Electronic Supplementary Information

Efficient and Flexible Approach for Local Distortion: Distortion Distribution Analysis enabled by Fragmentation

Zeyin Yan,§ Yunteng Sam Liao,§ Xin Li, and Lung Wa Chung*

Shenzhen Grubbs Institute, Department of Chemistry and Guangdong Provincial Key Laboratory of Catalysis, Southern University of Science and Technology, Shenzhen 518055, China

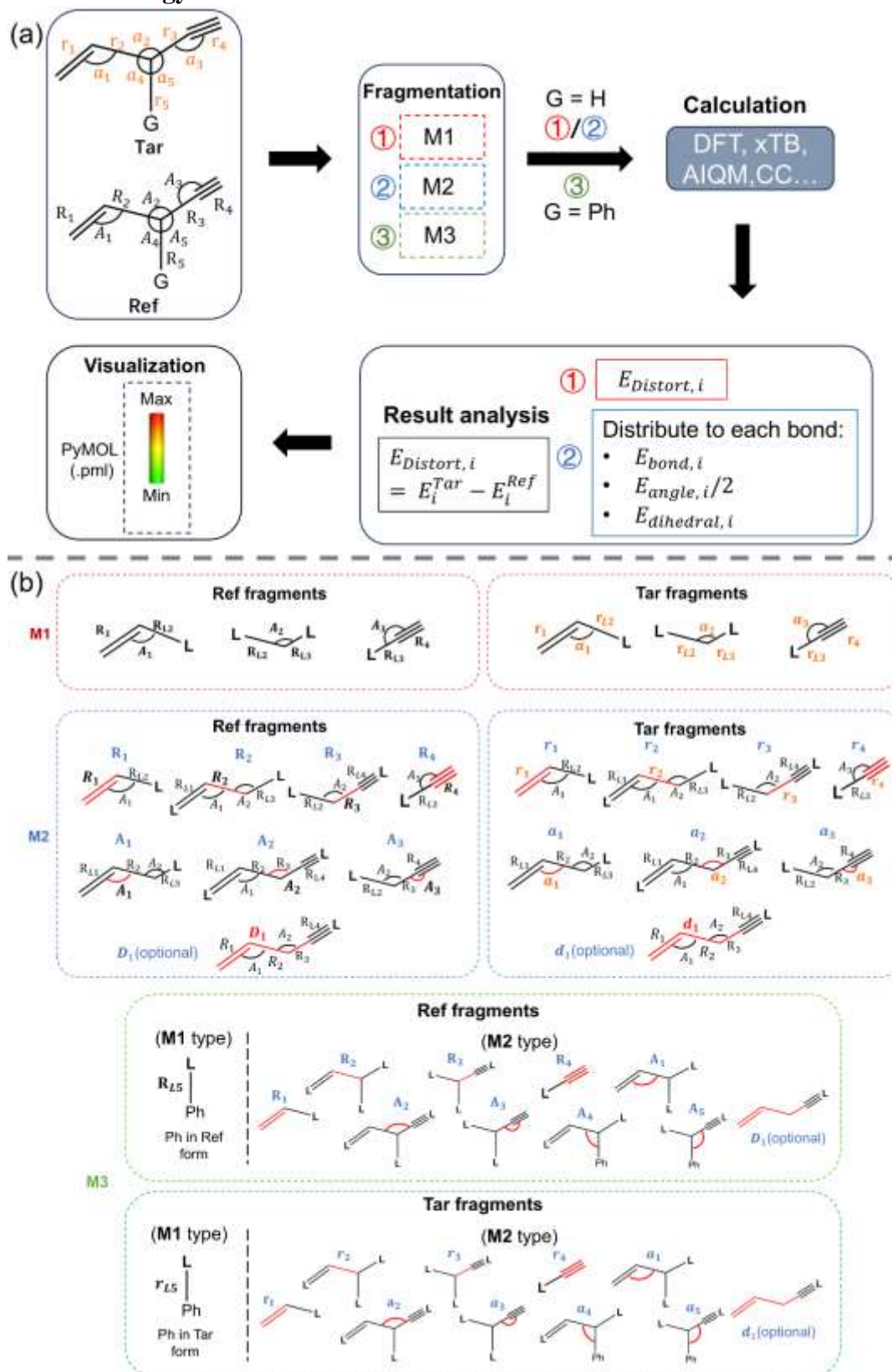§These authors contributed equally to this work

*E-mail: oscarchung@sustech.edu.cn

Table of Contents

1. Methodology details

Scheme S1 (a) Our workflow of distortion distribution analysis with fragmentation (D2AF). (b) Fragmentation examples for three approaches. **M3** scheme is a hybrid approach combined with **M1** and **M2** schemes, in which the only sole change of the coordinate label for the **M2**-type fragments is shown for clarity.

4

As shown in Scheme S1a, our fragmentation-based approach consists of three stages: fragmentation, calculation, and analysis. The selected target system (denoted as **Tar**, with more distortion) and its reference configuration (denoted as **Ref**, usually with less distortion) were used as input for D2AF. Distortion energy is the energy difference between the reference configuration and its target counterpart. The local distortion within a molecule can be analyzed by fragmenting the systems using one of three fragmentation and coordinate manipulation approaches, depending on the system's complexity (see below).

Method 1 (M1) Fragmentation/Coordinate scheme: In many-body expansion approximation used in prevalent fragmentation schemes,[S1, 2] the energy of a specific molecule can be expressed from the inclusion-exclusion principle as follows (eq. S1):

$$E_{Sys} \approx \sum_i E_i + \sum_{i,j} E_{ij} + \sum_{i,j,k} E_{ijk} + \cdots \quad \text{(S1)}$$

Where E_i, E_{ij} represent the one-body energy, two-body energy contribution, respectively. The energy difference between the whole **Tar** and **Ref** molecules can be decomposed by many-body expansion approximation as follows (eq. S2):

$$E_{Distort} \approx \sum_i (E_i^{Tar} - E_i^{ref}) + \sum_{i,j} (E_{ij}^{Tar} - E_{ij}^{ref}) + \sum_{i,j,k} (E_{ijk}^{Tar} - E_{ijk}^{ref}) \ldots \text{(S2)}$$

In this method (**M1**), the **Tar** and **Ref** molecules were divided into smaller fragments as much and reasonable as possible to get higher "resolution" of distortion map and reduce the computational costs (Scheme S1b). As we focus on the distortion energy of each subsystem/fragment (i.e., local distortion), the one-body energy part is considered only. Therefore, local distortion is estimated as the distortion energy of each fragment, i.e., the difference between the homologous fragment of the **Tar** and **Ref** systems (eq. S3).

$$E_{Distort,i} = E_i^{Tar} - E_i^{ref} \quad \text{(S3)}$$

To attain a higher resolution of distortion map, each system can be generally fragmentated in the smallest and reasonable species, consisting of a single heavy atom

as well its link atoms in most cases. Alternatively, users can customize the system's partition to expand the size of particular fragement(s) to include complex and key interactions (such as delocalization in conjugate aromatic or alkenyl/alkynyl groups, lone pairs repulsions). Consequently, **M1** typically cuts all single bonds of the **Tar** and **Ref** molecules except X-H bonds (X denotes a heavy atom), which serve as boundaries between fragments (Scheme S1b). For instance, as shown in Scheme S1b, both **Ref** and **Tar** molecules are partitioned into three fragments, each preserving the bonds and angles characteristic of its corresponding state. All boundaries of each fragment containing dangling bonds are also capped by link atoms (L).

When the system contains fused rings (e.g. indole in Figures S28 and S58), in order to maintain the minimum-size ring structure in **M1** scheme, the overlapped atoms of the rings may appear in multiple separated ring fragments. In the visualization, these overlapped/shared atoms are colored according to the last fragment list (including these shared atoms) defined in the input file.

Method 2 (M2) Fragmentation/Coordinate scheme: Inspired by molecular mechanics energy expression (eq. S4), an alternative perspective on decomposing the distortion energy into three bonding contributions. Within this molecular mechanics framework, the total energy of the system is conceptualized as follows:

$$E_{Sys} = E_{bond} + E_{angle} + E_{dihedral} + E_{elec} + E_{VDW} \qquad \text{(S4)}$$

Where E_{bond}, E_{angle} and $E_{dihedral}$ are bonding components contributed from each bond, angle and dihedral energy; E_{elec} and E_{VDW} are non-bonding/noncovalent components contributed from each electrostatic and van der Waals interaction. Only the bonding components were considered in the **M2** scheme, as the local distortion energy should mainly be dictated by the three bonding components and focus on the local geometrical effects. Therefore, the (relative) distortion energy of the target system (**Tar**) referring to its **Ref** counterpart can be approximated as follows (eq. S5):

$$\Delta E_{Distort} \approx \Sigma\Delta E_{bond} + \Sigma\Delta E_{angle} + \Sigma\Delta E_{dihedral} \qquad \text{(S5)}$$

In this method (**M2**), we aim to estimate these above energy changes by altering only one bonding internal coordinate (such as bond, angle and (optionally) dihedral) from its **Ref** to **Tar** state while keeping the other coordinates same as in **Ref** (Scheme S1c). To achieve such distortion decomposition, fragmentation including two, three or four heavy atoms (with their link atoms) for one target internal coordinate (bond, angle or dihedral, respectively) taken from the reference form (R_i, bond/ A_i, angle/ D_i, dihedral) is performed to generate the **Ref** fragments. Then, only one target internal coordinate ($r_i/a_i/d_i$) of each generated fragment is altered to be identical as that coordinate value in the target system to set up all combinations of the **Tar** fragments (Scheme S1c).

Scheme S1b illustrates all fragments obtained to reveal each individual local distortion energy contribution by altering one specific internal coordinate. Therefore, the **Tar** fragments differ from their corresponding **Ref** fragments by the one altered internal coordinate ($R_i/A_i/D_i$ *vs* $r_i/a_i/d_i$): each coordinate change corresponds to individual local distortion energy contribution for one specific bond, angle or dihedral. Accordingly, in contrast with the **M1** method, local distribution of each bonding coordinate within one fragment is evaluated individually and should not rely on the LA approximation in the **M2** method.

Method 3 (M3) Fragmentation/Coordinate scheme: Using **M2** can elucidate the intricate relationship between (relative) energy changes and alterations in bonds/angles (/optional dihedrals). This approach provides an intuitive and high-resolution perspective of (relative) local distortion energy. However, this approach exhibits limitations to apply to cyclic conjugated groups containing delocalized electrons (such as π electrons) in a ring. **M2** encounters challenges in strictly decoupling the conjugated moiety into bonds and angles in a ring by altering the only one specific bonding coordinate in the **Tar** form and keeping its remaining coordinates unchanged as in the **Ref** form. Therefore, when dealing with such complex systems, a hybridized partition method (refer as **M3**) was applied for two parts: a conjugated moiety using the **M1** method and the rest part(s) using **M2** (Scheme S1c). As shown in Scheme S1b, when the functional group (G) is aromatic Ph, the Ph group is categorized into the **M1** part,

and the rest part(s) is(are) treated using **M2**.
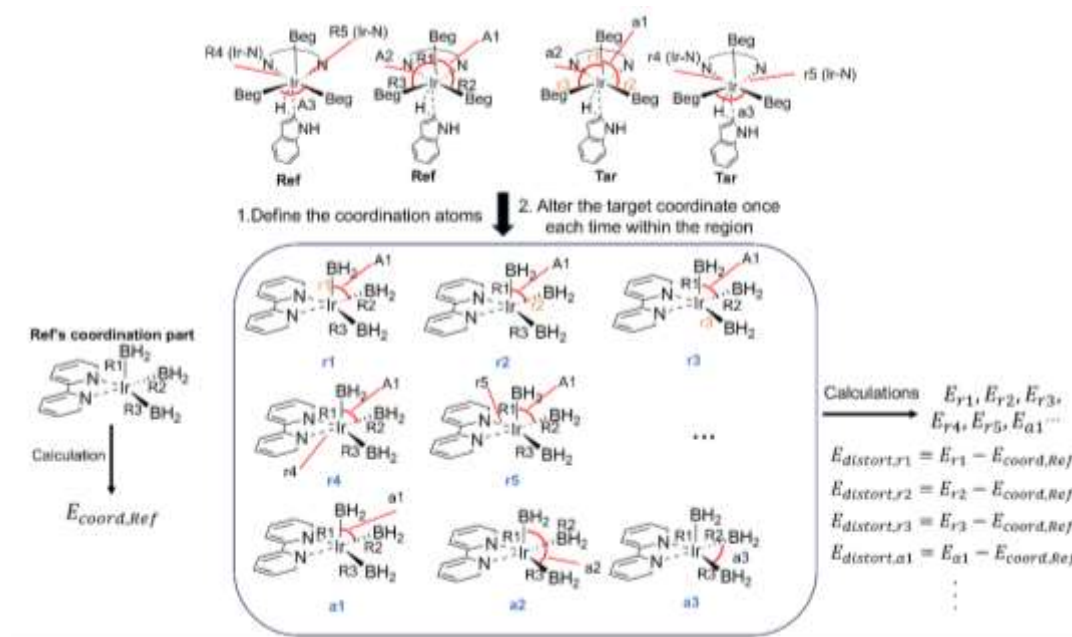
Scheme S2 Schematic fragmentations for metal coordination systems by the **M3** scheme.

Moreover, according to ligand field theory, metal coordination compounds exhibit distinctive and complex metal-ligand interactions. The key orbitals for this system are generally influenced and sensitive by ligands (coordinate numbers/elements), directly reflecting the electronic structure around the transition metal. Therefore, maintaining consistent metal-ligand interactions becomes crucial when analyzing distortions related to the transition metal compounds. To tackle this challenge, we propose a novel feature in **M3**, in which users can define a minimum and reasonable metal coordination region as one special fragment to capture the key metal-ligand interactions (Scheme S2) using **M2**. The **Ref** fragment was generated including all atoms in this special fragment (coordination region with their link atom(s)) taken from the reference form. Then, only one target internal coordinate in the coordination region is altered to be identical as to that coordinate value in the target system to set up all combinations of the **Tar** fragments. Accordingly, in local distribution distortion of each bonding coordinate within the coordination region is evaluated individually. After calculations to obtain energy of these fragments, (relative) local distortion energy of each individual target M-L bond can be estimated. It should be noted that this coordinate treatment cannot work perfectly if any large coordinate change involving the metal-ligand chelation is

8

involved between the **Ref** and **Tar** cooridinate fragment. This approach should help facilitate our analysis of local distortion distribution in metal coordination compounds.

Likewise, for the system contains fused rings (e.g. indole in Figures S28 and S58) in **M3** scheme, the overlapped atoms of the rings may appear in multiple separated ring fragments. In the visualization, these overlapped/shared atoms are colored according to the last fragment list (including these shared atoms) defined in the input file.

Link-atom Treatment: When any fragment generated by any one of the fragmentation schemes (**M1**-**M3**) contain any dangling bonds, the resulting dangling bonds should be saturated using the ONIOM-type link atom (LA, denoted as L in Scheme S1c) approach. Hence, the position of the link atom is determined by the following formula (eq. S6). Such an LA approach can project local distortion of the bonding boundary to the link atom in **M1**.

$$\mathbf{R}_{LA} = \mathbf{R}_A + g(\mathbf{R}_B - \mathbf{R}_A) \tag{S6}$$

where A is the link-atom connection (LAC, belonging to the fragment), B is the link-atom host (LAH), g is the scaling factor, $\mathbf{R}_{LA}$ is the position vector of the link atom. Following our benchmark analysis of various molecular models, hydrogen and nitrogen atoms are generally selected to cap single and triple dangling bonds, respectively. Additionally, the carbon atom was generally used for double dangling bond, except only one atom as the double- or triple-bond LAH (without other connection) to maintain a consistent chemical environment.

