

# Rationalizing the Activity of an “Artificial Diels-Alderase”: Establishing Efficient and Accurate Protocols for Calculating Supramolecular Catalysis

Tom A. Young,<sup>†</sup> Vicente Martí-Centelles,<sup>‡</sup> Jianzhu Wang,<sup>‡</sup> Paul J. Lusby,<sup>\*,‡</sup> and Fernanda Duarte<sup>†\*</sup>

<sup>†</sup>Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford, OX1 3TA, U.K

<sup>‡</sup>EaStCHEM School of Chemistry, University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, Scotland, EH9 3FJ, U.K

**ABSTRACT:** Self-assembled cages have emerged as novel platforms to explore bio-inspired catalysis. While many different size and shape supramolecular structures are now readily accessible, only a few are known to accelerate chemical reactions under sub-stoichiometric conditions. These limited examples point to a poor understanding of cage catalysis in general, limiting the ability to design new systems. Here we show that a simple and efficient density functional theory-based methodology, informed by explicitly solvated molecular dynamics and coupled cluster calculations is sufficient to accurately reproduce experimental guest binding affinities ( $MAD = 1.9 \text{ kcal mol}^{-1}$ ) and identify the catalytic Diels-Alder proficiencies ( $>80 \%$  accuracy) of two homologous  $\text{Pd}_2\text{L}_4$  metallocages with a variety of substrates. This analysis reveals how subtle structural differences in the cage framework affect binding and catalysis. These effects manifest in a smaller distortion and more favorable interaction energy for the catalytic cage compared to the inactive structure. This study gives a detailed insight that would otherwise be difficult to obtain from experiments, providing new opportunities in the design catalytically active supramolecular cages.

## INTRODUCTION

Self-assembly of molecular building blocks is ubiquitous in nature. Examples of this appear in protein cages,<sup>1,2</sup> enzyme complexes,<sup>3</sup> and bacterial nanocompartments.<sup>4</sup> The remarkable properties of these structures have inspired the design of biomimetic systems that exhibit similar properties at a minimalistic scale. In this context, a large variety of self-assembled cages constructed from simple building blocks via metal–ligand interactions,<sup>5,6</sup> hydrogen bonds,<sup>7–9</sup> and other non-covalent interactions<sup>10</sup> have been reported in recent decades.

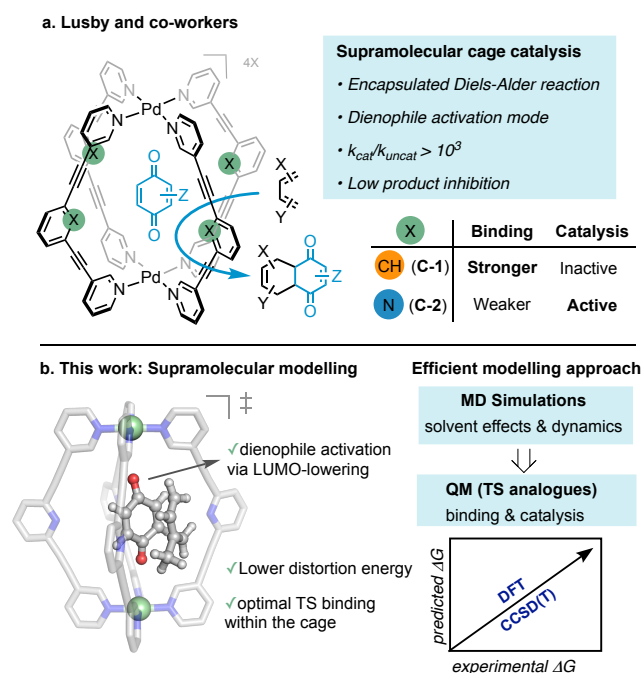
While many cages have been used successfully for recognition and sensing, far fewer have mimicked the catalytic proficiency and selectivity observed in enzymes. More than twenty years have passed since Sanders raised the question regarding the scarcity of effective supramolecular catalysts.<sup>11</sup> Still, his words remain a timely reminder of both the challenges and opportunities within this field. To date, only a handful of self-assembled catalytic metallocages exist,<sup>12</sup> most notably the  $[\text{Ga}_4\text{L}_6]^{12-}$  tetrahedron originally developed by Raymond.<sup>13</sup> Through a fruitful collaboration with Bergman and recently Toste, this water-soluble anionic assembly has been shown to catalyze a number of transformations, including aza-Cope<sup>14,15</sup> and Prins rearrangements,<sup>16,17</sup> the Nazarov cyclisation,<sup>18</sup> and hydrolysis of acid-labile compounds under basic conditions,<sup>19</sup> among others.<sup>12</sup> The prototypical self-assembled metallocages developed in Makoto Fujita’s group have also been shown to catalyze Diels-Alder and Knoevenagel reactions,<sup>20,21</sup> while other water soluble systems can promote different hydrolysis reactions.<sup>22–24</sup> Several trends have emerged from these investigations. First, substrate encapsulation occurs due to the hydrophobic effect. Second, many of the reactions produce water soluble products to avoid inhibition. Thirdly, acceleration is commonly a consequence of

high effective concentration of substrates or functional groups, either through co-encapsulation, ion-pairing or constrictive binding, and/or coulombic interactions between the charged host and electrostatically matched intermediates.<sup>25,26</sup> These approaches while effective in providing acceleration, in many cases suffer from product inhibition (turnover) or are specific to water-soluble cages.

Despite notable examples of cage-catalysis, there has been a distinct lack of complementary computational investigation, probably due to their large size and the presence of multiple metal centers. Indeed, only recently, the catalytic activity of the  $[\text{Ga}_4\text{L}_6]^{12-}$  cage has been explored computationally by Head-Gordon and Ujaque, respectively.<sup>27,28</sup> Other catalytic activity studies have almost exclusively focused on organic systems.<sup>7,25,26,29,30</sup> In this work, we focus on the  $[\text{Pd}_2\text{L}_4]^{4+}$  architecture, which occupies a prominent place in supramolecular chemistry.<sup>31–35</sup> Originally reported by Steel and co-workers,<sup>36</sup> the  $[\text{Pd}_2\text{L}_4]^{4+}$  topology is one of the simplest and most versatile architectures.<sup>32,34</sup> To date, several homo- and heteroleptic variants (consisting of one or different types of ligand, respectively) have been reported by Hooley,<sup>37</sup> Shionoya,<sup>38</sup> Clever,<sup>39–42</sup> Crowley,<sup>43,44</sup> and others,<sup>45–47</sup> with applications in molecular recognition,<sup>47</sup> drug delivery,<sup>31</sup> stabilization of reactive species,<sup>48</sup> and recently catalysis.<sup>49</sup> Despite their promising applications and scope for redesign, no computational investigation has been carried out to explore their catalytic power. The few prior computational studies on this system have only focused on structural<sup>33</sup> and spectroscopic analysis.<sup>50</sup>

In 2018, Lusby and co-workers evaluated the ability of simple  $[\text{Pd}_2\text{L}_4]^{4+}$  capsules **C-1** ( $[\text{Pd}_2\text{L}^{\text{CH}_4}]^{4+}$ ) and **C-2** ( $[\text{Pd}_2\text{L}^{\text{N}_4}]^{4+}$ ) to catalyze Diels-Alder (DA) reactions, using quinone substrates as

dienophiles in dichloromethane (DCM) solvent at room temperature (**Figure 1**).<sup>49</sup> Unlike most other capsule-catalyzed reaction, this method principally exploits enthalpic activation, where C–H hydrogen bond interactions were postulated to activate the dienophile. It was observed that while **C-2** was catalytic (rate accelerations ( $k_{\text{cat}}/k_{\text{uncat}}$ ) of up to  $10^3$ ), the homologous **C-1** cage was inactive, despite **C-1** binding quinones in the same mode. The contrasting catalytic ability was postulated to arise as a result of two factors: weakened substrate binding in **C-2** due to a repulsive interaction in the ground-state between the nitrogen lone pair ( $n$ ) and the  $\pi$ -bond of the guest, and stabilization of the TS through the formation of favorable  $\text{N}\cdots\text{HC}$  hydrogen bonds. The latter was inferred from the stronger binding of a TS mimic (the DA adduct of benzoquinone and cyclopentadiene) in **C-2** compared to **C-1**.<sup>49</sup> However, the difference in catalysis has been purely rationalized on the basis of thermodynamic parameters, such that the precise origin of acceleration, and by extension lack of activity for **C-1**, are still poorly understood. This substantially limits our ability to design supramolecular catalysts for new chemical reactions.



**Figure 1.** (a) Diels-Alder reaction studied by Lusby and co-workers employing the  $[\text{Pd}_2\text{L}_4]^{2+}$  capsules, **C-1** and **C-2**. (b) Schematic representation of our computational approach.

