Eliminating Imaginary Vibrational Frequencies in Quantum-Chemical Cluster Models of Enzymatic Active Sites

Paige E. Bowling, Saswata Dasgupta, and John M. Herbert*

ABSTRACT: In constructing finite models of enzyme active sites for quantum-chemical calculations, atoms at the periphery of the model must be constrained to prevent unphysical rearrangements during geometry relaxation. A simple fixed-atom or “coordinate-lock” approach is commonly employed but leads to undesirable artifacts in the form of small imaginary frequencies. These preclude 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. Here, we compare that approach to an alternative strategy in which fixed-atom contributions to the Hessian are simply limited thermochemical calculations to enthalpies only. Full-dimensional vibrational frequency calculations are possible by replacing the fixed-atom constraints with harmonic confining potentials. These preclude evaluation of finite-temperature free-energy corrections, limiting thermochemical calculations to enthalpies only. 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.42-43

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 often results in rearrangements that do not reflect the crystal structure. To avoid such unphysical rearrangements, the Cα 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.44 We refer to these as fixed-atom constraints,44 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.45–47 This obviates some of the advantage of

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 that, 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 combined with all-atom MD simulations.6-9 Setup of QM/MM calculations requires considerable care,9-15 and an under-appreciated aspect is just how slowly thermochemical predictions converge with respect to the size of the QM region, typically requiring hundreds of QM atoms.16-28 Some progress has been made toward automated selection of QM model regions.26-34

An alternative to QM/MM simulations for mechanistic studies of enzyme catalysis is to use limited “cluster” models of the active site.35-41 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.42-43

Received: February 8, 2024
Revised: April 5, 2024
Accepted: April 8, 2024
Published: April 22, 2024
using an optimized geometry of a finite model system in lieu of MD simulations, which would sample the low-lying minima.

Fixed-atom constraints invariably give rise to a large number of imaginary vibrational frequencies upon structure relaxations. These have been noted in numerous publications, yet a satisfactory explanation has never been provided. Their magnitude is typically small, being variously reported (for the modes exhibiting negative curvature) as $|b_l| \sim 10$ cm$^{-1}$, $|b_l| \lesssim 20$ cm$^{-1}$, $|b_l| \lesssim 30$ cm$^{-1}$, $|b_l| \lesssim 40$ cm$^{-1}$, $|b_l| \lesssim 50$ cm$^{-1}$, $|b_l| \lesssim 60$ cm$^{-1}$, $|b_l| \lesssim 70$ cm$^{-1}$, or $|b_l| < 100$ cm$^{-1}$. (Sometimes they are simply reported as “small”). Imaginary frequencies engendered by the application of constraints have been noted in QM/MM calculations as well, where they have been characterized as “unavoidable”. The present work demonstrates how these imaginary frequencies can be rigorously eliminated.

When the imaginary frequencies are small, a case can be made to ignore them for the purpose of obtaining zero-point energy (ZPE) corrections to reaction enthalpies and barrier heights. This may be a workable strategy to obtain thermochemistry and kinetics at absolute zero, but the harmonic partition function affording the finite-temperature vibrational entropy diverges in the presence of imaginary frequencies. While it has been suggested that thermally populated modes contribute little to enzymatic thermodynamics, it is not obvious why this should be true. Low-frequency modes can be significantly populated at finite temperature, and (for example) a single $30$ cm$^{-1}$ frequency contributes 1.7 kcal/mol to $\Delta G^\circ$. Alternatives to the harmonic approximation have been suggested for the entropy associated with low-frequency motion, yet these alternatives still require harmonic frequencies as inputs. Finally, it has been argued that entropic effects are generally small (meaning $\Delta H^\circ \approx \Delta G^\circ$) in the context of quantum-chemical cluster modeling of enzymatic reactions, yet this assumption cannot be systematically tested so long as cluster models are beset by imaginary frequencies.