2. Computational details

For the LA benchmark calculations, the realistic molecules were first fully optimized and then rigid scan calculations (using 'scan' keyword in Gaussian16) or relaxed scan calculations (using 'opt=modredundant" and "scan' keywords in Gaussian16) using M06-2X/6-31G(d) and/or B3LYP-D3/6-31G(d) methods. For each bond's scanning, a total of 14 steps with the step size of 0.05Å were taken. For each angle's bending scanning, a total of 12 steps with the step size of 5° strategy were taken. Moreover, structures from these rigid/relaxed scan calculations of the realistic systems were taken and then performed with the link-atom replacement and single-point energy calculations using the same computational method to obtain the related energy profiles using the LA approximation.

All these distortion calculations were performed using our open-source python package (D2AF, https://github.com/oscarchung-lab/D2AF), which employs Open Babel package to treat internal coordinates of fragments.[S3] Several interfaces of calculators such as Gaussian16 for DFT,[S4] GFN2-xTB as semi-empirical method (SE),[S5] AIQM1,[S6] and ANI-series[S7, 8] as MLPs, ORCA5.0 for CCSD(T) can also be used for energy computations.[S9, 10] Scripts for PyMOL[S11] were finally generated by our code to visualize (relative) distortion distribution (or so-called distortion map). The choice of the DFT method and basis sets as well as structures for each application system (unless stated otherwise) were taken from the previous publications, which are discussed in the below sections 4-5.

3. LA Benchmark details

A few possible choices of link atoms for single-, double- and triple-bond as well as special cases in our selected systems were systematically investigated and conducted to see the effect of link atoms on energetic profiles.
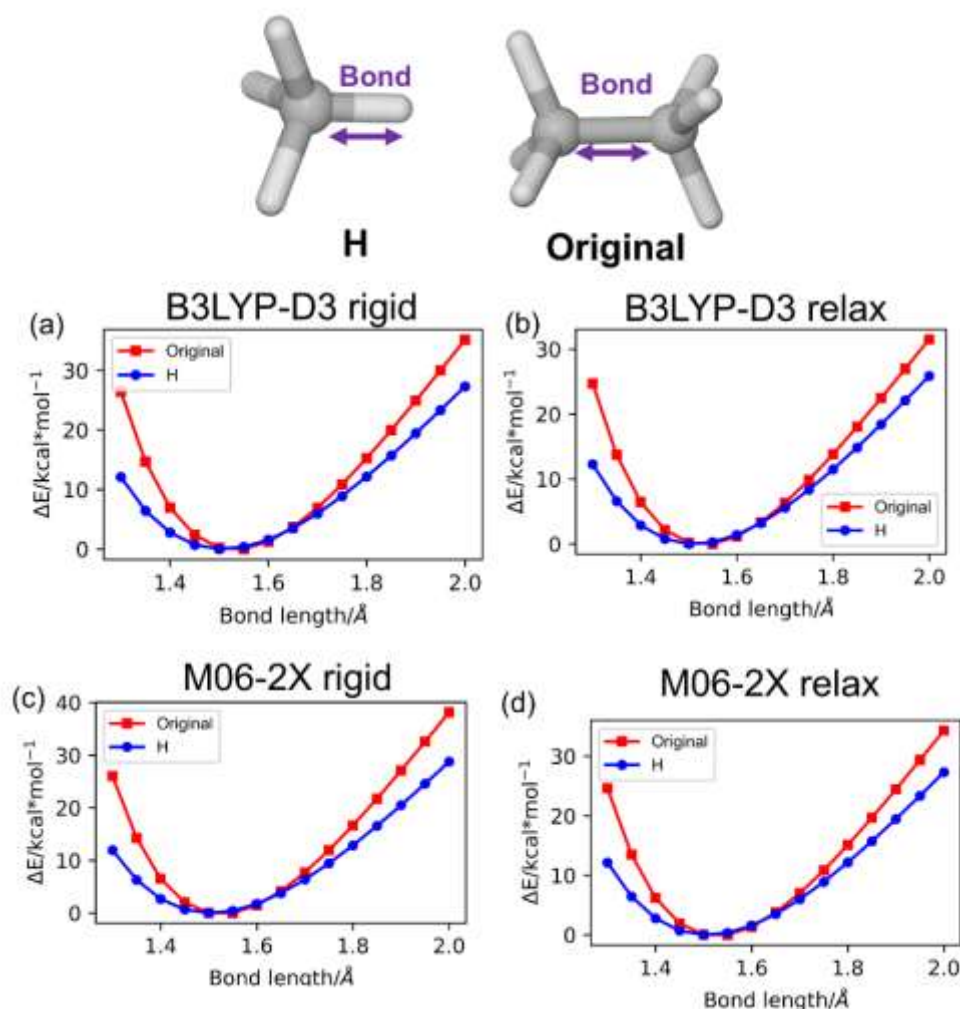
3.1. Single-bond LA and boundaries

Figure S1 Benchmark of link-atom approximation for C-C single bond. CH_3-CH_3 as real system and its corresponding model (CH_4) with H as the link atom were used to get the energy profile for their varying bond length (C-H vs C-C). (a, c) Rigid scan and (b, d) relaxed scan from 1.3 Å to 2.0 Å. All the energy profiles derived from scan calculations of the real system followed by replacement by the link atom and single-point energy calculations computed at (a, b) B3LYP-D3/6-31G(d) and (c, d) M06-2X/6-31G(d) levels.

Figure S2 Benchmark of link-atom approximation for C-C single bond. C_3H_8 as real system and its corresponding model (CH_4) with H as the link atoms were used to get the energy profile for their varying angle (H-C-H vs C-C-C). (a, c) Rigid scan and (b, d) relaxed scan from from 90° to 150°. All the energy profiles were derived from scan calculations of the real system followed by replacement by the link atom and single-point energy calculations computed at (c, d) B3LYP-D3/6-31G(d) and (a, b)M06-2X/6-31G(d) levels.
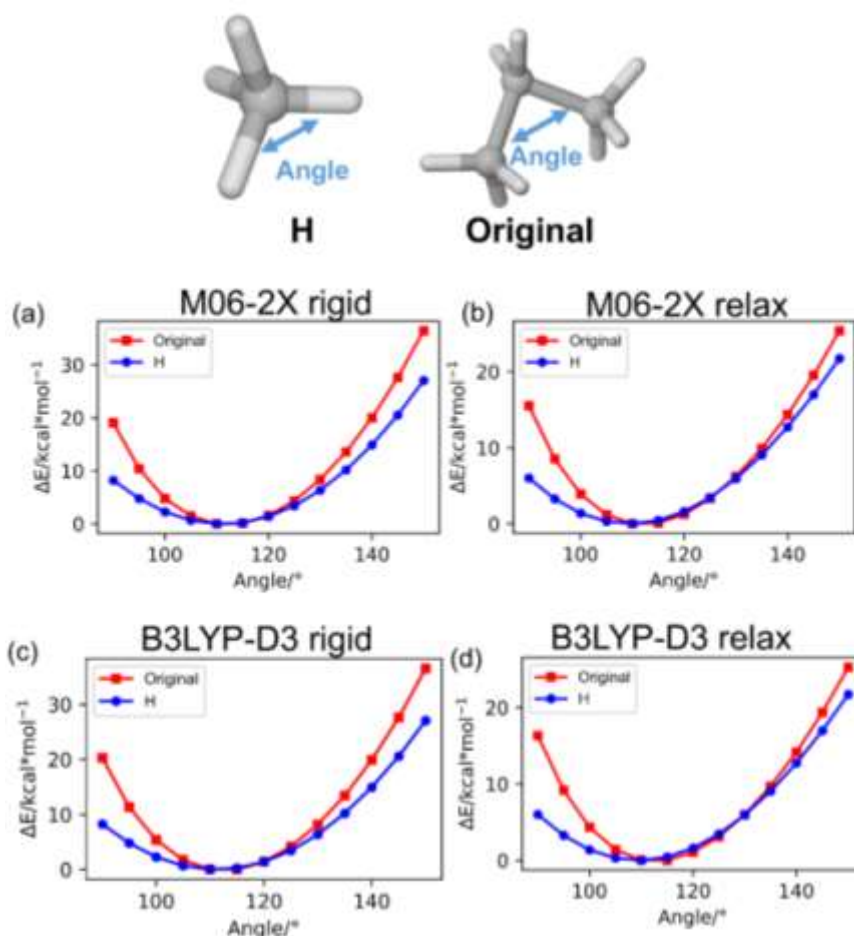
As shown in Figure S4S2, the common H-LA for the single-bond boundary was first assessed and can qualitatively reproduce the potential energy curve for a realistic C-C bond and C-C-C angle at B3LYP-D3/6-31G(d) and M06-2X/6-31G(d) levels. When the C-C bond changes from equilibrium, the H-LA begins to increasingly underestimate the C-C bond energy curve, (Figure S1: with the error of ~5.6 and ~9.4 kcal/mol at 2.0 Å for rigid and relaxed scans, respectivly.). Similarly, when the C-C-C angle bending from equilibrium, the H-LA begins to increasingly underestimate the C-C-C angle energy curve, (Figure S2: with the error of ~9.4 and ~3.7 kcal/mol at 150° for rigid and relaxed scan, respectivly.). Generally, H-LA can qualitatively replicate the overall behavior of the original potential energy curve.
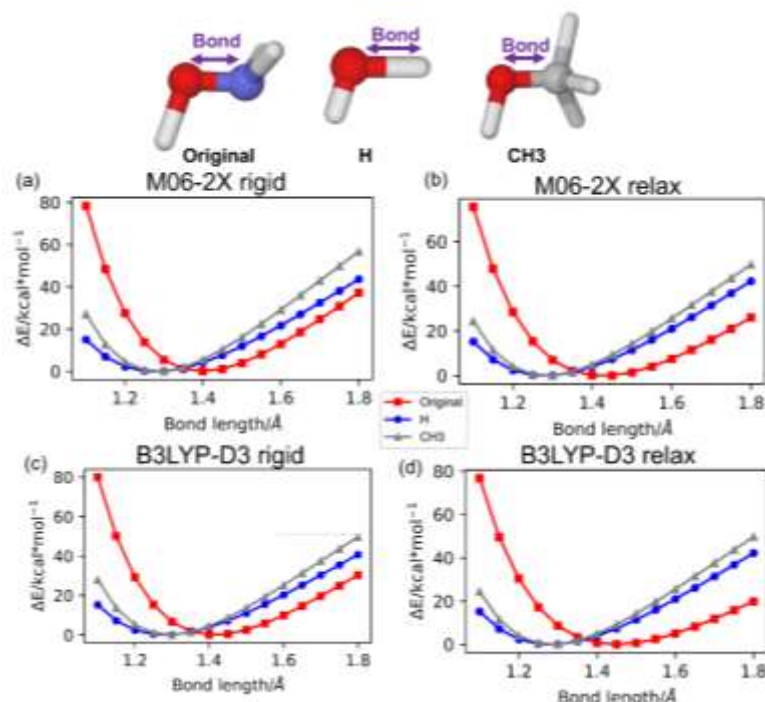
12

Figure S3 Benchmark of link-atom approximation for the O-N single bond in **S1** case (Figure 3). $HONH_2$ (a fragment in **M1**) as real model and its scaled model systems (HOH, $HOCH_3$) with H and CH_3 as the link atom, respectively. (a, c) Rigid scan and (b, d) relaxed scan from 1.1Å to 1.8Å. Energy profiles for the varying bond length derived from (a, c) rigid scan and (b, d) relaxed scan calculations of the real system followed by replacement by the link atom and single-point energy calculations computed at (a, b) M06-2X/6-31G(d) and (c, d) B3LYP-D3/6-31G(d) levels.
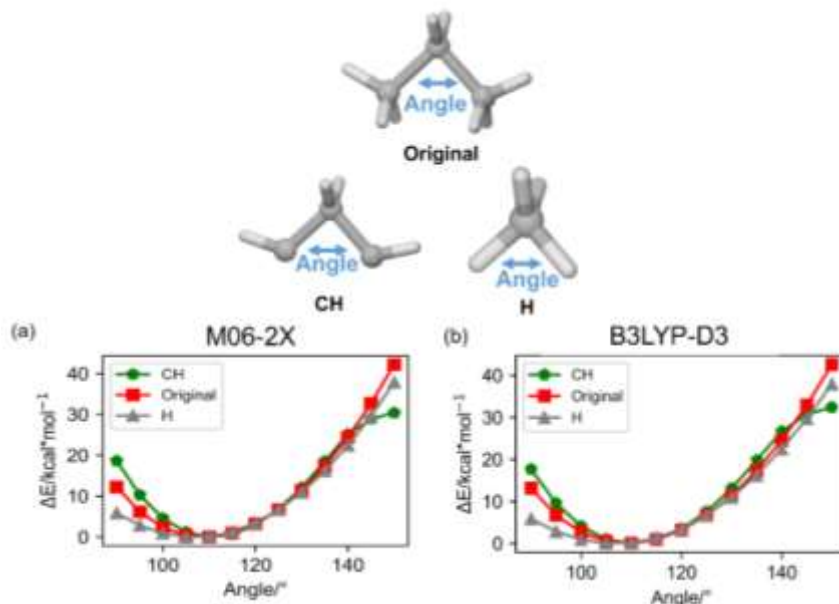

Figure S4 Benchmark of link-atom approximation for C-C single bond in the fragment **I** of the [4+2] Diels-Alder addition reaction (Figure 4). C_3H_8 as the original system and its correspongding models (CH_4 and $CH_2(CH)_2$) with H and CH as the link atom, respectively. Energy profiles for the varying angle derived from rigid scan calculations of the real system followed by replacement by the link atom and single-point energy calculations computed at (a) M06-2X/6-31G(d) and (b) B3LYP-D3/6-31G(d) levels.