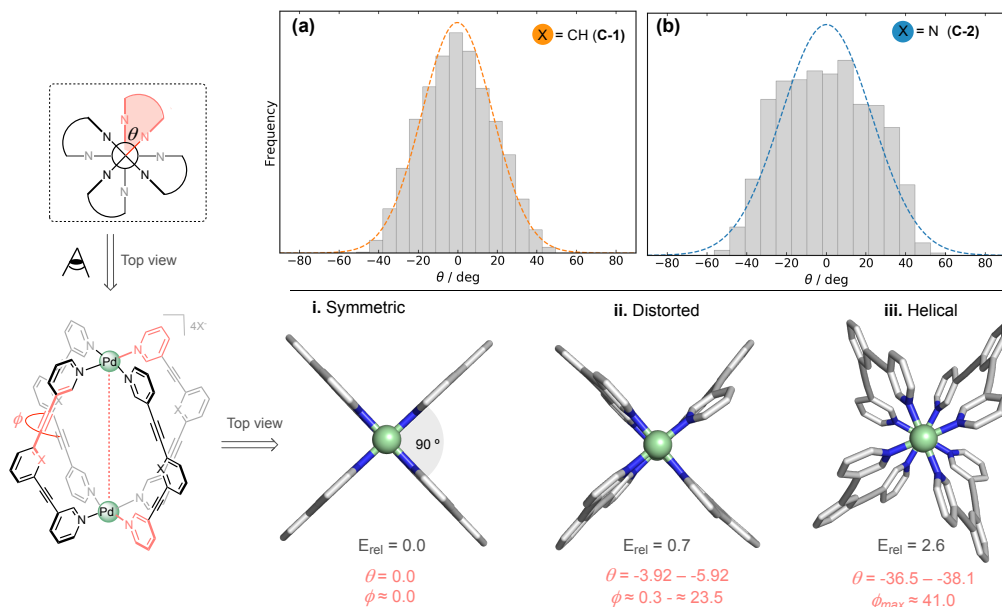
Computational studies that could quantitatively rationalize such observations are highly desirable to optimize and design of novel catalytic cages. However, the ability to routinely study and predict binding and catalysis in solution remains an open challenge. This is particularly true for the latter, where both thermodynamic and kinetic aspects must be considered. In this context, this work aims to rationalize the observed differences in binding and catalysis between metallocages **C-1** and **C-2**, paving the way to an efficient computational protocol to understand related systems. To achieve these goals, we employ explicitly solvated molecular dynamics (MD) and density functional theory (DFT) methods, which are validated against experiments and coupled cluster [CCSD(T)] calculations. Using this protocol, we elucidate the effect of intermolecular interactions, solvent and structural flexibility on the guest binding, and provide molecular-level insight into the catalytic properties of

these systems. Our approach provides an affordable route to explore novel metallocage designs as non-covalent catalysts for applications in synthetic methodology.

## RESULTS AND DISCUSSION

**Dynamic Properties.** The self-assembled structures of **C-1** and **C-2** have been assumed to be highly rigid, even though they contain several potential rotatable  $\sigma$ -bonds (sixteen C–C bonds at either side of the acetylene group and eight Pd–pyridine coordination interactions).<sup>51</sup> To examine the flexibility of these systems, we performed explicitly solvated (DCM) molecular dynamics (MD) simulations in the presence and absence of a quinone guest. A modified version of the  $\text{Pd}^{2+}$  dummy-model – where the square planar geometry is obtained by adding four dummy atoms each with  $+0.5 e$  charge around the Pd center – was used (**Table S1**).<sup>52</sup> This model provided Pd–N distances in the cages **C-1** (avg. = 1.95 Å) and **C-2** (avg. = 1.95 Å) similar to those found for Pd–pyridyl containing cages in the Cambridge Structural Database ( $2.02 \pm 0.01$  Å)<sup>31,53</sup> and also those obtained from DFT calculations (2.02 Å, **Table S2**). However, this model was found to substantially underestimate the Pd–O distances in aqua-complexes (avg. 1.85 Å *cf.* expt.<sup>54</sup> 2.01 Å), a measurement that is usually used to parameterize metal complexes and for which extensive amount of experimental data is available (**Figure S1**).<sup>55</sup> This demonstrates the challenges of directly transferring these parameters between different solvent environments. It is also important to note that while standard soft-sphere models (with no dummy atoms) have been used to model aqua-complexes, it was found that they failed to maintain stable metallocage assemblies in DCM beyond the picosecond time-scale.

With suitable force field parameters to describe the metal and organic building blocks, we then assessed the flexibility of the two cages by monitoring several geometric parameters (**Figures 2** and **S2-S6**). They include the Pd–Pd and Pd–N(pyr) distance, twist angle ( $\theta$ ) and a ‘squareness’ ( $\Delta I$ ) estimate (see **SI §1.2**). Despite the subtle difference between ligands  $\text{L}^{\text{CH}}$  and  $\text{L}^{\text{N}}$ , slightly different Pd–Pd distances are observed, which vary between 11.3 – 12.9 Å for **C-1** and 10.9 – 12.5 Å for **C-2** (avg. 12.11 and 11.74 Å respectively). The slightly shorter Pd–Pd distance of **C-2** has also been observed in several crystal structures. We hypothesize this difference arises from a subtle variation in ligand geometries, with  $\text{L}^{\text{N}}$  having a more concave angle than  $\text{L}^{\text{CH}}$  (**Figure S7**). This also manifests in contrasting helical flexibility, as defined by the twist angle, wherein  $\theta \sim 0^\circ$  corresponds to the highest symmetry conformation and non-zero values are twisted structures. MD simulations show that this twisting can be as high  $\pm 60^\circ$  for both cages. However, the distribution of twisted states is different. For the vacant cages, **C-1** is biased towards the symmetric structure ( $\theta \sim 0^\circ$ , **Figure 2a**) while **C-2** shows a spread of states ( $\pm 30^\circ$ , **Figure 2b**) with similar energy. DFT geometry optimization of representative conformations for **C-1** found during MD simulation, confirm several local minima that could be thermally accessible at room temperature. They include structures with almost perfect  $D_{4h}$  symmetry ( $E_{\text{rel}}(\text{i}) = 0.0 \text{ kcal mol}^{-1}$ ), a slight distortion ( $E_{\text{rel}}(\text{ii}) = 0.7 \text{ kcal mol}^{-1}$ ) and a pronounced helical twist ( $E_{\text{rel}}(\text{iii}) = 2.6 \text{ kcal mol}^{-1}$ , **Figure 2**). Overall, this analysis shows that correlated motion of several partially rotatable bonds gives rise to an overall macromolecular flexibility. The accessibility of helical geometries also provides potential new opportunities in enantioselective catalysis, which have recently started to be explored in alternate metallocage ( $\text{M}_4\text{L}_6$ ) assemblies.<sup>56</sup>

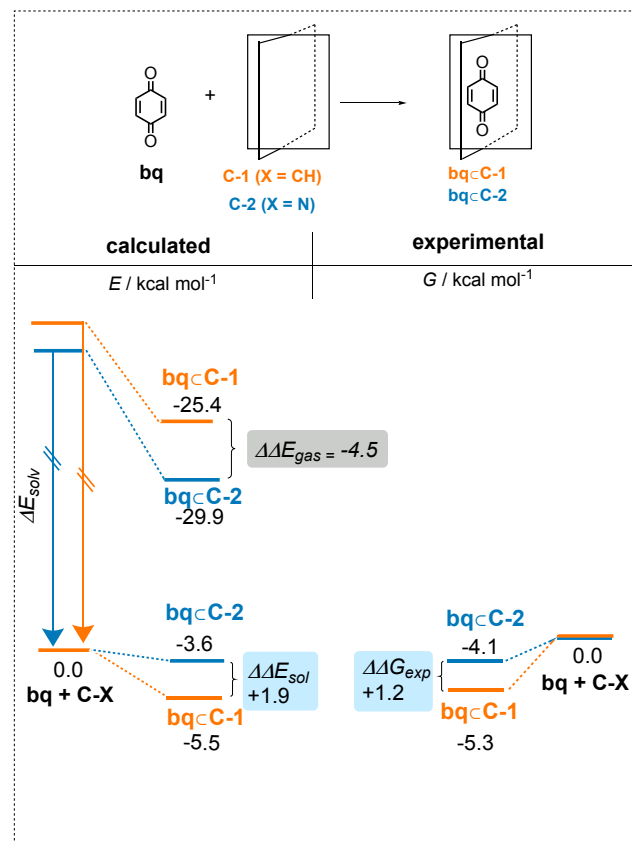


**Figure 2.** (a) Twist angle ( $\theta$ ) frequency for C-1 and C-2 calculated in explicit DCM solvent, over 30 ns of cumulative MD simulations and (b) local minima for C-1 (i-iii) calculated at the PBE0-D3BJ/def2-SVP level of theory. Relative energies ( $E_{rel}$ ) in kcal mol<sup>-1</sup>.

**Efficient Protocol for Binding Affinity Calculations.** To quantitatively rationalize catalysis and binding for C-1 and C-2, each with ~150-200 atoms, a computationally efficient protocol is required. Here, we targeted a total time less than the corresponding experiments would take (~1 day). We selected the M06-2X/def2-TZVP (M2) level of theory, which has been shown to provide accurate association energies for large supramolecular systems.<sup>57</sup> Furthermore, to compare the binding affinity across a range of quinone/Pd<sub>2</sub>L<sub>4</sub> cages, and considering the challenges associated with entropy calculations,<sup>58-60</sup> we analyzed relative potential energy differences ( $\Delta\Delta E_{bind}$ ) rather than Gibbs free energy differences ( $\Delta\Delta G_{bind}$ ), *i.e.* without considering entropic or zero-point-energy corrections, which cancel when comparing both cages (see detailed discussion in SI §3.3).<sup>61</sup>

