F.; Yu, J.-G.; Liu, R.-Z. Theoretical study of the RNA hydrolysis mechanism of the dinuclear zinc enzyme RNase Z. *Eur. J. Inorg. Chem.* **2009,** *2009,* 2967–2972.
[45] Liao, R.-Z.; Yu, J.-G.; Himo, F. Reaction mechanism of the trinuclear zine enzyme phospholipase C: A density functional theory study. *J. Phys. Chem. B* **2010,** *114,* 2533–2540.
[46] Liao, R.-Z.; Yu, J.-G.; Himo, F. Phosphate mono- and diesterase activities of the trinuclear zinc enzyme nuclease P1—insights from quantum chemical calculations. *Inorg. Chem.* **2010,** *49,* 6883–6888.
[47] Liao, R.-Z.; Yu, J.-G.; Himo, F. Reaction mechanism of the dinuclear zinc enzyme N-acyl-L-homoserine lactone hydrolase: A quantum chemical study. *Inorg. Chem.* **2009,** *1442–1448,*.
[48] Liao, R.-Z.; Yu, J.-G.; Himo, F. Mechanism of tungsten-dependent acetylene hydratase from quantum chemical calculations. *Proc. Natl. Acad. Sci. USA* **2010,** *107,* 22523–22527.
[49] Liao, R.-Z.; Himo, F. Theoretical study of the chemoselectivity of tungsten-dependent acetylene hydratase. *ACS Catal.* **2011,** *1,* 937–944.
[50] Liao, R. Z.; Yu, J. G.; Himo, F. Quantum chemical mod-