For the LA testing on the C-C-C bending of fragment **I** of the [4+2] Diels-Alder

addition reaction, H or CH as the link atom were applied (Figure S4). The H or CH as the link atom can reasonable reproduce the energy profiles, especially for the increased C-C-C angle bending. However, the H and CH as the link atom modestly under- and over-estimates the energies for the small C-C-C bending.
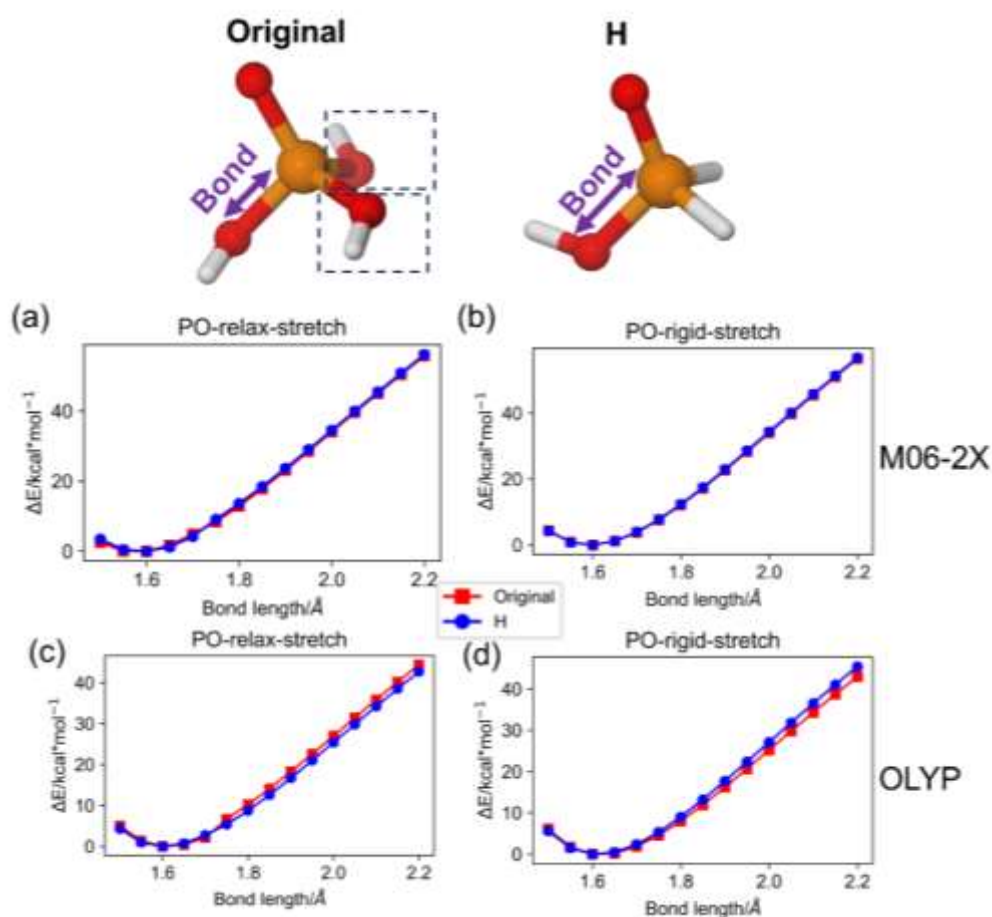

Figure S5 Link atom benchmark results of the P-O single bond from H_3PO_4 (fragment in case **SS2**) as the original system and its correspongding model $O=PH_2(OH)$ with H as the link atom. Energy profiles for the varying bond length derived from relaxed (a, c) and rigid (b, d) scan calculations of the real system followed by replacement by the link atom and single-point energy calculations at (a, b) CPCM M06-2X/def2-TZVP and (c, d) CPCM OLYP/def2-TZVP levels.

Figure S6 Link atom benchmark results of the P-O single bond from H_3PO_4 (fragment in case **SS2**) as the original system and its correspongding model $O=PH(OH)_2$ with H as the link atom. Energy profiles for the varying bond length were derived from relaxed (a, c) and rigid (b, d) scan calculations of the real system followed by replacement by the link atom and single-point energy calculations at (a, b) CPCM M06-2X/def2-TZVP and (c, d) CPCM OLYP/def2-TZVP levels.
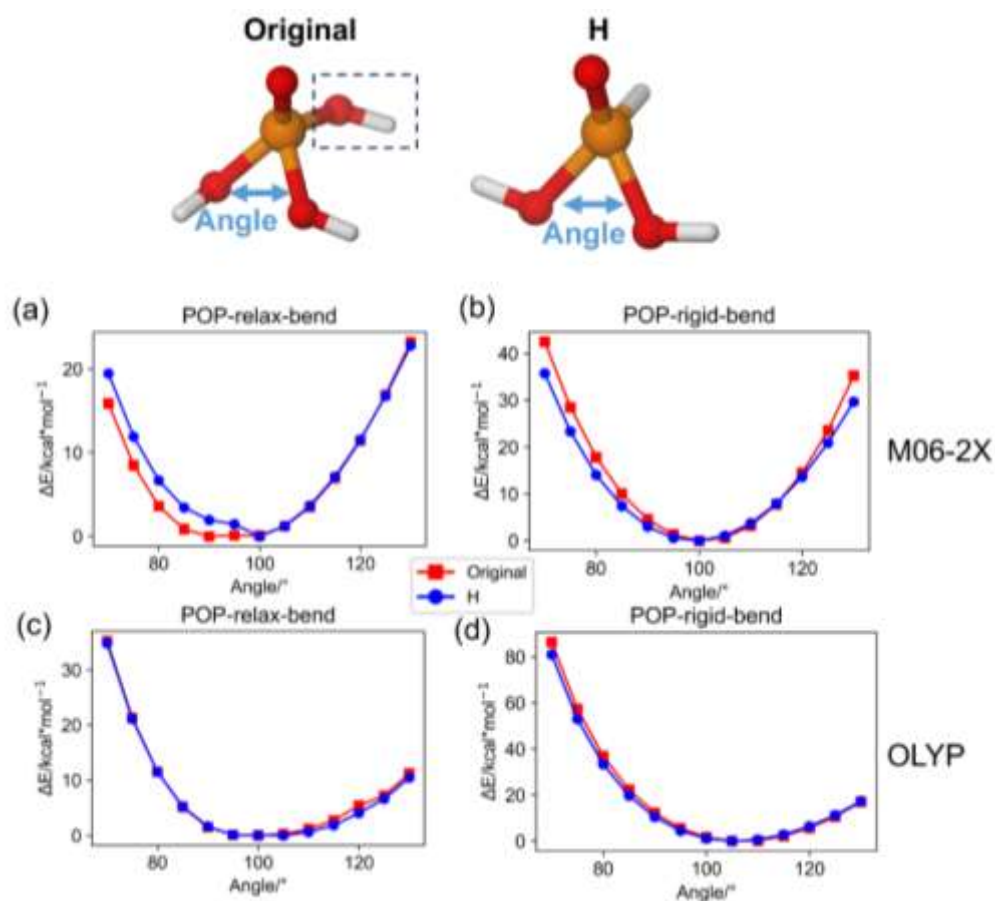
Figure S7 Link atom benchmark results of the P-O single bond from H_3PO_4 (fragment in case **SS2**) as the original system and its correspongding model $O=PH_2(OH)$ with H as the link atom. Energy profiles for the varying angle were derived from relaxed (a, c) and rigid (b, d) scan calculations of the real system followed by replacement by the link atom and single-point energy calculations at (a, b) CPCM M06-2X/def2-TZVP and (c, d) CPCM OLYP/def2-TZVP levels.
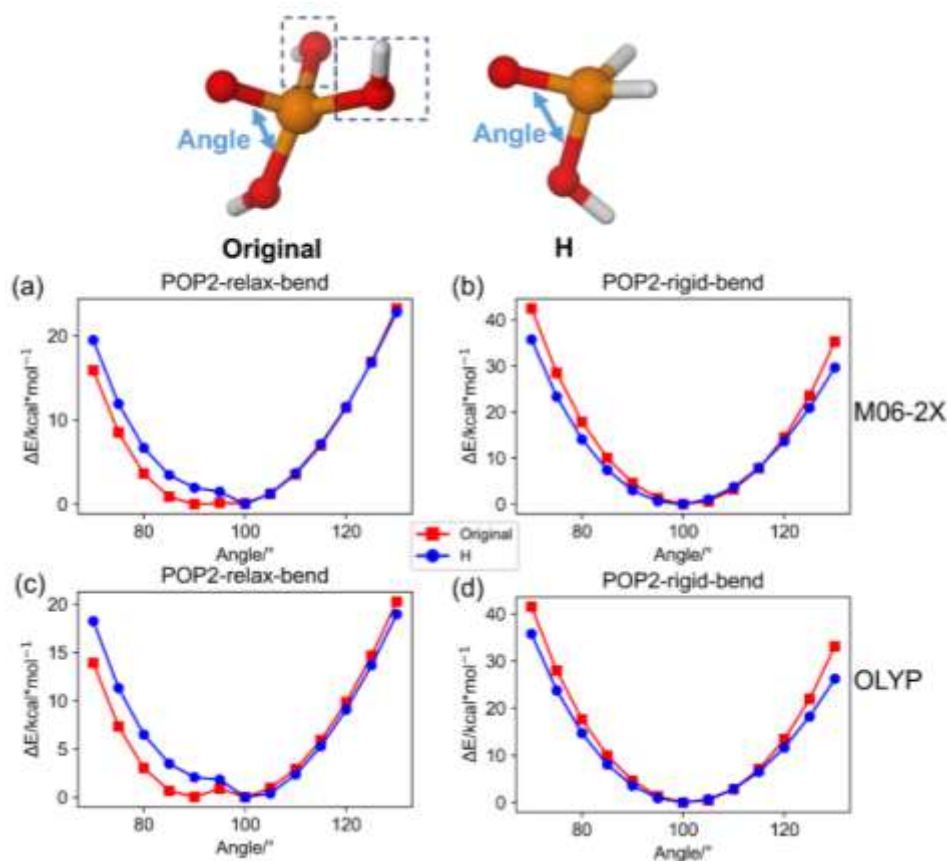
Figures S5-S7 show the benchmark results of link-atom approximation for P-O single bond in case of triphosphate hydrolysis via S_N2 (**SS2**, Figure S40). It reveals the energy profile for the varying bond length and angles can be reasonably reproduced using H-LA.
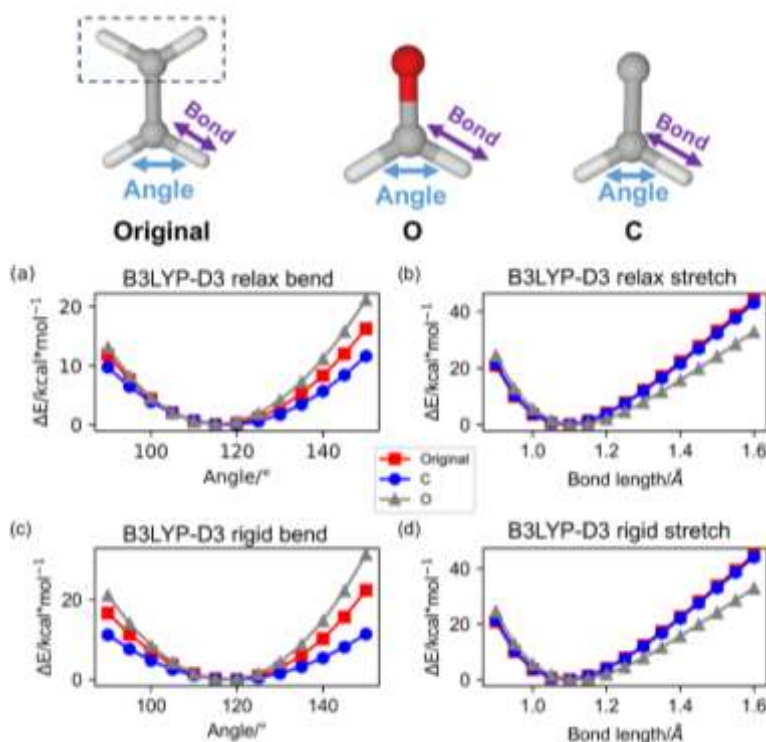
3.2. Double-bond LA and boundaries

Figure S8 Benchmark of link-atom approximation for C-C double bond. Ethene as the original system and its two corresponding models ($CH_2=O$, $CH_2=C$) with either C (radical) or O as the link atom, respectively. Energy profiles for their varying bond length (b, d) and angle (a, c) derived from relaxed (a, b) or rigid (c, d) scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at B3LYP-D3/6-31G(d) level.

Figure S9 Benchmark of link-atom approximation for C-C double bond. Ethene as the original system and its two corresponding models ($CH_2=O$, $CH_2=C$) with either C (radical) or O as the link atom, respectively. Energy profiles for their varying bond length (b, d) and angle (a, c) derived from relaxed (a, b) or rigid (c, d) scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at M06-2X/6-31G(d) level.
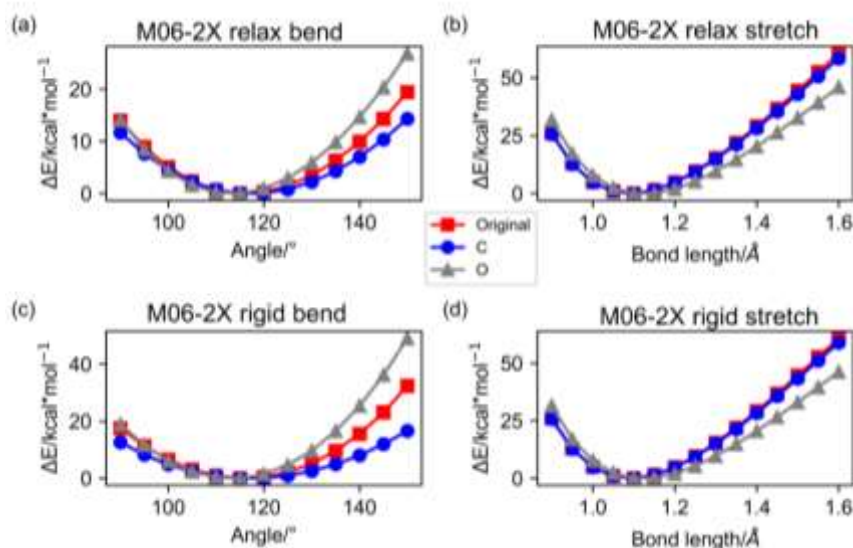
Similar to the results and discussion in Figures 2a-2c, our additional results further support that C-LA exhibits a smaller energy deviation than O-LA for the C-H bond change (Figures S8b,d and S9b,d). In addition, O-LA and C-LA give a large error for the large H-C-H angle bending, but error become smaller for small H-C-H bond angles (Figures S8a,c and S9a,c).