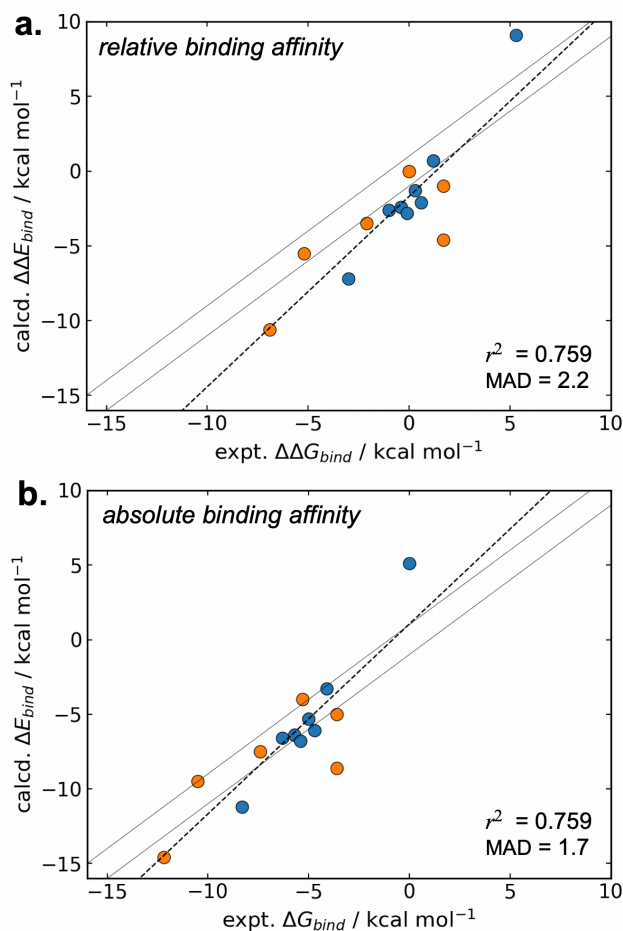
The binding energy of a general quinone (qn) is defined as  $\Delta E_{bind}(qn \subset C-X) = E(qn \subset C-X) - E(qn) - E(C-X)$ , with negative values suggesting favorable binding; the relative binding affinity between the two cages is then defined as  $\Delta\Delta E_{bind}(qn) = \Delta E_{bind}(qn \subset C-2) - \Delta E_{bind}(qn \subset C-1)$ , where positive values indicate a preference of the guest to bind within C-1 over C-2. From **Figure 3a** and **Table S3**, it can be observed that for benzoquinone (bq), in the gas phase, binding is preferred in C-2 over C-1 ( $\Delta\Delta E_{bind} = -4.5$  kcal mol<sup>-1</sup>), in contrast to the solution-phase experimental observation. The gas-phase C-2 preference arises from the H-bond interactions between the nitrogen lone pairs (*n*) and the antibonding  $\sigma^*(C-H)$  orbitals in bq, which are stronger than the interaction between the C-H group of the meta-substituted benzene and the  $\pi$  bond in bq (*vide infra*). However, when implicit DCM solvent is introduced, these interactions are almost entirely masked, leading to a preference for C-1 over C-2, in agreement to experimental results (calcd.  $\Delta\Delta E_{bind} = +1.9$  kcal mol<sup>-1</sup> *cf.* expt.  $\Delta\Delta G_{bind} = +1.2$  kcal mol<sup>-1</sup>). For the larger guest anthraquinone (aq), for which the *n* →  $\sigma^*(C-H)$  interactions are not possible, binding in C-1 is favored in both gas and solvent phase, in good agreement with experiment (calcd.  $\Delta\Delta E_{bind} = +3.7$  kcal mol<sup>-1</sup> *cf.* expt.  $\Delta\Delta G_{bind} = +5.1$  kcal mol<sup>-1</sup>). Pleasingly, the absolute binding affinity of bq was also predicted to within chemical accuracy (**Figure 3**),

which suggest that entropic corrections to the free energy of binding are negligible for this system *i.e.* calcd.  $\Delta E_{bind} \approx$  expt.  $\Delta G_{bind}$ . (see detailed discussion in SI §3.3).<sup>61</sup>



**Figure 3.** Calculated and experimental absolute and relative binding affinities for (a) bq⊂C-1 and bq⊂C-2. Experimental values were obtained in dichloromethane (DCM) solvent at room temperature. Gas phase and solution phase calculations performed at the M06-2X/def2-TZVP and SMD(DCM)-M06-2X/def2-TZVP level of theory, respectively. Values reported in **Table S3**.

The absolute binding affinities of **aq** are overestimated ( $\Delta E_{\text{bind}} = -16.8$  (–13.1) vs  $\Delta G_{\text{bind}} = -10.5$  (–5.4) kcal mol<sup>–1</sup> for **C-1**(**C-2**), **Table S3**), potentially arising from the significant loss in rotational entropy of **aq** compared to **bq** upon binding, in which case  $\Delta S$  is no longer negligible. Nevertheless, the relative binding energies for both cages compare well to experiment. These results demonstrate that the predicted relative binding affinities for **bq** and **aq** between **C-1** and **C-2** are accurate to within ~2 kcal mol<sup>–1</sup>. To probe whether this result was general for this class of system, we considered 14 experimental binding affinities (guest molecules shown in **SI §4**) for **C-1** and **C-2**, including a new data point corresponding to pentacenedione (**q4**) bound to **C-2** (**SI §10**). Bearing in mind the computational expense of performing geometry optimizations for this set at the SMD(DCM)-M2 level of theory (five days on a CPU with 8 cores), calculations were performed at the more efficient M1 level, followed by single point energy evaluations at SMD(DCM)-M2. We refer to this approach as SMD(DCM)-M2//M1 (**Table S8**).



**Figure 4.** Correlation plots of (a)  $\Delta\Delta E_{\text{bind}} = \Delta E_{\text{bind}}(\mathbf{qn} \subset \mathbf{C-X}) - \Delta E_{\text{bind}}(\mathbf{bq} \subset \mathbf{C-1})$ , and (b)  $\Delta E_{\text{bind}}$  calculated at the SMD(DCM)-M06-2X/def2-TZVP//PBE0-D3BJ/def2-SVP level of theory to experimental free energies. Orange and blue markers correspond to **C-1** and **C-2** cages, respectively. The different diagonals bracket the  $\pm 1$  kcal mol<sup>–1</sup> area of accuracy. The dash line represents the line of best fit

In these systems, both relative ( $\Delta\Delta E_{\text{bind}}$ ) and absolute binding energies ( $\Delta E_{\text{bind}}$ ) were calculated (**Figure 4**). Once again, not only are the relative binding affinities obtained to within a reasonable mean absolute deviation ( $\text{MAD} = 2.2$  kcal mol<sup>–1</sup>), the

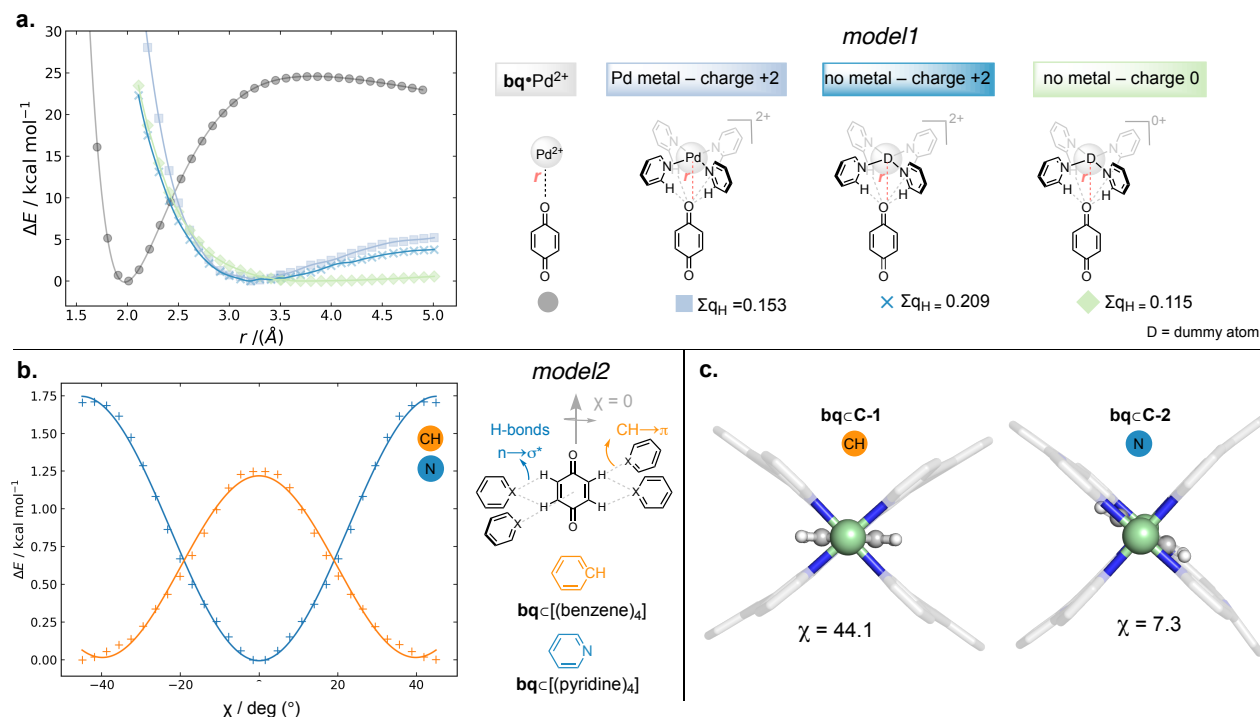
absolute binding affinities are close to chemical accuracy ( $\text{MAD} = 1.7$  kcal mol<sup>–1</sup>, **Figure 4**, **Table S8**), with calculations taking less than a day on a CPU with 8 cores. Compared to predictions in organic host-guest complexes, *e.g.* via the Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) challenge, where errors of at least 1–4 kcal mol<sup>–1</sup> and correlations ( $r^2$ ) below 0.5 have been reported, our correlation to experimental binding affinities ( $r^2 = 0.759$ ) can be considered very good.<sup>57,62,63</sup>

The binding affinity outlier is *tert*-butyl benzoquinone (top right in **Figure 4a**), for which we find binding within the cavity to be unfavorable. Experimentally, this quinone has been hypothesized to bind to the outer ‘pocket’ of the cage, *i.e.* expt.  $\Delta G_{\text{bind}}$  of encapsulation may, in fact, be positive. We also hoped to further accelerate the methodology by obtaining structures at a semi-empirical (PM7<sup>64</sup>) or tight-binding DFT (GFN-xTB<sup>65</sup>) level of theory. Unfortunately, these methods lead to relatively poor correlation ( $r^2(\text{SMD(DCM)-M2//PM7}) = 0.309$ ,  $r^2(\text{SMD(DCM)-M2//GFN-xTB}) = 0.173$ , **Figures S9-S11**). Note the latter is in spite of the relatively small root mean squared deviation (RMSD) to M1 geometries ( $0.3 \pm 0.2$  Å).