To eliminate the 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

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

for a system in which $N_{\text{fix}}$ of $N_{\text{atoms}}$ 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 because the artificial imaginary frequencies that are removed likely correspond to low-frequency vibrations with significant thermal populations at room temperature.

In a previous study, we demonstrated that the use of soft harmonic confining potentials for anchor atoms achieves the same objective as immobilizing those atoms, without producing imaginary frequencies. 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 anchor 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

Large enzyme models often require a significant number of fixed anchor-atom constraints, e.g., up to $N_{\text{fix}} = 31$ in the examples considered herein. To define all of these constraints via internal coordinates would be cumbersome at best. Moreover, constrained geometry optimization algorithms are typically much less efficient than unconstrained optimization. As such, we regard Cartesian constraints as the only generally viable option for large enzyme models.

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

$$\mathbf{g}_i = \begin{cases} 0 & \text{if } x_i \text{ is constrained} \\ \frac{dE}{dx_i} & \text{otherwise} \end{cases}$$

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, fixed-atom vibrational frequency calculations were performed in that way.

Alternatively, problems with imaginary frequencies can be somewhat mitigated by means of a makeshift approach based on a modified Hessian matrix ($\mathbf{H}$) whose elements are

$$\mathbf{H}_i = \begin{cases} 0 & \text{if } x_i \text{ or } x_j \text{ is constrained} \\ \frac{d^2E}{dx_idx_j} & \text{otherwise} \end{cases}$$

This amounts to computing the usual Hessian ($\mathbf{H} = \frac{d^2E}{dx_idx_j}$), then setting certain matrix elements to zero. We refer to this procedure as the “zeroed-out Hessian” technique. A recent review suggests that this approach is widely available in QM/MM programs, and it typically avoids imaginary frequencies. In a sense, however, this procedure merely hides the problem by removing the contributions of the anchor atoms to the normal mode analysis, thereby reducing the number of meaningful frequencies to $N_{\text{vib}}$ (eq 1). It has also been suggested to treat the imaginary frequencies as if they were real, but that approach does not seem to have any sound theoretical basis and is not used here.

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. In this harmonic-confiner model, an additional classical energy term is introduced, namely

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

where $\mathbf{r}_i$ indicates the Cartesian coordinates of the $i$th restrained atom, whose anchor position is $\mathbf{r}^0_i$. In this work, the same force constant is used for each restrained atom, and its value $k = 450$ N/m reflects a typical C–C single bond. (Previous work has explored variations in the value of $k$). All restrained atoms are $C_\alpha$ carbons of the peptide backbone. Analytic derivatives of $V_{\text{conf}}$ were incorporated into the gradient and Hessian calculations, and no gradient or Hessian...
3. RESULTS AND DISCUSSION

3.1. 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 ($\epsilon = 1$), with dihedral $\psi$ and $\phi$ 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 Figure 1, for use in subsequent restrained-atom optimization and vibrational frequency calculations. For those calculations, we defined four anchor atoms: two C$_A$ atoms at the points of separation from the tetrapeptide, as well as two C$_\beta$ atoms on side chains R$_A$ and R$_B$; see Figure 1b. Restricted-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 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 that is obtained ($n_{im}$). In two cases, a single imaginary frequency remains when harmonic confining potentials are used. This is associated with the methyl group on residue A and 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 $n_{im} = 1$ or 2, suggesting that most of the imaginary frequencies originate from the use of anchor atoms rather than dihedral constraints.

When the zeroed-out Hessian is used at the geometry that was optimized using fixed-atom constraints, the number of imaginary frequencies is much larger, $n_{im} = 12$ or 13 (see Table 1). Defining

$$n_{im}^0 = n_{im} - 3N_{fix}$$

which reduces the imaginary frequency count by the $3N_{fix}$ imaginary values that we expect from fixing $N_{fix}$ atoms in space, we find that $n_{im}^0 = 0$ or 1 for the zeroed-out Hessian technique. This suggests that one should expect $3N_{fix}$ imaginary frequencies, in a fixed-atom calculation with $N_{fix}$ anchor atoms, even when the corresponding matrix elements of $H$ are set to zero. This observation alone explains the prevalence of imaginary frequencies in cluster-QM enzymology.