Figure S10 Benchmark of link-atom approximation for C-N double bond. $CH_2=NCH_3$ as the original system and its two corresponding models ($CH_3N=C$, $CH_3N=O$) with either C (radical) or O as the link atom, respectively. Energy profiles for their varying bond length derived from relaxed (a, c) or rigid (b, d) scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at B3LYP-D3/6-31G(d) (a, b) and M06-2X/6-31G(d) (c, d) levels.
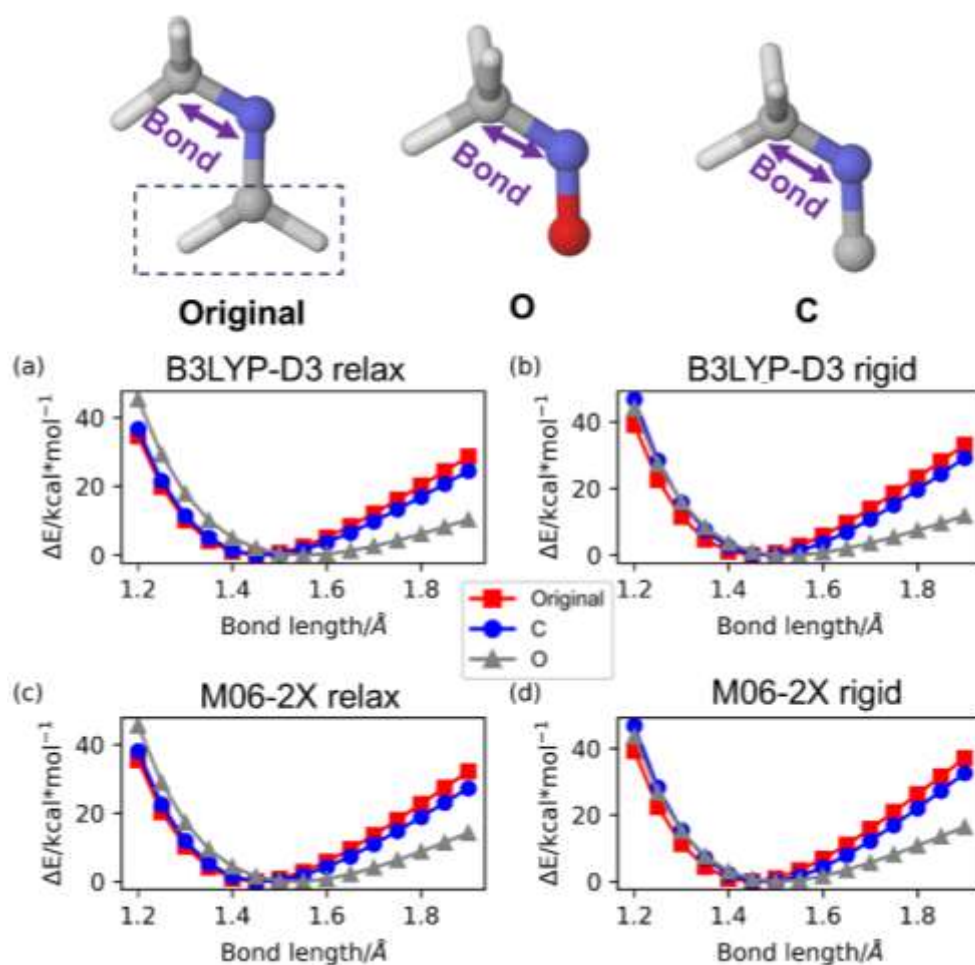
Figure S11 Benchmark of link-atom approximation for C-N double bond. $C_2H_5N=CH_2$ as the orginal system and its two corresponding models ($C_2H_5N=O$, $C_2H_5N=C$) with either O or C (radical) as the link atom, respectively. Energy profiles for their varying angle derived from (a,c) relax and (b,d) rigid scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at M06-2X/6-31G(d) and B3LYP-D3/6-31G(d) level.

Figure S12 Benchmark of link-atom approximation with two conjugating C-C double bonds as boundaries. $CH_2=CH-CH=CH_2$ as the original system and its corresponding models ($CH_2=CH-CH=O$ for angle bending/$O=CH-CH=O$ for bond stretching, $CH_2=CH-CH=C$ for angle bending/$C=CH-CH=C$ for bond stretching) with either O or C (radical) as the link atom, respectively. Energy profiles for their varying angle (a, c) and bond length (b, d) derived from (a, b) relaxed and (c, d) rigid scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at B3LYP-D3/6-31G(d) level.
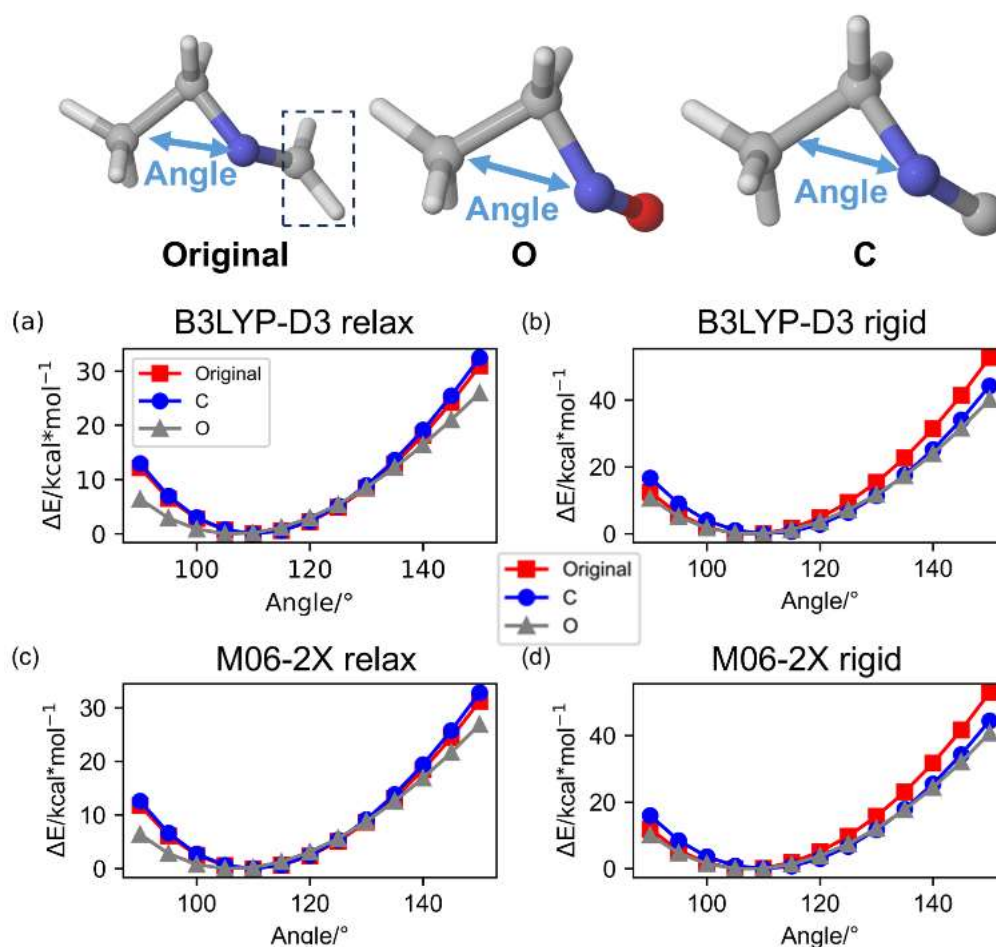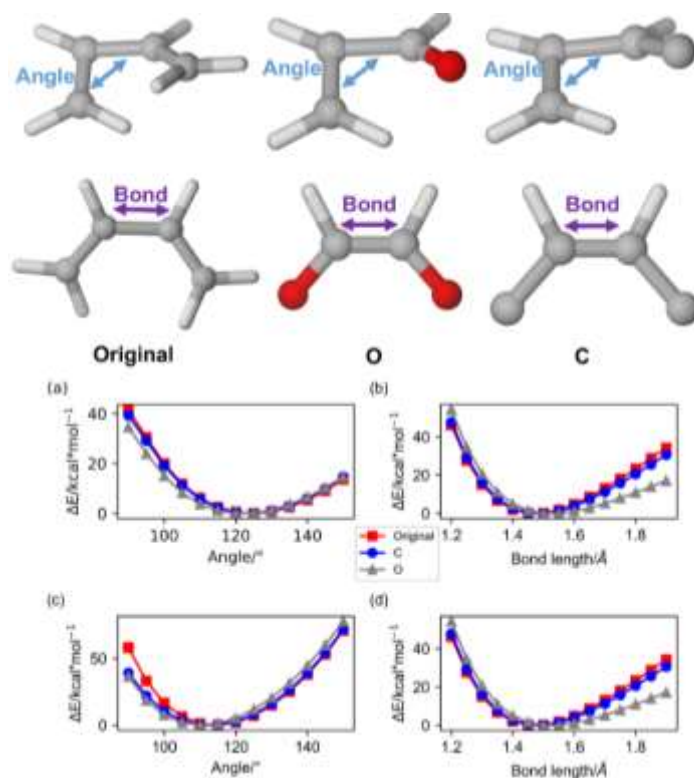

Figure S13 Benchmark of link-atom approximation with two conjugating C-C double bonds as boundaries. $CH_2=CH-CH=CH_2$ as the original system and its corresponding models ($CH_2=CH-CH=O$ for angle bending/$O=CH-CH=O$ for bond stretching, $CH_2=CH-CH=C$ for angle bending/$C=CH-CH=C$ for bond stretching) with either O or C (radical) as the link atom, respectively. Energy profiles for their varying angle (a, c) and bond length (b, d) are derived from (a, b) relaxed and (c, d) rigid scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at M06-2X /6-31G(d) level

20

In addition, C-LA can be a good LA for the conjugating C-C double bonds as it generally gives reasonable energetics to approximate changes in the C-C bond stretching (Figures S12b, d and S13b, d) and C-C-C angle bending (Figures S12a, c and S13a, c). Furthermore, C-LA and O-LA can generally give reasonable energetics to approximate changes in the C-C bond (Figures S12b and S13b), but C-LA is a slightly better choice for the large C-C-C angle bending (Figures S12c and S13c).
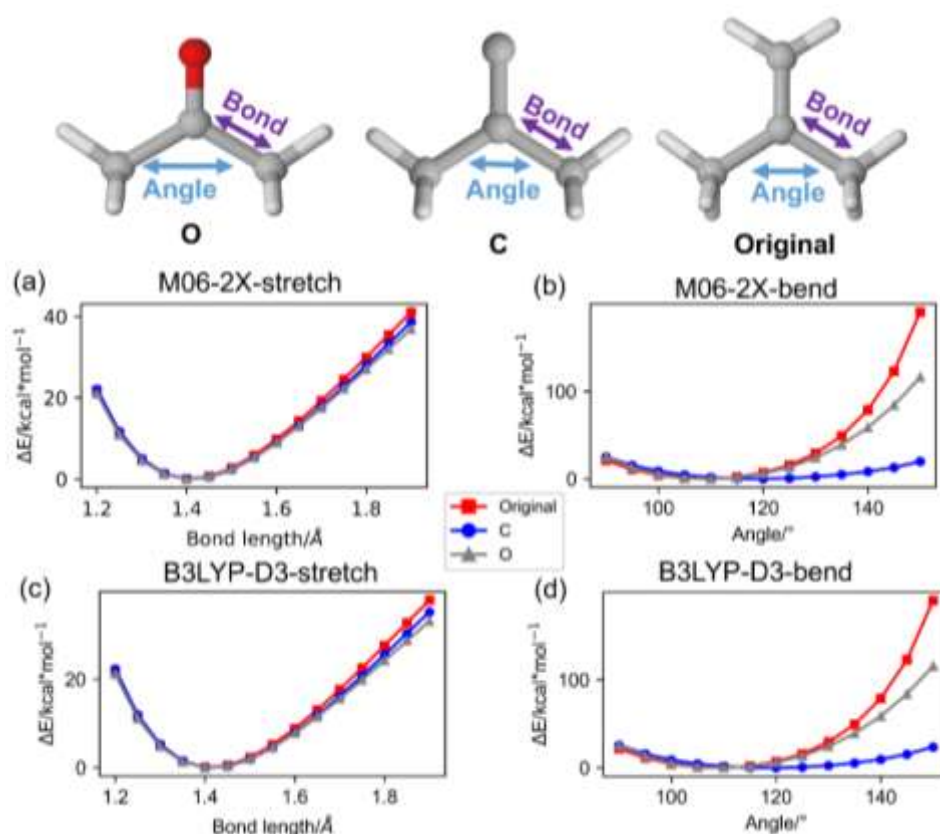
Figure S14 Benchmark of link-atom approximation for C-C double bond. $CH_2=C(CH_3)_2$ as the real system and its two corresponding models ($O=C(CH_3)_2$, $C=C(CH_3)_2$) with either O or C (radical) as the link atom, respectively. Energy profiles for their varying bond length (a, c) and angle (b, d) derived from rigid scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at (a, b) M06-2X/6-31G(d) and (c, d) B3LYP-D3/6-31G(d) levels.
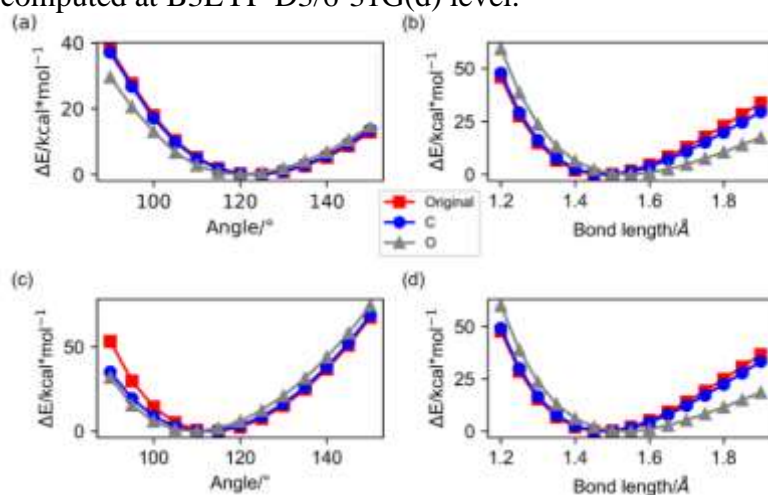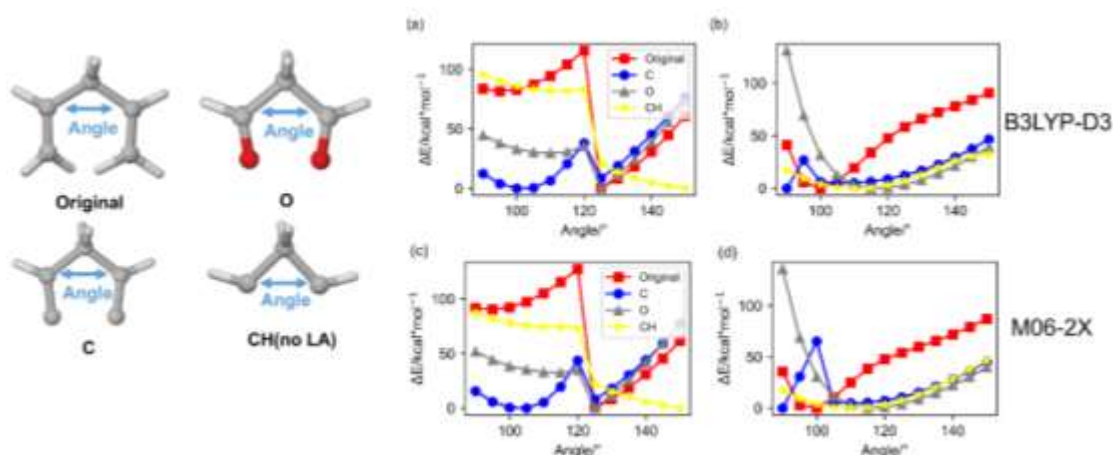
Figure S15 Benchmark of link-atom approximation for C-C double bond of **A1** case in [4+2] Diels-Alder addition reaction (Figure 4). $CH_2=CH-CH_2-CH=CH_2$ as the real system and its four corresponding models ($O=CH-CH_2-CH=O$, $C=CH-CH_2-CH=C$, $CH-CH_2-CH$) with either O or C (radical) as the link atom and no-LA (carbiene), respectively. Energy profiles for their varying angle derived from (a, c) relaxed and (b, d) rigid scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at (a, b) B3LYP-D3/6-31G(d) and (c, d) M06-2X/6-31G(d) levels.

The distortion energy relating to the small C-C-C angle bending (**A1**) of cyclopentadiene in the **S2**-[4+2] Diels-Alder addition reaction (Figure 4) is hard to be evaluated, due to the highly coupling of a few internal coordinates inside the cyclopentadiene ring. Thus, an additional manual step is necessary to mitigate this coupling effect. As shown in Figure S15, the energy curves of C-C-C show a large difference when C-C-C becomes smaller from the equilibrium position for the O-, C-LA and the original cases. These sudden changes may come from the two too-closed link atoms or CH_2.