**Rationalizing Differences in Binding Affinities.** To elucidate the nature of the non-covalent interactions driving quinone binding, and the differences between **C-1** and **C-2**, we constructed two reduced models. The first uses  $[\text{Pd}(\text{pyridine})_4]^{2+}$  as a model of the ‘top’ and ‘bottom’ of the cage (*model 1*, **Figure 5a**), while the second describes the interaction between the quinone guest with the central meta-substituted benzene or 2,6-pyridyl groups in **C-1** and **C-2**, respectively (*model 2*, **Figure 5b**).

Using *model 1* we analyze the effect of the metal and the +2 charge it introduces (**Figure 5a**), by comparing the canonical complex ( $[\text{Pd}(\text{pyridine})_4]^{2+}$ , ■ light blue) to a complex without metal and an overall +2 charge ( $[(\text{pyridine})_4]^{2+}$ , ×, cyan), and without metal and an overall zero charge ( $[(\text{pyridine})_4]^0$ , ◆ green). This model demonstrates that the metal influences binding by mainly introducing a positive charge that polarizes the adjacent *o*-pyridine (C)H donor groups (**Figure 5a** and **S16**). The second model, *model 2*, describes the interaction between the quinone guest with the ‘equatorial’ meta-substituted phenyl or pyridyl moieties. In this system, rotation of **bq** around the *z*-axis (defined by the angle  $\chi$ , **Figure 5b-c**) shows a minimum at  $\chi = 0^\circ$  for the **bq**⊂ $[(\text{pyridine})_4]$  system (blue line), while for the **bq**⊂ $[(\text{benzene})_4]$  system (orange line), minima are found at  $\chi \pm 45^\circ$ . In the former, electron donation from the nitrogen lone pair (*n*) to the antibonding  $\sigma^*(\text{C-H})$  orbital leads to the formation of weak hydrogen bonds at  $\chi = 0^\circ$  (**Figure S17-S18**).<sup>43</sup> In contrast, for the **bq**⊂ $[(\text{benzene})_4]$  system, repulsion between CH groups produces a maximum at the same position (**Figure 5b**). This explains why in **C-1**, **bq** sits out of the plane of opposing ligands in ( $\chi \approx 45^\circ$ ), while in **C-2**, **bq** lies in the plane ( $\chi \approx 0^\circ$ , **Figure 5c**). The  $n \rightarrow \sigma^*(\text{C-H})$  interactions are strong in the gas phase; however, due to their electrostatic nature, they are lost once implicit solvent is included. In contrast, while the  $\pi \rightarrow \sigma^*(\text{C-H})$  interactions are relatively weak, they are less dependent on the solvent and contribute to the binding of the guest within **C-1** (**Figure S19**). These results explain the observed preference for binding **qn** within **C-1** over **C-2**, which originates from differences in solvation energy, with **C-2** being the intrinsically stronger binder.





**Figure 5.** (a) Comparative PES for **bq** with Pd<sup>2+</sup> (grey), the canonical complex ([Pd(pyridine)<sub>4</sub>]<sup>2+</sup>, light blue) complex without metal and an overall +2 charge ([pyridine]<sub>4</sub>]<sup>2+</sup>, cyan), and complex without metal and an overall zero charge ([pyridine]<sub>4</sub>]<sup>0</sup>, green). The sum of atomic partial charges on the (C)H atoms, using the Hirshfeld scheme, is shown in each case (b) Non-covalent interaction in model 2, including benzenes/pyridines and bq. (c) Global minima of **bqC-1** and **bqC-2**. All calculated at the M06-2X/def2-TZVP level of theory.

**Rationalizing the Differences in Catalytic Activity.** To assess the influence of the cages on DA reactivity, the activation energies for a set of 10 uncatalyzed reactions (Figure S20) were calculated using a highly accurate quantum chemical method. Pleasingly, the error to experiments<sup>49</sup> was within  $\pm 1$  kcal mol<sup>-1</sup> (MAD = 0.9 kcal mol<sup>-1</sup>). Among the different DFT functionals tested, the M2 method was found to have the lowest error relative to the *ab initio* results (MAD 1.2 kcal mol<sup>-1</sup> for reaction barriers and 1.7 kcal mol<sup>-1</sup> for reaction energies, Figure S21-S22). In all cases, a highly synchronous transition state (TS) was observed, with the endo cycloadduct being preferred (Figure S23, Table S13-S15). These results are in agreement with previous computational studies on the uncatalyzed reaction of *para*-quinone imine derivatives using the M06-2X functional in vacuum.<sup>66</sup>

With a chemically accurate computational methodology in hand, the reaction barriers for the DA reaction between **bq** and isoprene (**d1**) in the presence of **C-1** or **C-2** were calculated. Formation of an association complex between **d1** and **bqC-1/C-2** is unfavorable. For example, with **C-2** it leads to a complex 8.8 kcal mol<sup>-1</sup> higher in free energy than **bqC-2** and isoprene separately (Table S16). This result is in line with <sup>1</sup>H NMR data of the catalytic process, which shows that neither the cage nor the quinone signals shift upon addition of diene to the reaction,<sup>49</sup> and contrast with most supramolecular host-guest complexes exhibiting DA catalytic activity, where a so-called “ternary Michaelis complex” is observed.<sup>29</sup> As suggested experimentally, our calculations confirm that the complex **bqC-2** is first formed and then reacts intermolecularly with an incoming isoprene molecule.

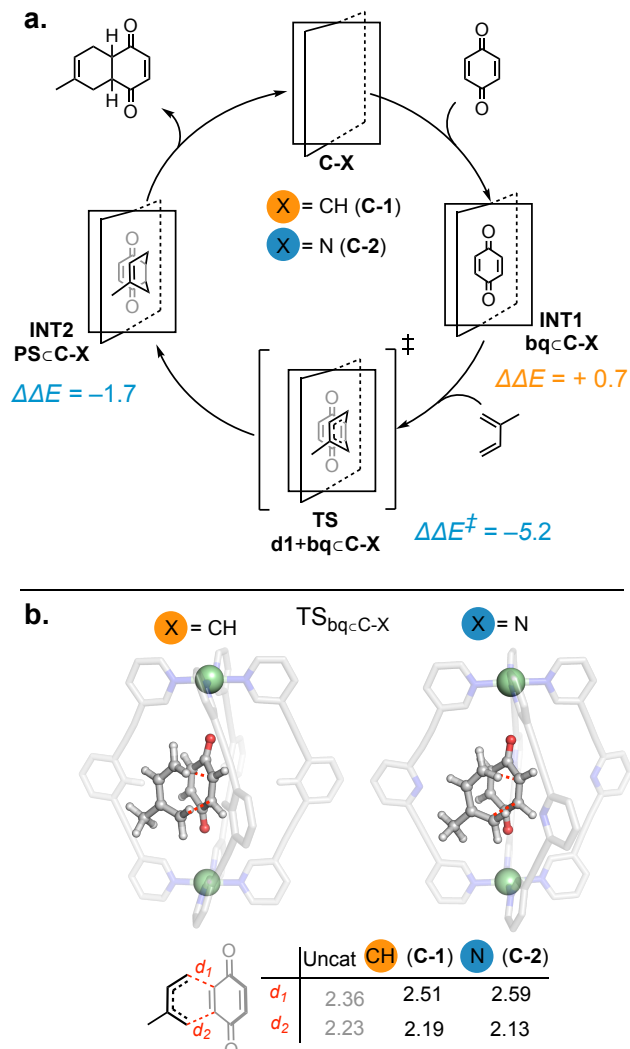
Defining the catalytic activity as  $\Delta\Delta E_{CA} = \Delta E_{uncat}^* - \Delta E_{cat}^*$ , such that positive values correspond to effective catalysis we calculate  $\Delta\Delta E_{CA}(\mathbf{d1+bqC-2}) = +5.4$  kcal mol<sup>-1</sup> (expt.  $\Delta\Delta G_{CA} = +3.6$  kcal mol<sup>-1</sup>) while  $\Delta\Delta E_{CA}(\mathbf{d1+bqC-1}) = -0.3$

kcal mol<sup>-1</sup> (expt. catalytically inactive, Table 1). As shown in Figure 6a, while **C-1** binds the substrate more strongly in the reactant state, **C-2** provides a better stabilization at the TS and slightly better stabilization of the product state (Figure S25). Analysis of the encapsulated TS geometries indicate that, for asymmetric dienes, the cage increases the asymmetric nature of the TSs compared to the uncatalyzed analogues, making it less pericyclic and more stepwise (Figure 6b). The active **C-2** cage affords a slightly more asynchronous reaction ( $\Delta d = 0.28/-0.17$ , 0.36/-0.23 Å for TS[d1+bqC-1] and TS[d1+bqC-2], respectively).