eling of enzymatic reactions: The case of decarboxylation. *J. Chem. Theory Comput.* **2011,** *7,* 1494–1501.
[51] Liao, R.-Z.; Yu, J.-G.; F.Himo, Tungsten-dependent formaldehyde ferredoxin oxidoreductase: Reaction mechanism from quantum chemical calculations. *J. Inorg. Biochem.* **2011,** *105,* 927–936.
[52] Manta, B.; Raushel, F. M.; Himo, F. Reaction mechanism of zinc-dependent cytosine deaminase from *Escherichia coli*: A quantum-chemical study. *J. Phys. Chem. B* **2014,** *118,* 5644–5652.
[53] Lind, M. E. S.; Himo, F. Theoretical study of reaction mechanism and stereoselectivity of arylmalonate decarboxylase. *ACS Catal.* **2014,** *4,* 4153–4160.
[54] Cassimjee, K. E.; Manta, B.; Himo, F. A quantum chemical study of the ω-transaminase reaction mechanism. *Org. Biomol. Chem.* **2015,** *13,* 8453–8464.
[55] Liao, R.-Z.; Himo, F.; Yu, J.-G.; Liu, R.-Z. Dipeptide hydrolysis by the dinuclear zinc enzyme human renal dipeptidase: Mechanistic insights from DFT calculations. *J. Inorg. Biochem.* **2010,** *104,* 37–46.
[56] Sheng, X.; Lind, M. E. S.; Himo, F. Theoretical study of the reaction mechanism of phenolic acid decarboxylase. *FEBS J.* **2015,** *282,* 4703–4713.
[57] Planas, F.; Sheng, X.; McLeish, M. J.; Himo, F. A theoretical study of the benzoylformate decarboxylase reaction mechanism. *Front. Chem.* **2018,** *6,* 205.
[58] Prejanò, M.; Škerlová, J.; Stenmark, P.; Himo, F. Reaction mechanism of human PAICS elucidated by quantum chemical calculations. *J. Am. Chem. Soc.* **2020,** *144,* 14528–14628.
[59] Kazemi, M.; Sheng, X.; Himo, F. Origins of enantiopreference in *Mycobacterium smegmatis* acyl transferase: A computational analysis. *Chem. Eur. J.* **2019,** *25,* 11945–11954.
[60] Sheng, X.; Kazemi, M.; Ząbło-Dobrowolska, A.; Kroutil, W.; Himo, F. Mechanism of biocatalytic Friedel–Crafts acylation by acyltransferase from *Pseudomonas protegens*. *ACS Catal.* **2020,** *10,* 570–577.
[61] Jensen, J. H. Predicting accurate absolute binding energies in aqueous solution: Thermodynamic considerations for electronic structure methods. *Phys. Chem. Chem. Phys.* **2015,** *17,* 12441–12451.
[62] Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.* **2012,** *18,* 9955–9964.
[63] Cremer, D.; Wu, A.; Larsson, A.; Kraka, E. Some thoughts about bond enegies, bond lengths, and force constants. *J. Mol. Model.* **2000,** *6,* 396–412.
[64] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate *ab initio* parameterization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **2010,** *132,* 154104.
[65] Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011,** *32,* 1456–1465.
[66] Gray, M.; Herbert, J. M. Density functional theory for van der Waals complexes: Size matters. *Annu. Rep. Comput. Chem.* **2024,** (accepted for publication; preprint available at DOI: 10.26434/chemrxiv-2023-xgr3q-v3).
[67] Gill, P. M. W.; Johnson, B. G.; Pople, J. A. A standard grid for density-functional calculations. *Chem. Phys. Lett.* **1993,** *209,* 506–512.
[68] Epifanovsky, E. *et al.* Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021,** *155,* 084801.
[69] Lange, A. W.; Herbert, J. M. Symmetric versus asymmetric discretization of the integral equations in polarizable continuum solvation models. *Chem. Phys. Lett.* **2011,** *509,* 77–87.
[70] Herbert, J. M.; Lange, A. W. Polarizable continuum models for (bio)molecular electrostatics: Basic theory and recent developments for macromolecules and simulations. In *Many-Body Effects and Electrostatics in Biomolecules*; Cui, Q.; Ren, P.; Meuwly, M., Eds.; CRC Press: Boca Raton, 2016; Chapter 11, pages 363–416.
[71] Herbert, J. M. Dielectric continuum methods for quantum chemistry. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021,** *11,* e1519.
[72] Bondi, A. Van der Waals volumes and radii. *J. Phys. Chem.* **1964,** *68,* 441–451.
[73] Rowland, R. S.; Taylor, R. Intermolecular nonbonded contact distances in organic crystal structures: Comparison with distances expected from van der Waals radii. *J. Phys. Chem.* **1996,** *100,* 7384–7391.
[74] Lange, A. W.; Herbert, J. M. Polarizable continuum reaction-field solvation models affording smooth potential energy surfaces. *J. Phys. Chem. Lett.* **2010,** *1,* 556–561.
[75] Lange, A. W.; Herbert, J. M. A smooth, nonsingular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach. *J. Chem. Phys.* **2010,** *133,* 244111.
[76] Ren, F.; Liu, F. Impacts of polarizable continuum models on the SCF convergence and DFT delocalization error of large molecules. *J. Chem. Phys.* **2022,** *157,* 184106.
[77] Bowling, P. E.; Broderick, D. R.; Herbert, J. M. Fragment-based calculations of enzymatic thermochemistry require dielectric boundary conditions. *J. Phys. Chem. Lett.* **2023,** *14,* 3826–3834.
[78] Slater, J. W.; Marguet, S. C.; Cirino, S. L.; Maugeri, P. T.; Shafaat, H. S. Experimental and DFT investigations reveal the influence of the outer coordination sphere on the vibrational spectra of nickel-substituted rubredoxin, a model hydrogenase enzyme. *Inorg. Chem.* **2017,** *56,* 3926–3938.
[79] Du, W.-G. H.; Götz, A. W.; Noodleman, L. DFT Fe_{a3}–O/O–O vibrational frequency calculations over catalytic reaction cycle states in the dinuclear center of cytochrome *c* oxidase. *Inorg. Chem.* **2019,** *58,* 13933–13944.
[80] Caserta, G.; Pelmenschikov, V.; Lorent, C.; Waffo, A. F. T.; Katz, S.; Lauterbach, L.; Schoknecht, J.; Wang, H.; Yoda, Y.; Tamasaku, K.; Kaupp, M.; Hildebrandt, P.; Lenz, O.; Cramer, S. P.; Zebger, I. Hydroxy-bridged resting states of a [NiFe]-hydrogenase unraveled by cryogenic vibrational spectroscopy and DFT computations. *Chem. Sci.* **2021,** *12,* 2189–2197.
[81] Pearson, A. D.; Mills, J. H.; Song, Y.; Nastertorabi, F.; Han, G. W.; Baker, D.; Stevens, R. C.; Schultz, P. G. Trapping a transition state in a computationally designed protein bottle. *Science* **2015,** *347,* 863-867.
[82] Summers, T. J.; Cheng, Q.; DeYonker, N. J. A transition state "trapped"? QM-cluster models of engineered threonyl-tRNA synthetase. *Org. Biomol. Chem.* **2018,** *16,* 4090–4100.
[83] Gray, M.; Bowling, P. E.; Herbert, J. M. In defense of (certain) Pople-type basis sets. *ChemRxiv* **2024,** DOI: 10.26434/chemrxiv-2024-jz473.
[84] Brown, G. M.; Williamson, J. M. Biosynthesis of riboflavin, folic acid, thiamine, and pantothenic acid. *Adv.*

Enzymol. Relat. Areas Mol. Biol. **1982,** *53,* 345–381.
[85] Lee, B. I.; Suh, S. W. Crystal structure of the Schiff basse intermediate prior to decarboxylation in the catalytic cycle of aspartate α-decarboxylase. *J. Mol. Biol.* **2004,** *340,* 1–7.
[86] Wagner, J. P.; Schreiner, P. R. London dispersion in molecular chemistry—Reconsidering steric effects. *Angew. Chem. Int. Ed. Engl.* **2015,** *54,* 12274–1229.
[87] Korambath, P. P.; Kong, J.; Furlani, T. R.; Head-Gordon, M. Parallelization of analytical Hartree-Fock and density functional theory Hessian calculations. Part I: Parallelization of coupled-perturbed Hartree-Fock equations. *Mol. Phys.* **2002,** *100,* 1771–1790.
[88] Baker, J.; Wolinski, K.; Malagoli, M.; Pulay, P. Parallel implementation of Hartree–Fock and density functional theory analytical second derivatives. *Mol. Phys.* **2004,** *102,* 2475–2484.
[89] Liu, K.-Y.; Liu, J.; Herbert, J. M. Accuracy of finite-difference harmonic frequencies in density functional theory. *J. Comput. Chem.* **2017,** *38,* 1678–1684.
[90] Ohio Supercomputer Center (http://osc.edu/ark:/19495/f5s1ph73), accessed 2024-02-07.

TOC Graphic