To minimize interactions between the links atoms, only the CH group without extra link atom was applied, resulting in a distortion energy of 3.1 kcal/mol for **A1**. This value is consistent with the value (3.0 kcal/mol) of fragment **I** (including **A1**) using the most reasonable link atom (CH) in the **M1** scheme (Figure S4).

3.3. Triplet-bond LA and boundaries
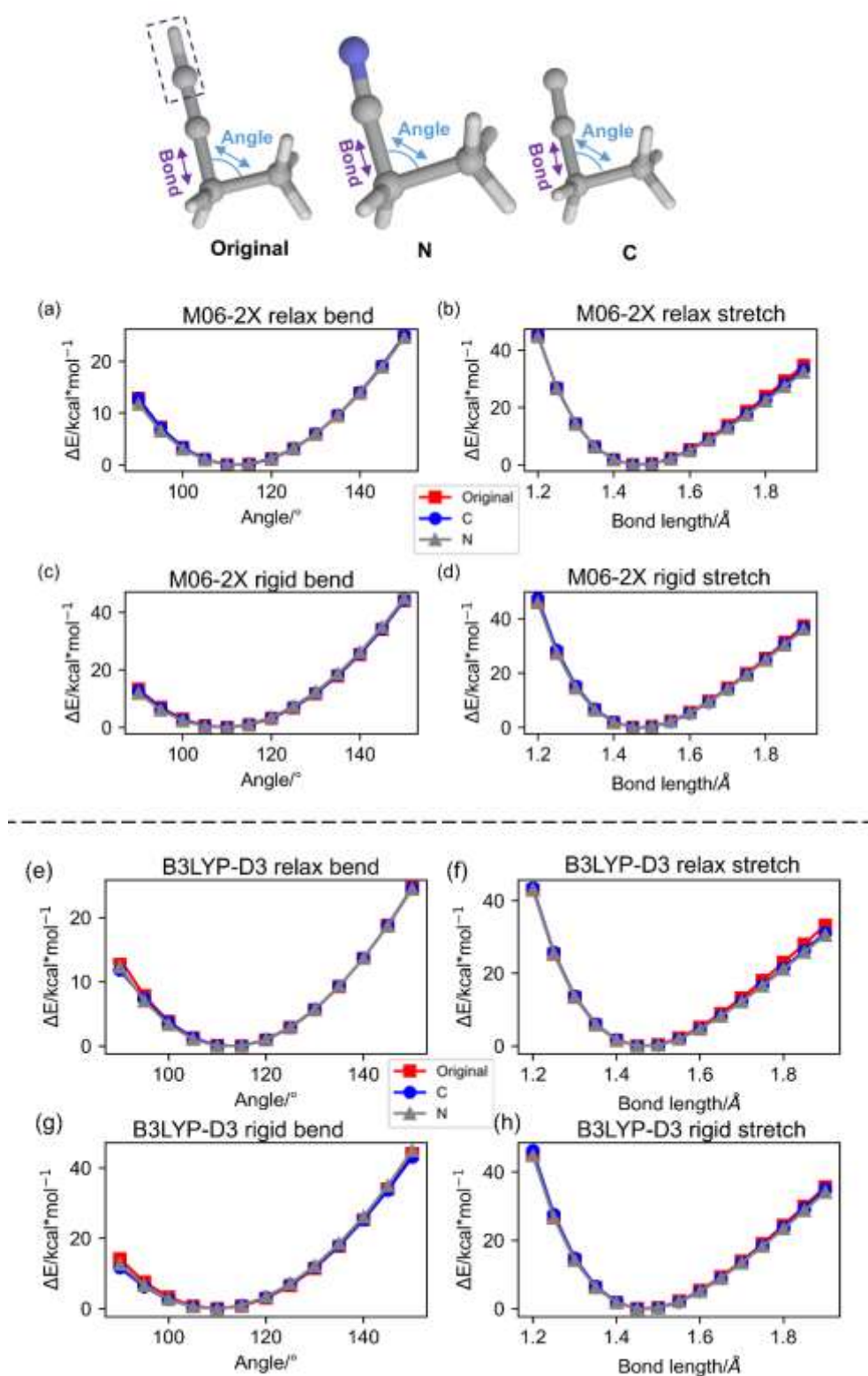
Figure S16 Benchmark of link-atom approximation for C-C triple bonds. 1-butyne as the original system and its two corresponding models (C_3H_5N, C_3H_5C radical) with either N or C (radical) as the link atom, respectively. Energy profiles for their varying (b,d;f,h)bond length and (a,c;e,g)angle derived from rigid scan calculations of the real systems followed by replacement by the link atom and single-point energy calculations computed at (e-h) B3LYP-D3/6-31G(d) and (a-d) M06-2X/6-31G(d) levels.

Similar to results and discussion in Figures 2d-2f, our additional results (using B3LYP-D3/6-31G(d)) further support that N-LA and radical-type C-LA can be a reasonable LA associated with a C-C triple bond boundary of 1-butyne (Figure S16).

4. Supplementary figures for the representative systems in the main text

4.1. Reverse Cope Elimination
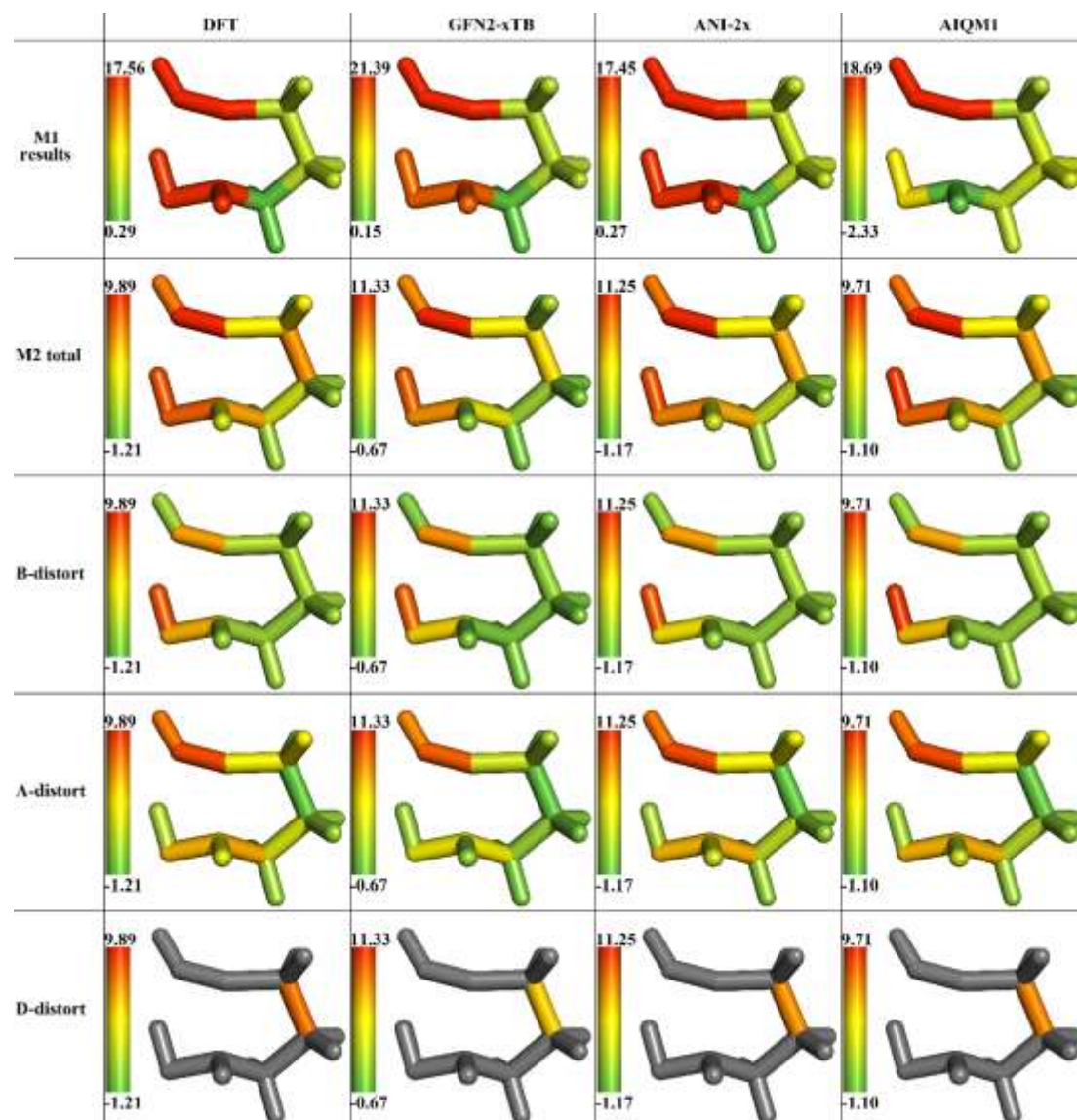
Figure S17 Distortion energy distribution (kcal/mol; coloring in term of log scale) of **S1**[S12] using the **M1** scheme and total distortion and its contribution from bond angle and dihedral terms using the **M2** scheme. The major local distortion contributed from the dihedral (**D1** in Figure 3) was also considered by using an expanded CH_2CH_2 fragment in **M1** scheme and adding **D1** in **M2** scheme). Energies were computed using DFT (B3LYP/6-31G(d)), GFN2-xTB, ANI-2x and AIQM1 methods.
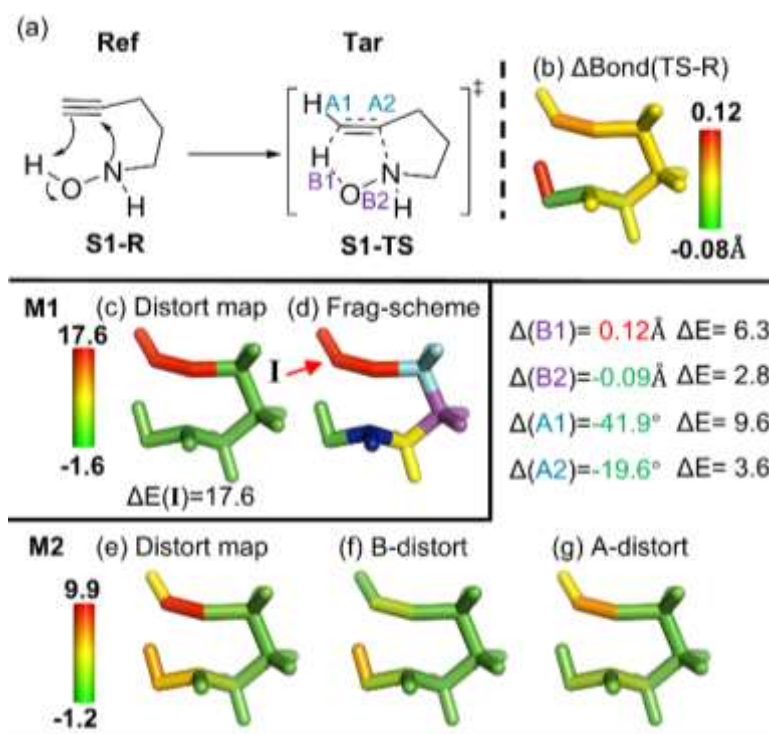
Figure S18 (a) Overview of **S1** (structures taken from ref S12). (b) Bond length change (Δbond in Å) from the reference form (**Ref**; reactant, **S1-R**) to the target form (**Tar**; TS, **S1-TS**). (c) Distortion energy distribution (kcal/mol) and (d) default fragmentation using **M1** scheme using the **default** fragmentation settings. (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M2** scheme using the **default** fragmentation settings. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at B3LYP/6-31G(d) level.

Figure S18 shows the **M1** & **M2** results without including one key dihedrals' contribution (**D1**), i.e. the default fragmentation settings for **M1** and **M2** methods. In addition, as mentioned in the main text, splitting of the N-O bond and introduction of H-LA modestly underestimates the distortion energy, due to the missing O/N lone-pairs repulsion within the smallest fragments (Figure S3).

4.2. Diels-Alder Addition

Figure S19 Distortion energy distribution (kcal/mol in log scale) of **S2**[S13] using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. Energy calculations were computed 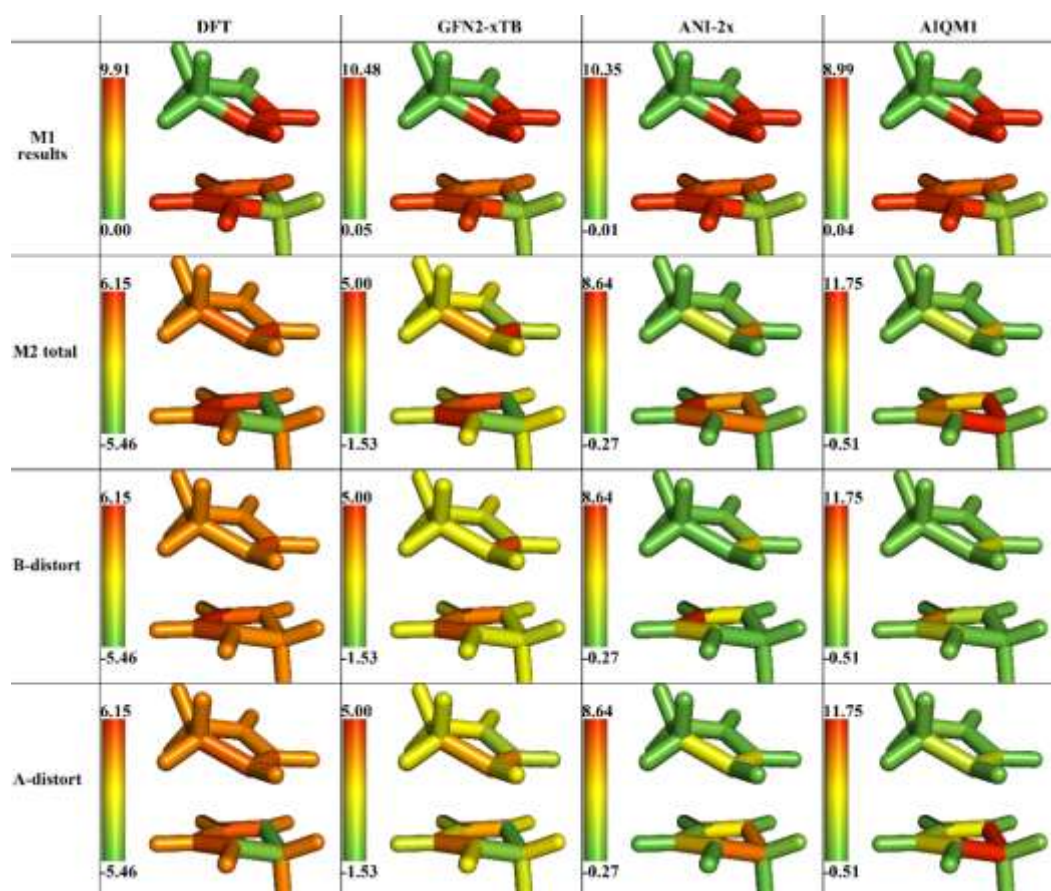using DFT (B3LYP/6-31G(d)), GFN2-xTB, ANI-2x, AIQM1 methods. These results are obtained from default fragmentation + LA settings (i.e. using default LAs) for **M2** & **M1.**