To rationalize these differences, we explored the electronic and steric contributions to the reaction energy in each system. In both cases, and as previously seen for Lewis acid-catalyzed [4+2] cycloadditions,<sup>67-69</sup> the cages enhance the electrophilic character of the dienophile bound in the cavity. Polarization of the dienophile within the cage leads to a net charge of +0.25 and +0.16 in **bq** for **C-1** and **C-2**, respectively,<sup>70</sup> leading to a lowering of the LUMO energy by 1.5 and 1.6 eV in **C-1** and **C-2**, respectively (Table S19). These results, demonstrate that both **C-1/C-2** can activate the dienophile by reducing its LUMO energy. However, as discussed below, other interactions in **C-1** offset this activating effect. To analyse this contrasting behaviour between **C-1** and **C-2**, we performed an energy decomposition analysis using the distortion/interaction framework.<sup>71</sup> Within this framework, the activation energy of a reaction ( $\Delta E^*$ ) is partitioned into the energy required to distort the reactants from their ground-state to their transition state geometry ( $\Delta E_{dist}^*$ ) and the energy of interaction between these fragments ( $\Delta E_{int}^*$ ). For the reaction under study, the optimal interaction between the activated dienophile and the incoming diene takes place within one of the cage portals (Figure 6b); this requires the ligand to distort significantly, thus suggesting that distortion of the cage and the substrate is required to reach the TS in the cage-catalyzed process.

**Table 1.** Distortion/Interaction Analysis for the [4+2] cyclisation of isoprene with benzoquinone calculated at the SMD(DCM)-M2 level of theory. Shown here are the distortion energies of the cage ( $\Delta E_{\text{d[cage]}}^*$ ), substrate ( $\Delta E_{\text{d[diene]}}^*$ ), dienophile ( $\Delta E_{\text{d[bq]}}^*$ ), and **bq**⊂**C-X** complex ( $\Delta E_{\text{d[bq⊂C-X]}}^*$ ). The interaction energy is defined as  $\Delta E_{\text{int}}^* = \Delta E^* - \Delta E_{\text{d}}^*$ , where  $\Delta E^*$  is the activation energies and total  $\Delta E_{\text{d}}^* = \Delta E_{\text{d[bq⊂C-X]}}^* + \Delta E_{\text{d[diene]}}^*$ . All energies are given in kcal mol<sup>-1</sup>. \*sum of diene and quinone distortion.

|                                 | Distortion Energy components |                                |                             |                                 | Total Distortion Energy               | Interaction Energy (Activation – Distortion) | Activation Energy |
|---------------------------------|------------------------------|--------------------------------|-----------------------------|---------------------------------|---------------------------------------|--|-------------------|
|                                 | $\Delta E_{\text{d[C-X]}}^*$ | $\Delta E_{\text{d[diene]}}^*$ | $\Delta E_{\text{d[bq]}}^*$ | $\Delta E_{\text{d[bq⊂C-X]}}^*$ | $\Delta E_{\text{d[bq⊂C-X+diene]}}^*$ | $\Delta E_{\text{int}}^*$                    | $\Delta E^*$      |
| Uncat.                          | --                           | 15.8                           | 7.1                         | --                              | 22.9*                                 | -13.0  | 9.9               |
| C-1                             | 5.2                          | 16.2                           | 8.8                         | 13.2                            | 29.4                                  | -19.2  | 10.2              |
| C-2                             | 5.1                          | 14.9                           | 9.3                         | 10.0                            | 24.9                                  | -20.4  | 4.5               |
| $\Delta\Delta E_{\text{C2-C1}}$ | -0.1                         | -1.3                           | 0.5                         | -3.2                            | -4.5                                  | -1.2   | -5.7              |



**Figure 6.** (a) Catalytic cycle for the reaction between benzoquinone and isoprene ( $\Delta\Delta E = \Delta E(\text{C-2}) - \Delta E(\text{C-1})$  in kcal mol<sup>-1</sup> at SMD(DCM)-M06-2X/def2-TZVP//PBE0-D3BJ/def2-SVP, negative(positive) values in blue(orange) refer to favorable stabilization for C-2 (C-1)). (b) Optimized TS geometries; Distances (Å) of the forming bonds at the TS for the uncatalyzed (grey) and encapsulated (black) reaction.

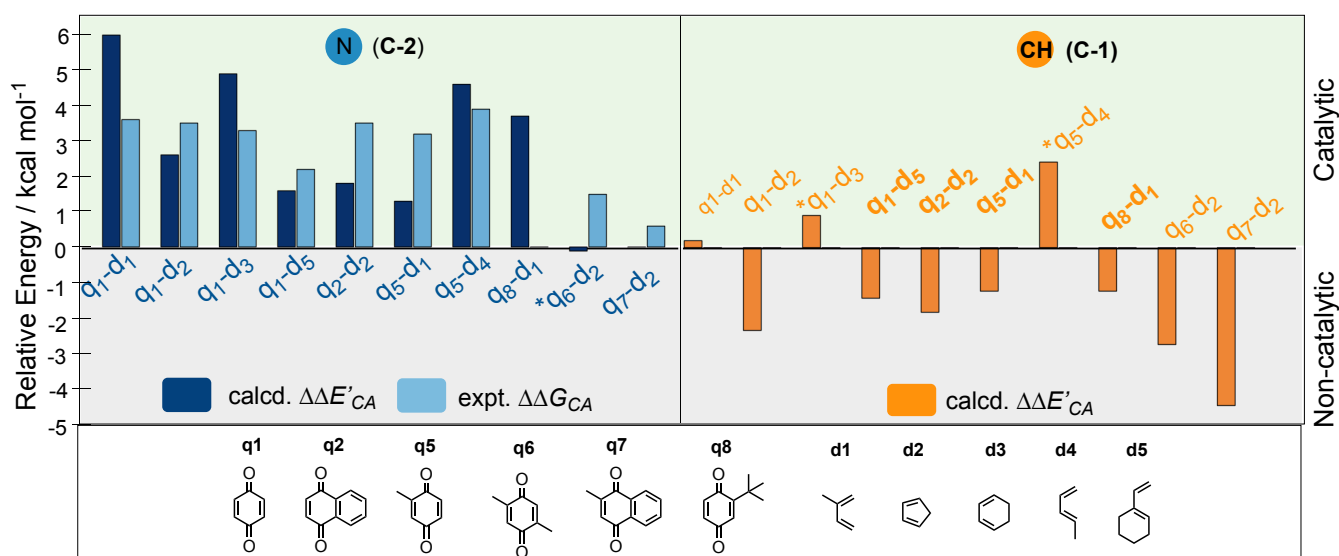
The energetic cost of distorting each component of the reaction is shown in **Table 1**. Here, we separated each TS structure into three fragments ( $\Delta E_{\text{d[C-X]}}^*$ ,  $\Delta E_{\text{d[diene]}}^*$ , and  $\Delta E_{\text{d[bq]}}^*$ ) and also evaluated the **bq**⊂**C-X** complex as a single fragment ( $\Delta E_{\text{d[bq⊂C-X]}}^*$ ). In C-2 a smaller distortion compared to C-1 is

obtained for both the diene (1.3 kcal mol<sup>-1</sup>) and the cage-dienophile complex (4.5 kcal mol<sup>-1</sup>). For C-1, the cost of aligning the dienophile **bq** with the cage portal is found to be twice the cost seen in C-2 (**Figure S29**). This can be associated to steric clashes arising between the central moiety of **L<sup>CH</sup>** and the dienophile. This effect can be visually understood through the use of non-covalent interaction (NCI) plots, which in C-1 shows steric clashes between the dienophile and the cage, while in C-2 these regions are smaller and compensated by a favourable N⋯HC hydrogen interaction between the diene and the cage. The latter may also lead to the slightly more positive interaction term for this cage (**Figure S29**). For both cages, the interaction energy ( $\Delta E_{\text{int}}^*$ ) is more favorable compared to the uncatalyzed reaction, which is in line with lowering of dienophile LUMO energy. However, distortion effects render C-1 uncatalytic. These results demonstrate that small differences in cage sterics and flexibility play an important role in determining the potential catalytic activity of very similar metallocages towards a given reaction. Such energetic costs will vary depending on the activation mode required and the nature of the substrates involved but can be easily quantified.

Finally, while our analysis in **Table 1** included only potential energies, Gibbs free energies can be obtained by adding thermal and entropic contributions (**Table S20**). By doing so, we obtain  $\Delta G_{\text{uncat}}^* = 23.5$  kcal mol<sup>-1</sup> and  $\Delta G_{\text{cat}}^* (\text{C-1/C-2}) = 23.6/18.8$  kcal mol<sup>-1</sup>, which are in good agreement with experimental results and confirm that catalysis in C-2 is enthalpic in origin.

**Efficient Protocol for Calculating Catalytic Activity.** Having rationalized the catalytic properties of the two cages for a single cycloaddition reaction, we sought to validate our methods with a wider range of substrates. Although fully optimized TSs for the DA reactions outlined in **Figure 7** can be located within these cages, the computational demand is significant, even using a low level of theory (>3 days on a CPU with 8 cores). With the aim to decrease this time, we proposed to evaluate catalytically promising cages considering the following approximations: (1) To rely on potential energy rather than free energies differences, and (2) to calculate the activation energy of ‘TS analogues’ rather than the true transition state. We use the term ‘TS analogues’ to indicate that the bonds being formed and broken are constrained to the distance values found in the uncatalyzed TS.

The catalytic activity ( $\Delta\Delta E'_{\text{CA}}$ ) is then found by performing a constrained minimization and comparing to the uncatalyzed variant (**Figure 7**). This computationally inexpensive approach can achieve 80% accuracy in predicting the catalytic proficiency towards the DA reactions tested (**Figure 7**), while providing a ten-fold reduction in computational time.



**Figure 7.** Calculated analogue activation energy ( $\Delta\Delta E'_{CA}$ , dark color) and experimentally observed catalytic efficiency ( $\Delta\Delta G_{CA}$ , light blue, only available for C-2). For C-1, the experimentally untested quinone-diene reactions are shown in bold. Two false positive for C-1 are indicated with an asterisk \*. Values tabulated in Table S20.