ZPE corrections in Table 1, which are computed using only the real-valued frequencies. Computed at the B3LYP+D3/6-31G(d,p) level.

<table>
<thead>
<tr>
<th>peptide (AB)</th>
<th>unmodified Hessian</th>
<th>zeroed-out Hessian</th>
<th>harmonic confiner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala-Arg</td>
<td>225.0</td>
<td>218.0</td>
<td>231.8</td>
</tr>
<tr>
<td>Asn-Asp</td>
<td>180.7</td>
<td>174.0</td>
<td>188.2</td>
</tr>
<tr>
<td>Cys-Glu</td>
<td>180.6</td>
<td>172.8</td>
<td>187.6</td>
</tr>
<tr>
<td>Gly-Gln</td>
<td>169.9</td>
<td>163.8</td>
<td>177.4</td>
</tr>
<tr>
<td>His-Ile</td>
<td>239.5</td>
<td>231.5</td>
<td>246.3</td>
</tr>
<tr>
<td>Leu-Lis</td>
<td>271.4</td>
<td>263.1</td>
<td>278.3</td>
</tr>
<tr>
<td>Met-Phe</td>
<td>240.6</td>
<td>232.8</td>
<td>248.0</td>
</tr>
<tr>
<td>Pro-Ser</td>
<td>179.4</td>
<td>170.3</td>
<td>185.6</td>
</tr>
<tr>
<td>Thr-Trp</td>
<td>243.6</td>
<td>234.9</td>
<td>250.5</td>
</tr>
<tr>
<td>Tyr-Val</td>
<td>242.6</td>
<td>234.7</td>
<td>245.9</td>
</tr>
</tbody>
</table>

*ZPE in kcal/mol, using only the real-valued frequencies. Computed at the B3LYP+D3/6-31G(d,p) level.
calculations. 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.\textsuperscript{131–133} The focus here is on calculation of reaction barriers (to elucidate mechanistic information), and our subsequent examples will explore realistic quantum-chemical cluster models of enzyme active sites.

3.2. 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).\textsuperscript{135} The charge-neutral cluster model is taken from ref \textsuperscript{135} and is shown in Figure 2, highlighting the biphenyl moiety. We constructed a one-dimensional scan around the indicated dihedral angle (\( \theta \)) between phenyl rings of the biphenyl moiety, relaxing the other degrees of freedom at each value of \( \theta \). These one-dimensional scans are plotted in Figure 3 using either vacuum boundary conditions (\( \epsilon = 1 \)) or dielectric boundary conditions with \( \epsilon = 4 \), 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,\textsuperscript{135} using B3LYP+D3/6-31G(d') and C-PCM with \( \epsilon = 4 \), reported a minimum-energy biphenyl angle \( \theta = -2.5^\circ \). Using fixed-atom constraints and scanning in 5\(^\circ\) increments, we obtain a minimum at \( \theta = -5.0^\circ \) using either \( \epsilon = 1 \) or \( \epsilon = 4 \) (see Figure 3). Our procedure differs in minor details from that in ref \textsuperscript{135}, with a slightly different basis set [6-31G(d,p)] and a somewhat different solute cavity construction, but we regard our results as 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 energy profile,\textsuperscript{136} but our purpose here is to make contact with previous studies that employed fixed-atom constraints.

In the present work, fixed-atom constraints result in nearly identical torsional potentials regardless of whether \( \epsilon = 1 \) or \( \epsilon = 4 \); see the overlay of the two potentials that is provided in Figure 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. Average root-mean-square deviations (RMSDs) between structures relaxed using \( \epsilon = 1 \) versus \( \epsilon = 4 \) are 0.102 and 0.138 Å, respectively. The structures with the largest RMSDs are those with the largest differences in energy, with values as large as 0.382 Å for \( \epsilon = 1 \) at \( \theta = 25^\circ \) and 0.285 Å for \( \epsilon = 4 \) at \( \theta = -25^\circ \).