As shown in Figure S19 and discussed in the main text, using default fragmentation settings for **M1** and **M2** schemes can lead to unreasonable results for fragments **I** and **A1** (Figure 4). As shown in the benchmark in Figure S4, H-LA can modestly underestimate the (**A1**) angle bending included in fragment **I**. Using CH as link atoms can provide a reasonable value for fragment **I**. Moreover, unreasonable results were also obtained by the default fragmentation settings for the **M2** scheme, (~-5.5 kcal/mol in A-distort (DFT results), Figure S18). To find a better model to describe such cases for **A1**, benchmarks shown in Figures S4 and S15 were carried out and suggest that a few LAs are incapable to well reproduce the distortion energy. Whereas, no introduction of link atoms (CH model) gives the most consistent results with the CH_3-CH_2-CH_3 (Figure S4). This may indicate that such model (CH) is so far the better model to represent distortion relating to the **A1**'s bending.
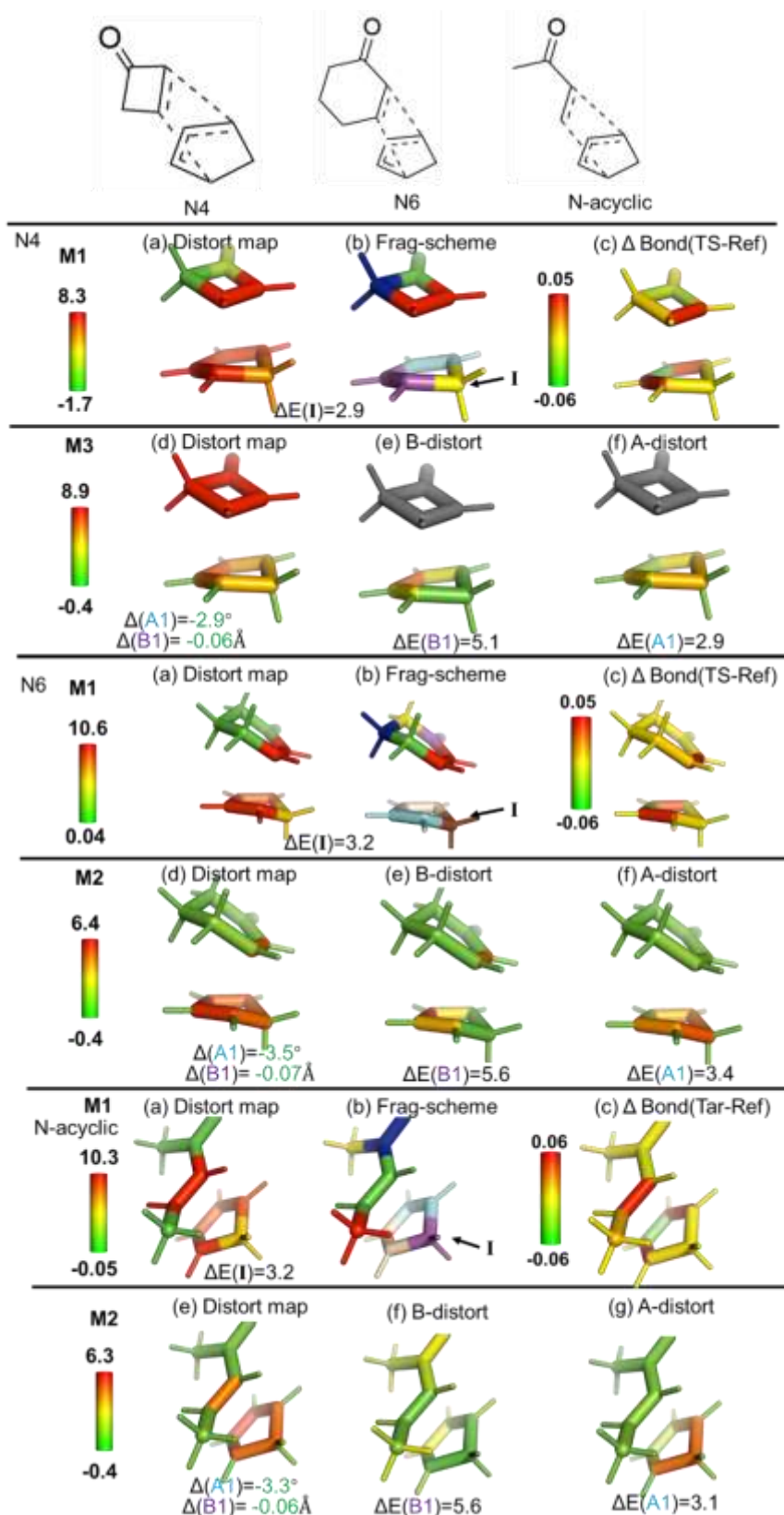
Figure S20 Results for Diels-Alder cycloaddition series (**N4**, **N6**, **N-acyclic** substrates; structures taken from ref S13) Same as the **S2**, the CH was chosen as the link atom for the fragment **I** in **M1** scheme and no link atom was applied to **A1** angle in **M2** scheme (Figure 4). All energies were computed using DFT (B3LYP/6-31G(d)) method.

4.3. Azide-alkyne Cycloaddition
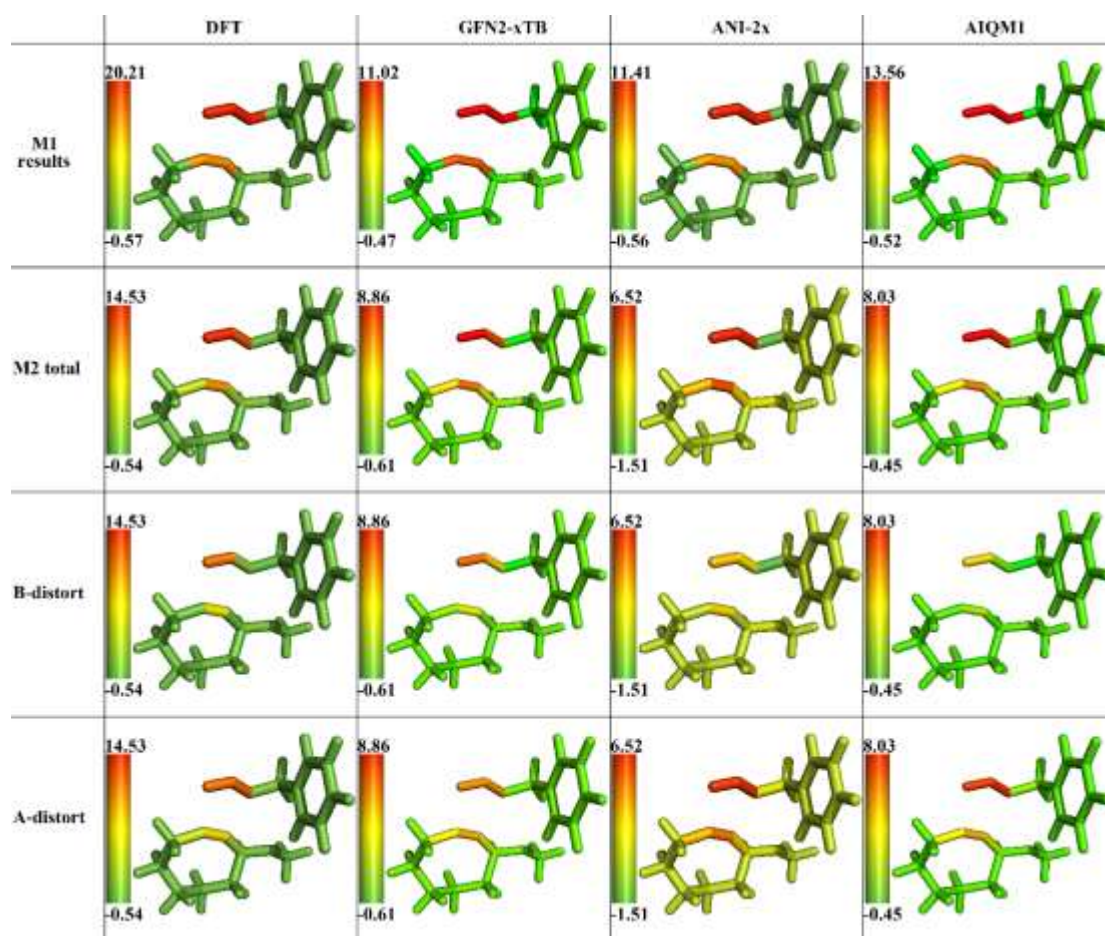
Figure S21 Distortion energy distribution (kcal/mol in log scale) of **S3**[S14] using **M1** and **M2** schemes and its contribution from bond and angle terms. Energy calculations were computed using SCS-MP2/6-31G(d), GFN2-xTB, ANI-2x, AIQM1 methods. SCS-MP2/6-31G(d) calculations were conducted by Gaussian16.

Figure S22 Distortion energy distribution (kcal/mol in log scale) of **S3** using **M3** scheme and its contribution from bond, angle and group fragment terms. Energy calculations were computed using SCS-MP2/6-31G(d), GFN2-xTB, ANI-2x, AIQM1 methods. SCS-MP2/6-31G(d) calculations were conducted by Gaussian16.

To get deeper understanding on this distortion-promoted addition, the same reaction with an acyclic alkyne substrate (2-butyne, **S3*-R**) was also studied by the same computational method (Figure 5h). Based on **S3*-R** structure, the potential energy surface (**S3*-TS** and **S3*-P**) was calculated. To compared with the strained azide-alkyne cycloaddition (**S3**), **S3model-R**, **S3model-TS** and **S3model-P** models were generated by truncating **S3-R**, **S3-TS** and **S3-P** (only keeping the alkyne part) followed by capping with H-LA. Single-point energies on the **S3model-R**, **S3model-TS** and **S3model-P** models were calculated directly to give a rough barrier (~9.6 kcal/mol), which is comparable to the full **S3** system (~7.7 kcal/mol).

4.4. **Triplet Excited-state Di-π-methane Rearrangement.**

Figure S23 Distortion energy distribution (kcal/mol in log scale) of **S4**[S15] using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. Energy calculations were computed using DFT (B3LYP/6-31+G(d)), GFN2-xTB, ANI-2x and AIQM1 methods.
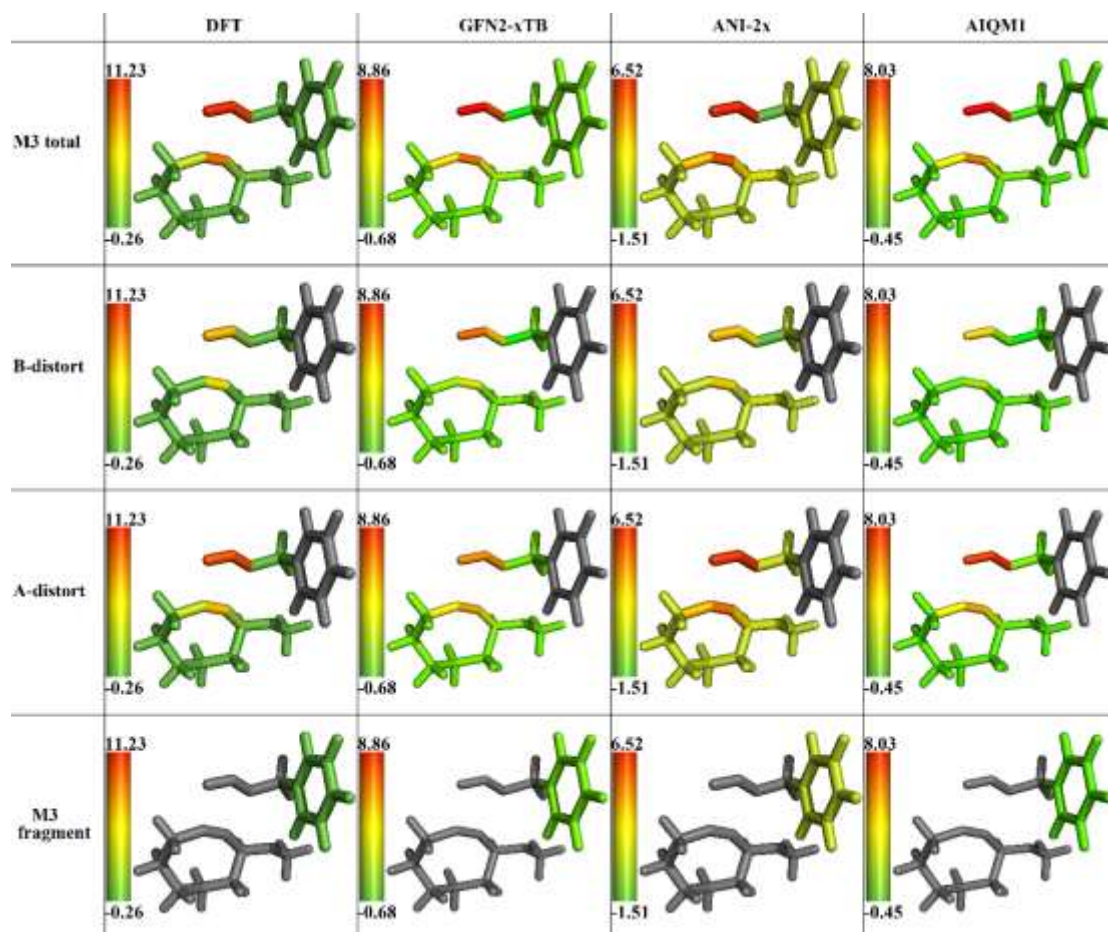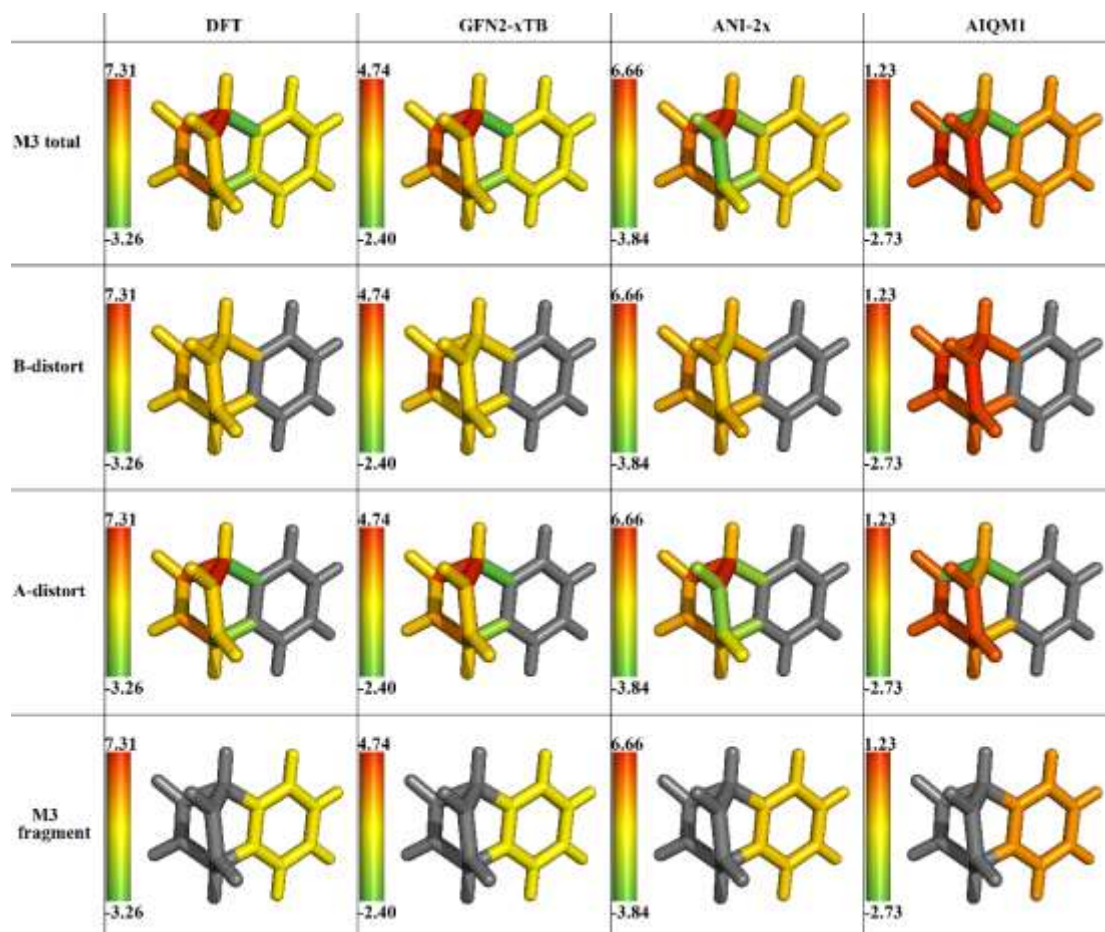
Figure S24 Distortion energy distribution (kcal/mol in log scale) of **S4** using **M3** scheme and its contribution from bond, angle and fragment terms. Energy calculations were computed using DFT (B3LYP/6-31G+(d), GFN2-xTB and ANI-2x and AIQM1 methods