Using this computationally inexpensive approach cage C-1 was found to be non-catalytic for seven out of ten reactions tested, while **q1-d1** and **q1-d3** combinations were within computational error. However, the substrate combination **q5-d4** is clearly predicted to lead to catalysis. As this cage-substrate combination was not reported in our original work,<sup>49</sup> we have now monitored this reaction and found no obvious acceleration compared to the cage-free control experiment (SI §10). This suggests that there are other subtleties involved in promoting catalysis, which will need to be explored in further studies.

We have compared these results to those obtained using full optimization (Figure S30). Selecting the set of substrates with the largest discrepancy to experiments (**q1+d1** **q5+d4**) and **q7+d2** within C-1 and C-2, we observed that the full transition states provided almost identical results to those obtained using TS analogues ( $r^2=0.94$ ,  $MAD=0.6$  kcal mol<sup>-1</sup>, Figure S30). Moreover, the number of CPU-hours (CPUh) required to calculate a single  $\Delta\Delta E$  value using TS analogues (~100 CPUh, up to 12 hours on 8CPUs) is at least an order of magnitude lower than when using fully optimised TS (~1000 CPUh, 5 days on 8 CPUs using the TS analogues as starting point), thus demonstrating the substantial computational cost saved by our approach without compromising accuracy.

## CONCLUSION

Supramolecular metallogages have emerged as promising biomimetic catalysts. However, a theoretical understanding of the structural and electronic factor that determine efficient binding and catalysis has been lacking. Here, we have rationalized the binding and significantly different Diels-Alder catalytic activity for two highly-homologous metallogages (C-1 and C-2) through complementary classical simulations and quantum calculations. This is the first computational study that simultaneously explores binding and catalysis with multiple cage and substrate combinations. For the DA reactions studied here, we find that electronic activation of dienophile is observed for both cages. However, in the case of inactive C-1, this is offset by significant distortion energy which inhibits the interaction with the incoming diene. In contrast, catalytically active C-2 is much

more able to accommodate distortion approaching the transition state. This suggests that, in addition to the catalytic machinery required to activate the substrates undergoing chemical reactions, novel metallogages designs should consider the plasticity of the building blocks and the cage as a whole. While these features are now understood in enzyme catalysis, they have been much less studied in biomimetic catalyst design.

In addition to rationalizing the catalytic activity, this work introduces an efficient computational protocol to model metallo-supramolecular catalysts in solution. Further work is underway to automate this process within an open-source Python module (*cgbind*, <https://github.com/duartegroup/cgbind>). We expect this to motivate a closer interaction between experimentalists and computational chemists in the quest towards the discovery of novel catalysts.

## COMPUTATIONAL METHODS

**MD Simulations.** Configurational sampling of the cage structures in dichloromethane (DCM) was performed using classical MD simulations. For ligands and substrates, OPLS-AA-compatible force-field parameters were generated using the Macromodel ffd14 version (Schrödinger LCC).<sup>72</sup> Restrained electrostatic potential (RESP) charges were derived by fitting partial charges to HF/6-31G(d) electrostatic potentials calculated using Gaussian 09 (version D.01).<sup>73</sup> Geometries for these calculations were obtained from optimizations at the TPSS-D3BJ/def2TZVP/fit level of theory. The topology for DCM was taken from the GROMACS Molecule & Liquid Database.<sup>74,75</sup>

All simulations were performed using the GROMACS package (version 2019.1)<sup>76,77</sup> using three-dimensional periodic boundary conditions. Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) approach<sup>78</sup> with a cut-off length of 1.0 nm. A dispersion correction was applied to energy and pressure terms to account for truncation of van der Waals terms.

The systems were immersed in a box of solvent with a distance from the border of at least 15 Å. Following steepest descent minimization, the systems were equilibrated in two steps; the

first phase involved simulating for 100 ps under a constant volume (NVT) ensemble with position restraints applied to heavy atoms. The temperature was maintained at 298 K using the V-Rescale method. This was followed by 100 ps of constant-pressure (NPT) equilibration using the Parrinello–Rahman pressure coupling algorithm with the compressibility set to  $4.5 \times 10^{-5}$  bar<sup>-1</sup> and the time constant set to 5 ps. A 1 fs time step was used during these stages to allow potential inhomogeneities to self-adjust. Each system was then equilibrated initially for 10 ns with a 1 fs time step at constant pressure. Three different runs, differing in the initial random seed were run per system. The first 1 ns was removed from the analysis.

**Electronic Structure Calculations.** All density functional theory (DFT) calculations were performed in the ORCA v. 4.1 software<sup>79</sup> package. Initial geometry optimizations for cages (C-1, C-2) and cage–substrate complexes (bq⊂C-1, bq⊂C-2) were performed with the M06-2X functional<sup>80</sup>, the Ahlrichs def2-TZVP<sup>81, 82</sup> basis set which includes the ECP28MWB<sup>81</sup> on Pd. Solvent effects were accounted for with the SMD<sup>83</sup> solvent model with parameters appropriate for DCM. Subsequent optimizations were performed with the PBE0<sup>84</sup> functional with the Becke-Johnson damped D3 dispersion correction (D3BJ)<sup>85</sup> in combination with a def2-SVP basis set. Where stated vibrational frequencies were computed at the optimization level of theory to confirm whether the structures correspond to minima or a transition states and to evaluate the zero-point vibrational energy (ZPVE) and thermal corrections at 298 K. Gibbs free energies in which a change of molecularity occurs include a 1 atm to 1 M standard state correction.

Single-point energies were obtained at the M06-2X/def2-TZVP level of theory in DCM solvent using the SMD solvent model. Single point energies for non-catalyzed DA reactions have also been carried out using PBE0,<sup>84</sup> TPSS,<sup>86</sup> PW6B95,<sup>87</sup> ωB97X-D3<sup>88</sup> and M06-2X<sup>80</sup> functionals to analyze the influence of the functional and dispersion upon optimized structures and stabilities when compared to the DLPNO-CCSD(T)<sup>89</sup> linear-scaling coupled cluster implementation, which was used in combination with the minimally augmented doubly polarized ma-def2-TZVPP basis set. The chain of spheres resolution of identity (RIJCOSX in ORCA) approximation was used for the Hartree-Fock exchange component of hybrid functionals, with the default fitting bases. NBO charge and second order perturbation analyses were carried out in NBO<sup>90</sup> v. 6.0. Multiwfn v. 3.6<sup>91</sup> has been used for electron density analysis. Unless specified otherwise, all energy differences are calculated at SMD(DCM)-M06-2X/def2-TZVP//PBE0-D3BJ/def2-SVP level of DFT theory.

## ASSOCIATED CONTENT

### Supporting Information

Analysis of MD simulations; discussion on entropic contributions, absolute binding and reaction energies; energies (a.u.) selected structures (PDF). Archive with cartesian coordinates (in xyz format) and MD topologies and input files (ZIP). The Supporting Information is available free of charge on the ACS Publications website.

## AUTHOR INFORMATION

### Corresponding Author

\*fernanda.duartegonzalez@chem.ox.ac.uk

\*Paul.Lusby@ed.ac.uk

**Author Contributions:** The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

## ACKNOWLEDGMENT

We acknowledge the EPSRC Centre for Doctoral training, Theory and Modelling in Chemical Sciences (EP/L015722/1) for a studentship to TAY generously supported by AWE and for access to the Dirac cluster at Oxford. This work was also supported by the Leverhulme Trust (VMC; RPG-2015-232) and the University of Edinburgh (JW, Principal's Career Development Scholarship). The authors acknowledge the use of the EPSRC Tier-2 National HPC Facility Service Cirrus (www.cirrus.ac.uk). We thank Alistair Sterling and Rebecca Spicer for insightful comments.

## ABBREVIATIONS

DA: Diels-Alder; bq: benzoquinone; DFT: density functional theory; MAD: mean absolute deviation; MD: molecular dynamics; SMD: solvent model density and TS: transition state.