Harmonic restraints alleviate artificial rigidity introduced by the fixed atoms, leading to softening of the torsional potential. The differences are more pronounced for \( \epsilon = 4 \), but even in the case of vacuum boundaries, one can observe this lowering of the potential for \( \theta \gtrsim 10^\circ \). The minimum-energy structure is not affected because energy changes within about 5\(^\circ\) of the minimum-energy torsion angle are no larger than \( \sim 0.1 \text{ kcal/mol} \) for the calculations using harmonic confining potentials. For larger values of \( \theta \), however, there is significant energy lowering under dielectric (\( \epsilon = 4 \)) 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 influence of dielectric boundary conditions wanes as the size of the QM-cluster model increases, especially for models with \( \geq 150 \) atoms. In our own work using sizable enzyme models, including some with ionic side chains, 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.

### 3.3. Reaction Energy Profiles for Enzyme Models

We next examined several active-site models for an enzymatic reaction in which \( \alpha \)-aspartate \( \alpha \)-decarboxylase (AspDC) catalyzes conversion of \( \alpha \)-aspartate to \( \beta \)-alanine. This is an essential process in the biosynthesis of vitamin B5 that produces a precursor to 4′-phosphopantetheine and coenzyme A in bacteria. Modeling by others has explored decarboxylation of \( \alpha \)-aspartate using various cluster models of AspDC, focusing on the C–C bond-breaking step leading to liberation of \( \text{CO}_2 \). One of these active-site models is depicted in Figure 4, derived from the crystal structure of AspDC taken from *Helicobacter pylori* in association with iso-asparagine (PDB: 1UHE). Here, we examine how various anchor-atom constraints impact the predicted barrier heights, ZPE, and vibrational entropy as the model size increases.

Four different model systems, designated I–IV and containing 76–189 atoms, were taken from ref 88. The largest of these (model IV, which is called model IV.2 in ref 88) consists of 15 amino acid residues, some of which are truncated, and a crystallographic water molecule. The total number of anchor atoms ranges from \( N_{\text{fix}} = 5 \) to 14 as listed in Table 2. These model structures were relaxed at the B3LYP +D3(BJ)/6-31G(d,p) level. Only minor changes are observed relative to the structures provided in ref 88. These are documented for model I in Figure 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(BJ)/6-311+G(2d,2p)//B3LYP+D3(BJ)/6-31G(d,p) level, where the basis set used for the single-point energy calculations affords error statistics comparable to much larger basis sets, in benchmark tests. In ref 88, 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. In ref 88, vibrational frequency calculations were omitted for model IV but they are reported here.

Table 2 presents computed values for the forward barrier height

\[
\Delta^\text{‡}E = E_{\text{TS}} - E_{\text{reactants}}
\]  

and for the reaction energy

\[
\Delta_{\text{rxn}}E = E_{\text{products}} - E_{\text{reactants}}
\]  

for models I–IV, using both fixed-atom and harmonic constraints. 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 \( \Delta^\text{‡}E \) nor \( \Delta_{\text{rxn}}E \) 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 \( \Delta^\text{‡}E \) and \( \Delta_{\text{rxn}}E \), 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, computed using fixed-atom constraints, increases from 8 kcal/mol (model III) to 14 kcal/mol (model IV), such that \( \Delta^\text{‡}E(\text{IV}) \) is closer to \( \Delta^\text{‡}E(\text{I}) \), after having decreased in the order \( \Delta^\text{‡}E(\text{I}) > \Delta^\text{‡}E(\text{II}) > \Delta^\text{‡}E(\text{III}) \). (This is consistent with a sizable jump in the energetics between model III and model IV.2 that is reported in ref 88, from 9 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 \( \Delta_{\text{rxn}}E \). 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 88, boundary conditions with \( \varepsilon = 4 \) increase \( \Delta^\text{‡}E \) for model IV by 4.0 kcal/mol and increase \( \Delta_{\text{rxn}}E \) by 5.6 kcal/mol. These corrections are similar for the smaller models, e.g., 3.9 kcal/mol (\( \Delta^\text{‡}E \)) and 2.7 kcal/mol (\( \Delta_{\text{rxn}}E \)) in