4.5. A Cyclization Reaction of the Fumaramide within [2]Rotaxane.
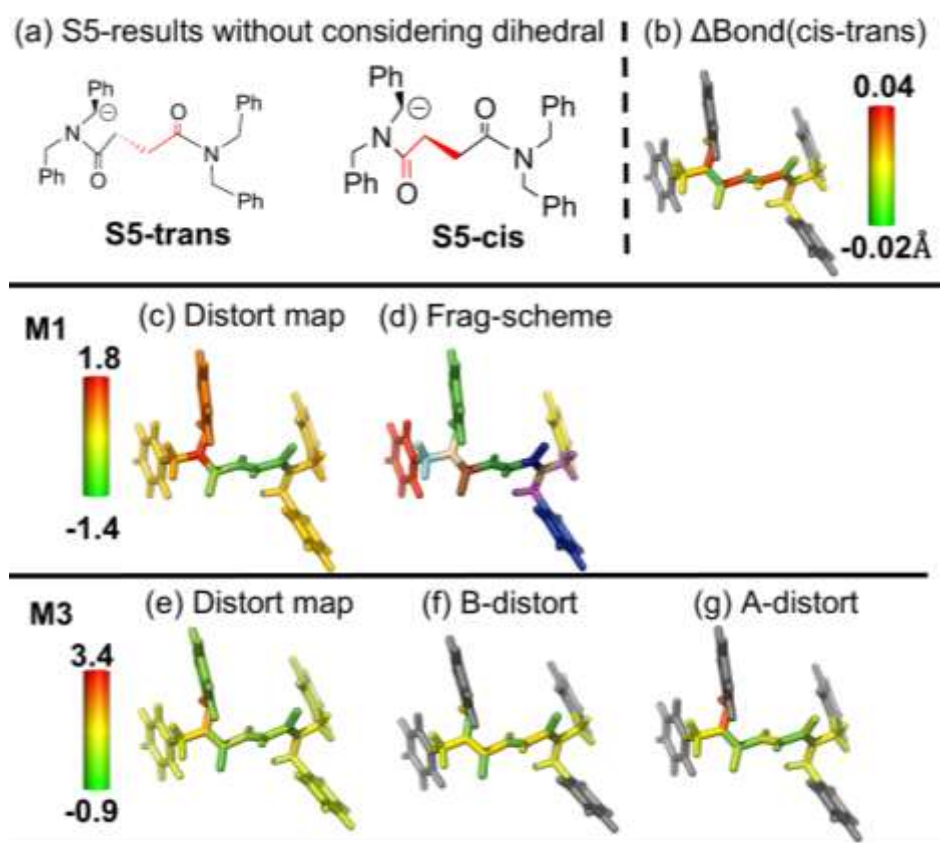
Figure S25 (a) Overview of the cis and trans isomers of an intermediate (**S5**) during the CsOH-promoted cyclization.[S16] (structures taken from ref S16) (b) Bond length change (Δbond in Å) from the reference form (**Ref**; **S5-trans**) to the target form (**Tar**; **S5-cis**). (c) Distortion distribution (kcal/mol) and (d) fragmentation using **M1** scheme. (e) Distortion distribution (kcal/mol), distortion contribution from (f) bond and (g) angle using **M3** scheme. **These results are obtained without consideration of the two key dihedral contribution.** All energies were computed using DFT (SMD M06/6-311++G(d,p)) method.

33

Table S1 Atom indexes and major distortion contributions (>0.5 kcal/mol) from bond and angle using **M3** scheme. These results are obtained without consideration of the two key dihedral contribution. All energies were computed using DFT (SMD M06/6-311++G(d,p)) method.
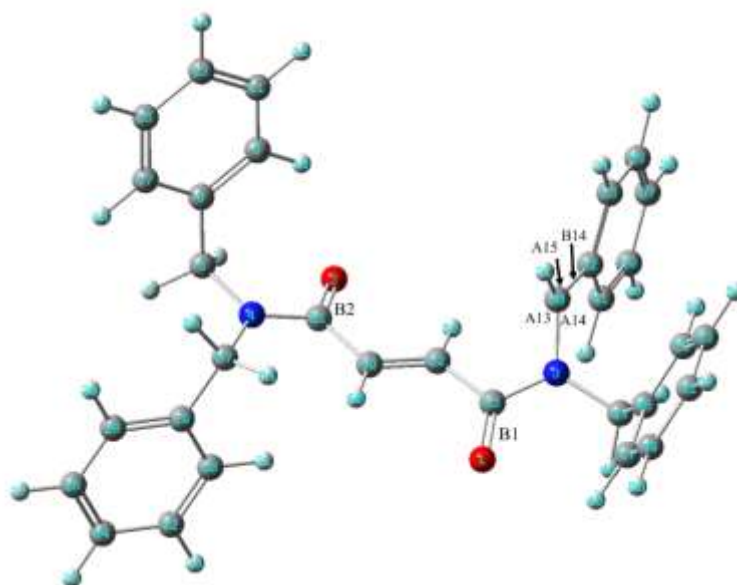
Coordinate ID	Atom indexes	Distortion /kcal*mol^{-1}	Coordinate change Bond: Å; Angle: °
A13	3,30,31	1.49	-5.7
A14	3,30,32	1.16	-4.4
A15	31,30,32	0.99	-7.5
B14	30,32	0.93	0.02
B1	1,65	-0.56	-0.01
B2	2,61	-0.63	-0.01
Sum	**NA**	**1.0^a**	**NA**

a. Summation of local distortion energy of all bonds and angles (including those with <0.5 kcal/mol are not listed above). Distortion of the whole molecule: **5.0 kcal/mol**
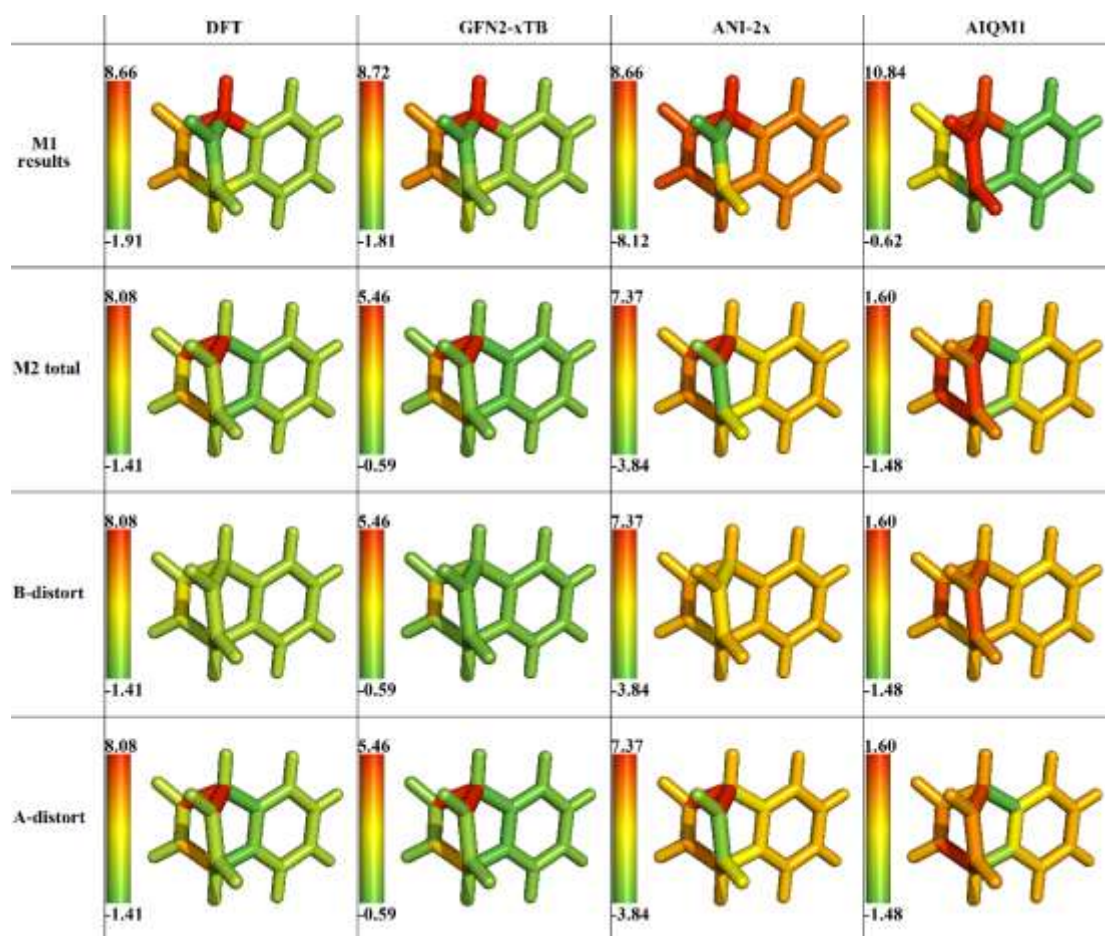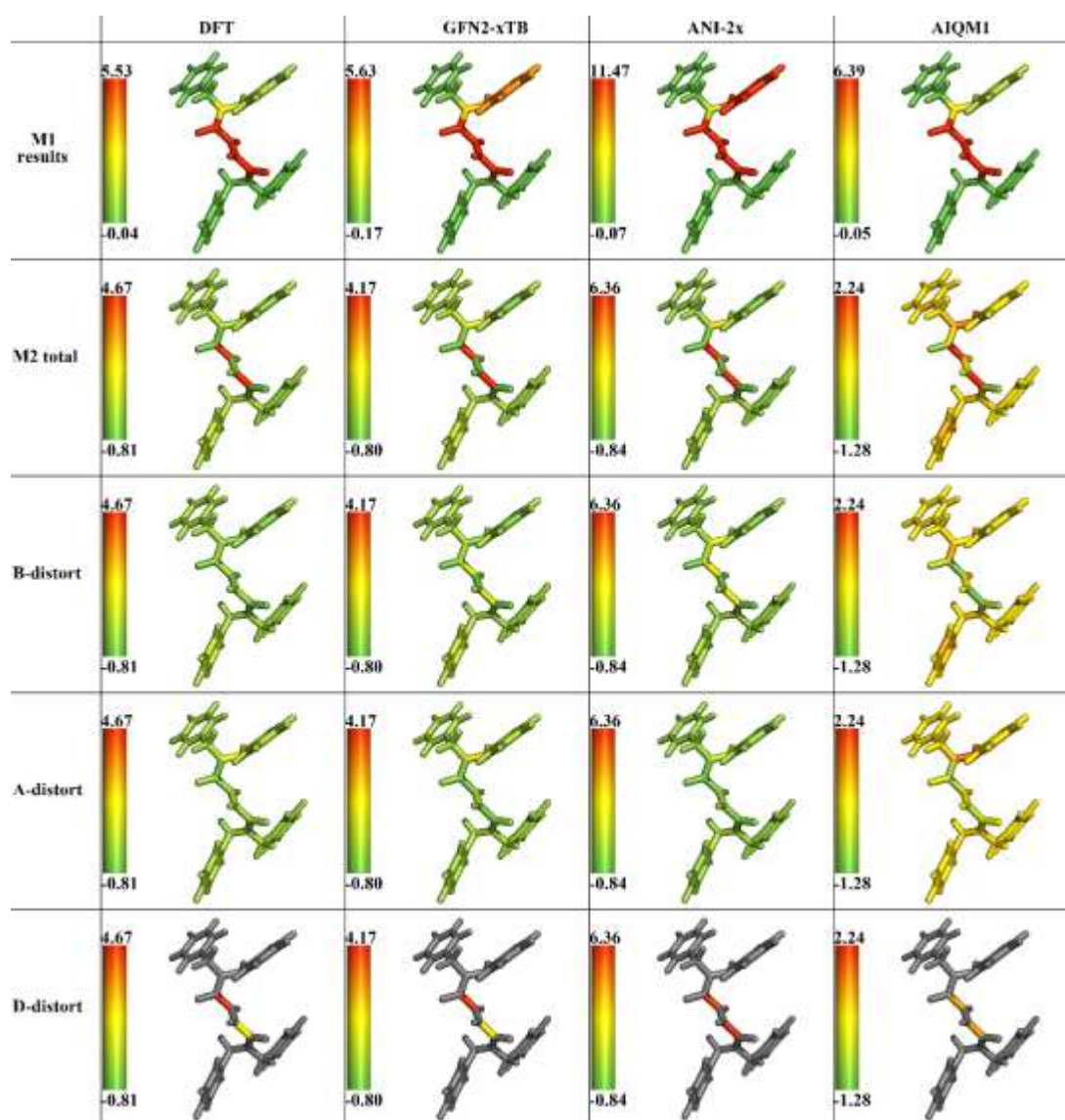
Figure S26 Distortion energy distribution (kcal/mol in log scale) of **S5** using **M1** scheme and total distortion and its contribution from bond, angle and dihedral terms using **M2** scheme. Energy calculations were computed using DFT (SMD M06/6-311++G(d,p)), GFN2-xTB and ANI-2x, AIQM1 methods. The larger fragment **I** (including **D1** and **D2** dihedrals) was further treated using the **M1** scheme and the two key dihedrals were included in the **M2** scheme (Figure 7).
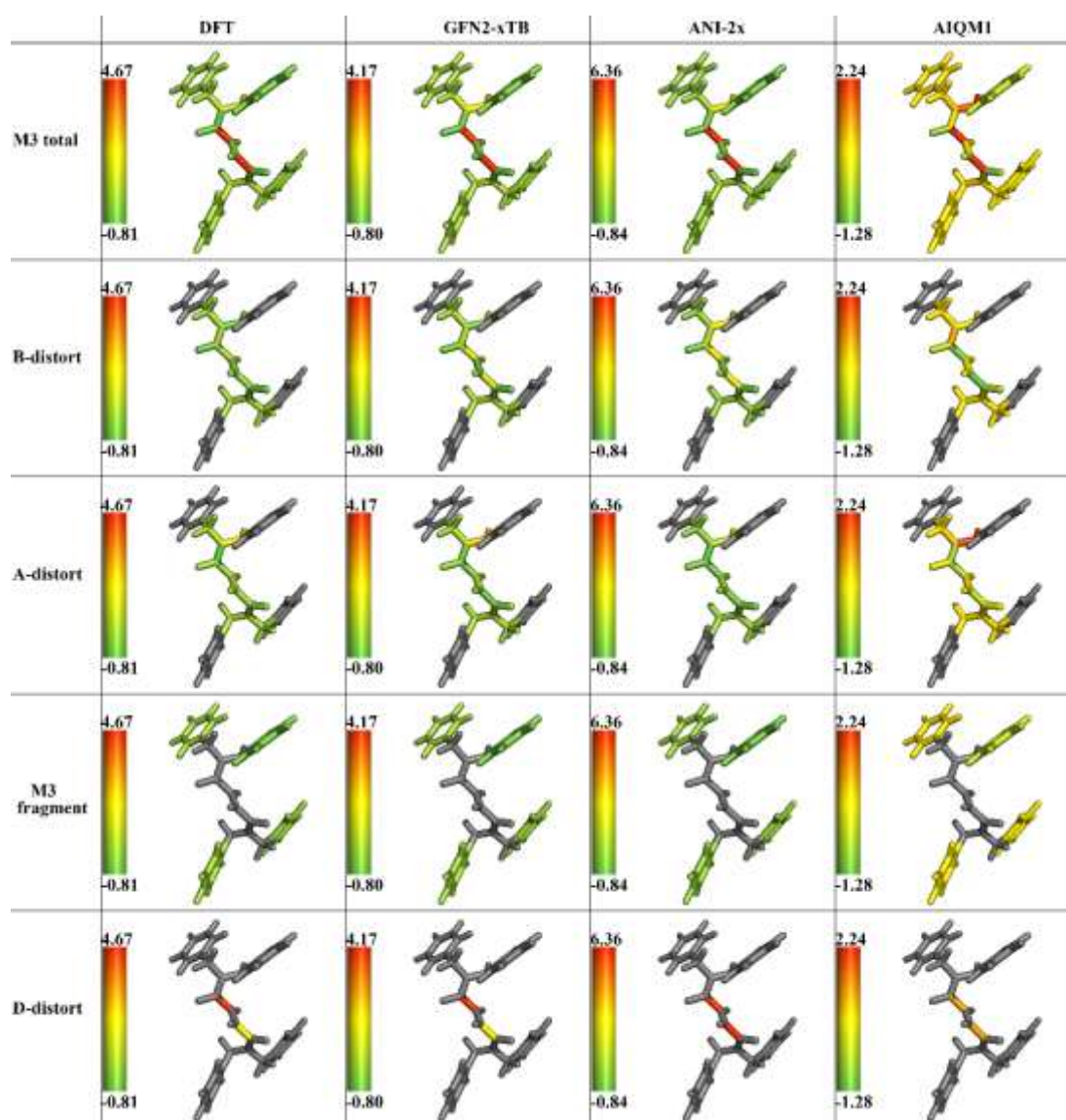
Figure S27 Distortion energy distribution (kcal/mol in log scale) of **S5** using **M3** scheme and its contribution from bond, angle and dihedral and group fragment terms using DFT (SMD M06/6-311++G(d,p)), GFN2-xTB and ANI-2x, AIQM1 methods. The two key dihedrals were included in the **M3** scheme (Figure 7).