## REFERENCES

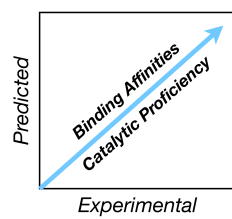
1. Theil, E. C., Ferritin Protein Nanocages Use Ion Channels, Catalytic Sites, and Nucleation channels to Manage Iron/oxygen Chemistry. *Curr. Opin. Chem. Biol.* **2011**, *15*, 304-311.
2. Azuma, Y.; Zschoche, R.; Tinzl, M.; Hilvert, D., Quantitative Packaging of Active Enzymes into a Protein Cage. *Angew. Chem., Int. Ed.* **2015**, *55*, 1531-1534.
3. Schmitt, D. L.; An, S., Spatial Organization of Metabolic Enzyme Complexes in Cells. *Biochemistry* **2017**, *56*, 3184-3196.
4. Rurup, W. F.; Snijder, J.; Koay, M. S. T.; Heck, A. J. R.; Cornelissen, J. J. L. M., Self-Sorting of Foreign Proteins in a Bacterial Nanocompartment. *J. Am. Chem. Soc.* **2014**, *136*, 3828-3832.
5. Wiester, M. J.; Ulmann, P. A.; Mirkin, C. A., Enzyme Mimics Based Upon Supramolecular Coordination Chemistry. *Angew. Chem., Int. Ed.* **2011**, *50*, 114-137.
6. Cook, T. R.; Zheng, Y.-R.; Stang, P. J., Metal–Organic Frameworks and Self-Assembled Supramolecular Coordination Complexes: Comparing and Contrasting the Design, Synthesis, and Functionality of Metal–Organic Materials. *Chem. Rev.* **2013**, *113*, 734-777.
7. Pahima, E.; Zhang, Q.; Tiefenbacher, K.; Major, D. T., Discovering Monoterpene Catalysis Inside Nanocapsules with Multiscale Modeling and Experiments. *J. Am. Chem. Soc.* **2019**, *141*, 6234-6246.
8. Zhang, Q.; Tiefenbacher, K., Hexameric Resorcinarene Capsule is a Brønsted Acid: Investigation and Application to Synthesis and Catalysis. *J. Am. Chem. Soc.* **2013**, *135*, 16213-16219.
9. Ajami, D.; Rebek, J., More Chemistry in Small Spaces. *Acc. Chem. Res.* **2013**, *46*, 990-999.
10. Kaanumalle, L. S.; Gibb, C. L. D.; Gibb, B. C.; Ramamurthy, V., A Hydrophobic Nanocapsule Controls the Photophysics of Aromatic Molecules by Suppressing Their Favored Solution Pathways. *J. Am. Chem. Soc.* **2005**, *127*, 3674-3675.
11. Sanders, J. K. M., Supramolecular Catalysis in Transition. *Chem-Eur J.* **1998**, *4*, 1378-1383.
12. Fang, Y.; Powell, J. A.; Li, E.; Wang, Q.; Perry, Z.; Kirchon, A.; Yang, X.; Xiao, Z.; Zhu, C.; Zhang, L.; Huang, F.; Zhou, H.-C., Catalytic reactions within the cavity of coordination cages. *Chem. Soc. Rev.* **2019**, *48*, 4707-4730.
13. Caulder, D. L.; Powers, R. E.; Parac, T. N.; Raymond, K. N., The Self-Assembly of a Predesigned Tetrahedral M4L6 Supramolecular Cluster. *Angew. Chem. Int. Ed.* **1998**, *37*, 1840-1843.
14. Fiedler, D.; Bergman, R. G.; Raymond, K. N., Supramolecular Catalysis of a Unimolecular Transformation: Aza-Cope Rearrangement within a Self-Assembled Host. *Angew. Chem., Int. Ed.* **2004**, *43*, 6748-6751.
15. Brown, C. J.; Bergman, R. G.; Raymond, K. N., Enantioselective Catalysis of the Aza-Cope Rearrangement by a Chiral Supramolecular Assembly. *J. Am. Chem. Soc.* **2009**, *131*, 17530-17531.
16. Hart-Cooper, W. M.; Zhao, C.; Triano, R. M.; Yaghoubi, P.; Ozores, H. L.; Burford, K. N.; Toste, F. D.; Bergman, R. G.; Raymond, K. N., The Effect of Host Structure on the Selectivity and Mechanism of Supramolecular Catalysis of Prins Cyclizations. *Chem. Sci.* **2015**, *6*, 1383-1393.



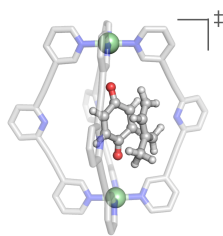
17. Kaphan, D. M.; Toste, F. D.; Bergman, R. G.; Raymond, K. N., Enabling New Modes of Reactivity via Constrictive Binding in a Supramolecular-Assembly-Catalyzed Aza-Prins Cyclization. *J. Am. Chem. Soc.* **2015**, *137*, 9202-9205.
18. Hastings, C. J.; Pluth, M. D.; Bergman, R. G.; Raymond, K. N., Enzymelike Catalysis of the Nazarov Cyclization by Supramolecular Encapsulation. *J. Am. Chem. Soc.* **2010**, *132*, 6938-40.
19. Pluth, M. D.; Bergman, R. G.; Raymond, K. N., Acid Catalysis in Basic Solution: A Supramolecular Host Promotes Orthoformate Hydrolysis. *Science* **2007**, *316*, 85-88.
20. Yoshizawa, M.; Tamura, M.; Fujita, M., Diels-alder in Aqueous Molecular Hosts: Unusual Regioselectivity and Efficient Catalysis. *Science* **2006**, *312*, 251-4.
21. Murase, T.; Nishijima, Y.; Fujita, M., Cage-Catalyzed Knoevenagel Condensation under Neutral Conditions in Water. *J. Am. Chem. Soc.* **2012**, *134*, 162-164.
22. Bolliger, J. L.; Belenguer, A. M.; Nitschke, J. R., Enantiopure Water-Soluble [Fe<sub>4</sub>L<sub>6</sub>] Cages: Host-Guest Chemistry and Catalytic Activity. *Angew. Chem. Int. Ed.* **2013**, *52*, 7958-7962.
23. Cullen, W.; Misuraca, M. C.; Hunter, C. A.; Williams, N. H.; Ward, M. D., Highly efficient catalysis of the Kemp elimination in the cavity of a cubic coordination cage. *Nat. Chem.* **2016**, *8*, 231-236.
24. Cullen, W.; Metherell, A. J.; Wragg, A. B.; Taylor, C. G. P.; Williams, N. H.; Ward, M. D., Catalysis in a Cationic Coordination Cage Using a Cavity-Bound Guest and Surface-Bound Anions: Inhibition, Activation, and Autocatalysis. *J. Am. Chem. Soc.* **2018**, *140*, 2821-2828.
25. Chakraborty, D.; Chattaraj, P. K., Bonding, Reactivity, and Dynamics in Confined Systems. *J. Phys. Chem. A* **2019**, *123*, 4513-4531.
26. Goehry, C.; Besora, M.; Maseras, F., Computational Study on the Mechanism of the Acceleration of 1,3-Dipolar Cycloaddition inside Cucurbit[6]uril. *ACS Catal* **2015**, *5*, 2445-2451.
27. Vaissier Welborn, V.; Head-Gordon, T., Electrostatics Generated by a Supramolecular Capsule Stabilizes the Transition State for Carbon-Carbon Reductive Elimination from Gold(III) Complex. *J. Phys. Chem. Lett.* **2018**, *9*, 3814-3818.
28. Norjmaa, G.; Maréchal, J.-D.; Ujaque, G., Microsolvation and Encapsulation Effects on Supramolecular Catalysis: C-C Reductive Elimination inside [Ga<sub>4</sub>L<sub>6</sub>]<sup>12-</sup> Metallocage. *J. Am. Chem. Soc.* **2019**, *141*, 13114-13123.
29. Daver, H.; Harvey, J. N.; Rebek, J.; Himo, F., Quantum Chemical Modeling of Cycloaddition Reaction in a Self-Assembled Capsule. *J. Am. Chem. Soc.* **2017**, *139*, 15494-15503.
30. Daver, H.; Algarra, A. G.; Rebek, J.; Harvey, J. N.; Himo, F., Mixed Explicit-Implicit Solvation Approach for Modeling of Alkane Complexation in Water-Soluble Self-Assembled Capsules. *J. Am. Chem. Soc.* **2018**, *140*, 12527-12537.
31. Lewis, J. E. M.; Gavey, E. L.; Cameron, S. A.; Crowley, J. D., Stimuli-responsive Pd<sub>2</sub>L<sub>4</sub> metallosupramolecular cages: towards targeted cisplatin drug delivery. *Chem. Sci.* **2012**, *3*, 778-784.
32. Schmidt, A.; Casini, A.; Kühn, F. E., Self-assembled M<sub>2</sub>L<sub>4</sub> Coordination Cages: Synthesis and Potential Applications. *Coord. Chem. Rev.* **2014**, *275*, 19-36.
33. Preston, D.; Barnsley, J. E.; Gordon, K. C.; Crowley, J. D., Controlled Formation of Heteroleptic [Pd<sub>2</sub>(La)<sub>2</sub>(Lb)<sub>2</sub>]<sup>4+</sup> Cages. *J. Am. Chem. Soc.* **2016**, *138*, 10578-10585.
34. Vasdev, R. A. S.; Preston, D.; Crowley, J. D., Multicavity Metallosupramolecular Architectures. *Chem. Asian J.* **2017**, *12*, 2513-2523.
35. Preston, D.; Lewis, J. E. M.; Crowley, J. D., Multicavity [Pd<sub>4</sub>L<sub>4</sub>]<sup>2n+</sup> Cages with Controlled Segregated Binding of Different Guests. *J. Am. Chem. Soc.* **2017**, *139*, 2379-2386.
36. McMorran, D. A.; Steel, P. J., The First Coordinatively Saturated, Quadruply Stranded Helicate and Its Encapsulation of a Hexafluorophosphate Anion. *Angew. Chem. Int. Ed.* **1998**, *37*, 3295-3297.
37. Liao, P.; Langloss, B. W.; Johnson, A. M.; Knudsen, E. R.; Tham, F. S.; Julian, R. R.; Hooley, R. J., Two-component Control of Guest Binding in a Self-assembled Cage Molecule. *Chem. Commun.* **2010**, *46*, 4932-4934.
38. Clever, G. H.; Tashiro, S.; Shionoya, M., Inclusion of Anionic Guests inside a Molecular Cage with Palladium(II) Centers as Electrostatic Anchors. *Angew. Chem. Int. Ed.* **2009**, *48*, 7010-7012.
39. Bloch, W. M.; Holstein, J. J.; Hiller, W.; Clever, G. H., Morphological Control of Heteroleptic cis- and trans-Pd<sub>2</sub>L<sub>2</sub>L<sub>2</sub> Cages. *Angew. Chem. Int. Ed.* **2017**, *56*, 8285-8289.
40. Zhu, R. M.; Bloch, W. M.; Holstein, J. J.; Mandal, S.; Schafer, L. V.; Clever, G. H., Donor-Site-Directed Rational Assembly of Heteroleptic cis-[Pd<sub>2</sub>L<sub>2</sub>L<sub>2</sub>] Coordination Cages from Picolyl Ligands. *Chem-Eur J* **2018**, *24*, 12976-12982.
41. Li, R.-J.; Holstein, J. J.; Hiller, W. G.; Andréasson, J.; Clever, G. H., Mechanistic Interplay between Light Switching and Guest Binding in Photochromic [Pd<sub>2</sub>Dithienylethene<sub>4</sub>] Coordination Cages. *J. Am. Chem. Soc.* **2019**, *141*, 2097-2103.
42. Chen, B.; Holstein, J. J.; Horiuchi, S.; Hiller, W. G.; Clever, G. H., Pd(II) Coordination Sphere Engineering: Pyridine Cages, Quinoline Bowls, and Heteroleptic Binding One or Two Fullerenes. *J. Am. Chem. Soc.* **2019**, *141*, 8907-8913.
43. Lewis, J. E. M.; Crowley, J. D., Exo- and endo-hedral interactions of counteranions with tetracationic Pd<sub>2</sub>L<sub>4</sub> metallosupramolecular architectures. *Supramol. Chem.* **2014**, *26*, 173-181.
44. McNeill, S. M.; Preston, D.; Lewis, J. E. M.; Robert, A.; Knerr-Rupp, K.; Graham, D. O.; Wright, J. R.; Giles, G. I.; Crowley, J. D., Biologically active [Pd<sub>2</sub>L<sub>4</sub>]<sup>4+</sup> Quadruply-stranded Helicates: Stability and Cytotoxicity. *Dalton Trans.* **2015**, *44*, 11129-11136.
45. Jansze, S. M.; Wise, M. D.; Vologzhanina, A. V.; Scopelliti, R.; Severin, K., Pd<sup>II</sup>L<sub>4</sub>-type coordination cages up to three nanometers in size. *Chem. Sci.* **2017**, *8*, 1901-1908.
46. Chand, D. K.; Biradha, K.; Fujita, M., Self-assembly of a Novel Macrotricyclic Pd(II) Metallocage Encapsulating a nitrate ion. *Chem. Commun.* **2001**, 1652-1653.
47. Kishi, N.; Li, Z.; Yoza, K.; Akita, M.; Yoshizawa, M., An M<sub>2</sub>L<sub>4</sub> Molecular Capsule with an Anthracene Shell: Encapsulation of Large Guests up to 1 nm. *J. Am. Chem. Soc.* **2011**, *133*, 11438-11441.
48. Yamashina, M.; Sei, Y.; Akita, M.; Yoshizawa, M., Safe Storage of Radical Initiators within a Polyaromatic Nanocapsule. *Nat. Commun.* **2014**, *5*, 4662.
49. Marti-Centelles, V.; Lawrence, A. L.; Lusby, P. J., High Activity and Efficient Turnover by a Simple, Self-Assembled "Artificial Diels-Alderase". *J. Am. Chem. Soc.* **2018**, *140*, 2862-2868.
50. Schmidt, A.; Hollering, M.; Drees, M.; Casini, A.; Kühn, F. E., Supramolecular exo-functionalized palladium cages: fluorescent properties and biological activity. *Dalton Trans.* **2016**, *45*, 8556-8565.
51. we have found that the barrier to rotation around the acetylene bond is negligible (0.4 and 1.1 kcal mol<sup>-1</sup> for L 1 and L-2 respectively, cf. ethane ΔE<sub>rot</sub> = 2.7 kcal mol<sup>-1</sup> at the M06-2X/def2-TZVP level of theory, Figure S7).
52. Duarte, F.; Bauer, P.; Barrozo, A.; Amrein, B. A.; Purg, M.; Aqvist, J.; Kamerlin, S. C. L., Force Field Independent Metal Parameters Using a Nonbonded Dummy Model. *J. Phys. Chem. B* **2014**, *118*, 4351-4362.
53. August, D. P.; Nichol, G. S.; Lusby, P. J., Maximizing Coordination Capsule-Guest Polar Interactions in Apolar Solvents Reveals Significant Binding. *Angew. Chem. Int. Ed.* **2016**, *55*, 15022-15026.
54. Persson, I., Hydrated metal ions in aqueous solution: How regular are their structures? *Pure Appl Chem* **2010**, *82*, 1901-1917.
55. Li, P.; Merz, K. M., Metal Ion Modeling Using Classical Mechanics. *Chem. Rev.* **2017**, *117*, 1564-1686.
56. Zhao, C.; Sun, Q. F.; Hart-Cooper, W. M.; Dipasquale, A. G.; Toste, F. D.; Bergman, R. G.; Raymond, K. N., Chiral Amide Directed Assembly of a Diastereo- and Enantiopure Supramolecular Host and its Application to Enantioselective Catalysis of Neutral Substrates. *J. Am. Chem. Soc.* **2013**, *135*, 18802-18805.
57. Sure, R.; Grimme, S., Comprehensive Benchmark of Association (Free) Energies of Realistic Host-Guest Complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3785-3801.
58. Bootsma, A. N.; Wheeler, S., Popular Integration Grids Can Result in Large Errors in Dft-computed Free Energies. *ChemRxiv. Preprint*. **2019**.
59. Grimme, S., Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chem. Eur. J.* **2012**, *18*, 9955-9964.
60. Caldararu, O.; Olsson, M. A.; Riplinger, C.; Neese, F.; Ryde, U., Binding free energies in the SAMPL5 octa-acid host-guest challenge calculated with DFT-D3 and CCSD(T). *J. Comput. Aid. Mol. Des.* **2017**, *31*, 87-106.
61. G<sub>bind</sub>=E<sub>bind</sub>+ZPE+PAV-TAS. The last three terms are obtained from frequency calculations assuming an ideal gas model (IGM).
62. 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.
63. Rizzi, A.; Murkli, S.; McNeill, J. N.; Yao, W.; Sullivan, M.; Gilson, M. K.; Chiu, M. W.; Isaacs, L.; Gibb, B. C.; Mobley, D. L.; Chodera, J. D., Overview of the SAMPL6 Host-Guest Binding Affinity Prediction Challenge. *J. Comput. Aid. Mol. Des.* **2018**, *32*, 937-963.