![Figure 4. Transition-state structure for model II of AspDC (from ref 88), leading to CO₂ release. Anchor atoms are indicated by connections to cartoon springs.](image-url)
Table 2. Barrier Height ($\Delta^1E$) and Reaction Energy ($\Delta_{\text{rxn}}E$) for Models of AspDC Decarboxylation$^a$

<table>
<thead>
<tr>
<th>model</th>
<th>$N_{\text{atoms}}$</th>
<th>$N_{\text{fix}}$</th>
<th>$\Delta^1E$ (kcal/mol)</th>
<th>$\Delta_{\text{rxn}}E$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>fixed atom</td>
<td>harmonic confiner</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>unmodified$^b$</td>
<td>zeroed out$^b$</td>
</tr>
<tr>
<td>I</td>
<td>76</td>
<td>5</td>
<td>13.4 (−2.7)</td>
<td>8.5 (−3.1)</td>
</tr>
<tr>
<td>II</td>
<td>95</td>
<td>7</td>
<td>13.6 (−2.6)</td>
<td>8.7 (−2.9)</td>
</tr>
<tr>
<td>III</td>
<td>135</td>
<td>9</td>
<td>9.4 (−2.4)</td>
<td>5.4 (−2.5)</td>
</tr>
<tr>
<td>IV</td>
<td>189</td>
<td>14</td>
<td>14.2 (−0.9)</td>
<td>3.4 (−1.1)</td>
</tr>
</tbody>
</table>

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

Figure 5. Energy profiles for AspDC-catalyzed decarboxylation using model systems of increasing size (I–IV) with (a–d) fixed-atom constraints versus (e–h) harmonic confining potentials. All energies are relative to the reactant state (R) and were computed at the B3LYP+D3(BJ)/6-311+G(2df,2p)//B3LYP+D3(BJ)/6-31G(d,p) level.

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 oscillate with respect to model size. For the largest model, the fixed-atom vibrational entropy correction even differs in sign with respect to the harmonic-confiner values, which appear to converge as the model size increases. 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 the 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, the difference between the transition state and product energies becomes more pronounced in models II and III as compared to model I but then contracts significantly in model IV. In all cases, the difference between the transition state and products is smaller when harmonic confining potentials were used. This suggests that the harmonic restraints effectively mitigate rigidity and impart 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, 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 ZPE, which can hinder accurate prediction of reaction thermochemistry and kinetics. Convergence with respect to model size is sometimes erratic, and 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 constraints. The simplicity of the harmonic approach suggests that this should be the default paradigm for quantum-chemical cluster modeling of enzyme active sites.

ASSOCIATED CONTENT
Data Availability Statement
Computational methods used in this work are available in the Q-Chem program. A trial license for Q-Chem can be obtained from https://www.q-chem.com/try. Coordinates for all optimized structures are provided in the Supporting Information.


of two isomers of N-methyl-4-hydroxyproline. ACS Catal. 2024, 14, 4407–4422.


(72) Prejano, M.; Alberto, M. E.; Russo, N.; Toscano, M.; Marino, T. The effects of the metal ion substitution into the active site of metalloenzymes: A theoretical insight on some selected cases. Catalysts 2020, 10, 1038.


(126) 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.


(132) Du, W.-G. H.; Götz, A. W.; Noodlemann, L. DFT \( \text{Fe}^{3+}/\text{O}−\text{O}^{−} \) vibrational frequency calculations over catalytic reaction cycle states in the dinuclear center of cytochrome \( c \) oxidase. Inorg. Chem. 2019, 58, 13933−13944.