64. Stewart, J. J. P., Optimization of Parameters for Semiempirical Methods VI: more Modifications to the NDDO Approximations and re-Optimization of Parameters. *J. Mol. Model.* **2013**, *19*, 1-32.
65. Grimme, S.; Bannwarth, C.; Shushkov, P., A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements ( $Z = 1-86$ ). *J. Chem. Theory Comput.* **2017**, *13*, 1989-2009.
66. Uliana, M. P.; Servilha, B. M.; Alexopoulos, O.; de Oliveira, K. T.; Tormena, C. F.; Ferreira, M. A. B.; Brocksom, T. J., The Diels-Alder Reactions of para-benzoquinone Nitrogen-Derivatives: an Experimental and Theoretical Study. *Tetrahedron* **2014**, *70*, 6963-6973.
67. F. Fringuelli; Taticchi, A., *The Diels-Alder Reaction: Selected Practical Methods*. John Wiley: New York, 2002; p i-xvi.
68. Corey, E. J., Catalytic Enantioselective Diels-Alder Reactions: Methods, Mechanistic Fundamentals, Pathways, and Applications. *Angew. Chem., Int. Ed.* **2002**, *41*, 1650-1667.
69. Liu, Z.; Lee, J. H. Q.; Ganguly, R.; Vidović, D., A Well-Defined Aluminum-Based Lewis Acid as an Effective Catalyst for Diels-Alder Transformations. *Chem-Eur J* **2015**, *21*, 11344-11348.
70. Marti-Centelles, V.; Duarte, F.; Lusby, P. J., Host-Guest Chemistry of Self-Assembled Hemi-Cage Systems: The Dramatic Effect of Lost Pre-Organization. *Isr. J. Chem.* **2018**, *59*, 257-266.
71. Ess, D. H.; Houk, K. N., Distortion/interaction energy control of 1,3-dipolar cycloaddition reactivity. *J. Am. Chem. Soc.* **2007**, *129*, 10646-10647.
72. Schrödinger Release 2013-3: MacroModel version 10.2. Schrödinger LLC. New York 2013.
73. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, T.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., *Gaussian 09, Revision D.01*. **2013** Gaussian Inc., Wallingford CT.
74. van der Spoel, D.; van Maaren, P. J.; Caleman, C., GROMACS molecule & liquid database. *Bioinformatics* **2012**, *28*, 752-753.
75. Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D., Force field benchmark of organic liquids: density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant. *J. Chem. Theory Comput.* **2012**, *8*, 61-74.
76. M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GROMACS development team. GROMACS User Manual version 2019.1 <http://www.gromacs.org>.
77. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*:19-25.
78. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald: An  $N \cdot \log(N)$  Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.
79. Neese, F., Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1327.
80. Zhao, Y.; Truhlar, D. G., The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Gour M06-class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
81. Weigend, F.; Ahlrichs, R., Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.
82. Weigend, F., Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057-1065.
83. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378-6396.
84. Adamo, C.; Barone, V., Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158-6170.
85. Johnson, E. R.; Becke, A. D., A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. *J. Chem. Phys.* **2006**, *124*, 174104.
86. Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E., Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
87. Zhao, Y.; Truhlar, D. G., Design of Density Functionals That Are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. *J. Phys. Chem. A* **2005**, *109*, 5656-5667.
88. 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.
89. Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F., Sparse maps—A Systematic Infrastructure for Reduced-Scaling Electronic Structure Methods. II. Linear Scaling Domain Based Pair Natural Orbital Coupled Cluster Theory. *J. Chem. Phys.* **2016**, *144*, 024109.
90. Glendenning, E. D.; Landis, C. R.; Weinhold, F., NBO 6.0: Natural bond orbital analysis program. *J. Comput. Chem.* **2013**, *34*, 1429-1437.
91. Lu, T.; Chen, F., Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580-592.

### Accurate Methodology



### Reactivity Insight



Dynamics · Flexibility · Conformations · Solvent Effects  
Binding Affinity · Transition States · Catalytic (in)activity