4.6. Iridium-Catalyzed C−H Borylation
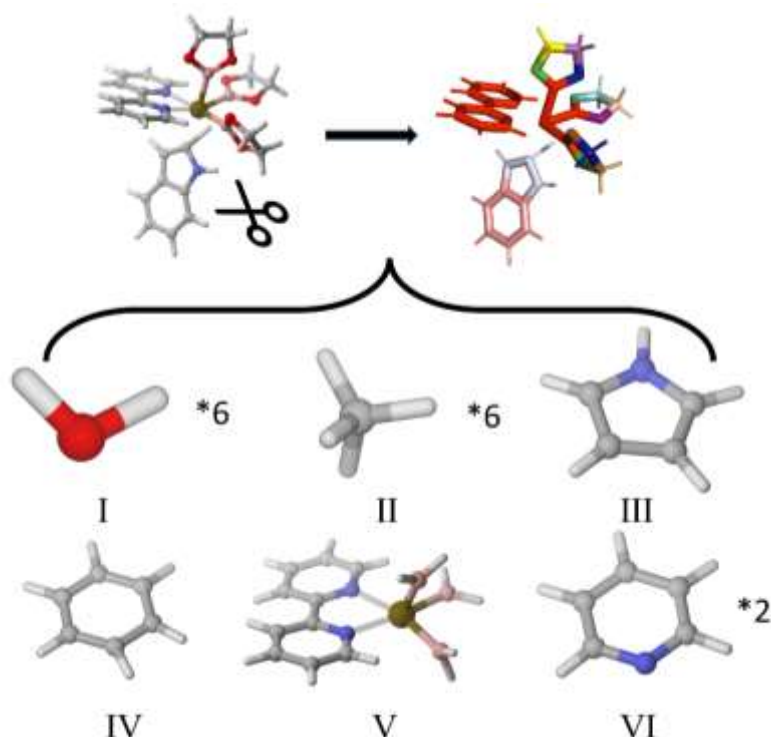
Figure S28 All molecular fragments generated by the **M1** scheme from the Ir-indole system (**S6**)

As shown in Figure S28, the whole system of the Ir-indole system is divided into six kinds of fragments using the **M1** scheme. Six fragments **I** come from six oxygen atoms in the Beg rings. Six fragments **II** come from the sp3 carbon atoms. Fragments **III** and **IV** come from separation of Indole ring in order to localize the distortion. The specialized fragment **V**, which includes the Ir center, along with its coordination boron atoms and the bpy are set to capture the crucial coordination interactions. Two ring fragments **VI** were obtained by separating the bpy to calculate the ring distortion.

4.7. Si-H Insertion enabled by Artificial Heme Enzyme
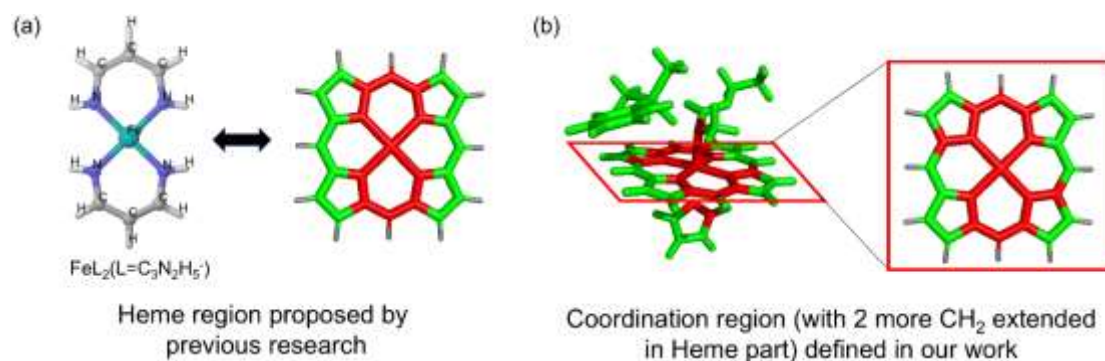
Figure S29 (a) Heme region proposed by previous research.[S17] (b) The specialized coordination region defined in our work (with two more CH_2 extension from the previous model (a) in the heme part as well as the truncated ligated parts) for the heme-catalyzed Si-H insertion. The remaining heme parts (green) are treated by the **M1** scheme with each fragment consisting of two sp2 carbon atoms and their hydrogen atoms.

As proposed previously,[S17] Figure S29a shows the simplified coordination region/fragment as the core heme part. In our work, the specialized coordination region includes the heme region (the red part in Figure S29b), the ligated C atom of the ester and the truncation part ($NHCH_2$) of the ligated imidazole to study the heme protein catalyzed Si-H insertion. This region keeps coordinating interactions around central Fe atom.
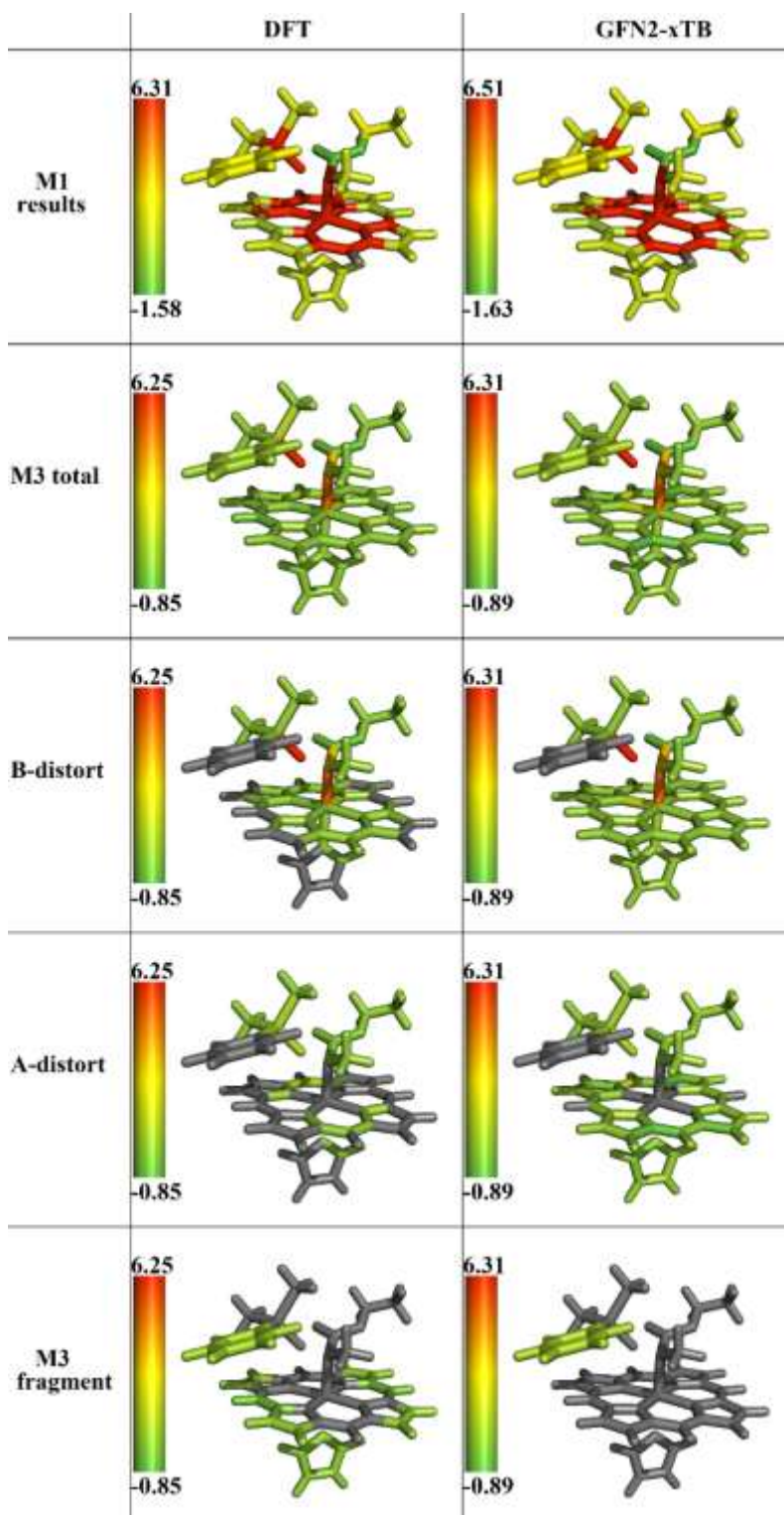
Figure S30 Distortion energy distribution (kcal/mol in log scale) of the Si-H insertion enabled by artificial Heme enzyme (**S7**)[S18] using **M1** scheme and total distortion and its contribution from bond, angle and group fragment terms using **M3** scheme. Energy calculations were computed using DFT (CPCM B3LYP/def2-TZVP) and GFN2-xTB methods

4.8. Imatinib in Spleen Tyrosine Kinase

The drug Imatinib in spleen tyrosine kinase from the X-ray and QR refined structures were studied, in which the latter one were obtained using ONIOM(ωB97-D3/6-31G(d):GFN2-xTB:Amber) method.[S19] As shown in Figure S31b, e, the **M1** and **M3** results show some difference in distortion energy related to the amide group, possibly due to the mis-counting contribution from **D1** (Δ(**D1**) = ~36.6°). After taking **D1** into account, distortion energy in this region by **M3** scheme becomes more consistent with that by **M1** scheme.
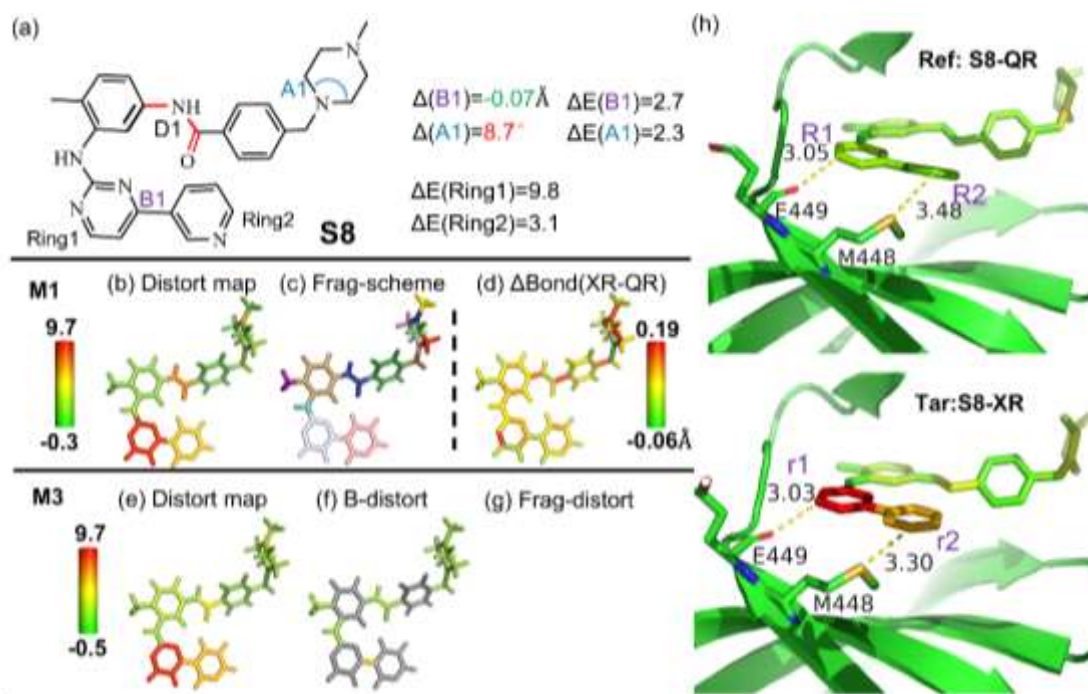
Figure S31 (a) Overview of structures of Imatinib-Spleen tyrosine kinase(**S8**)[S19] (structures from ref S19). (b) Distortion energy distribution (kcal/mol) and (c) fragmentation using **M1** scheme. (d) Bond length change (Δbond in Å) from quantum refinement (**Ref**, QR, **S8-QR**) structure to X-ray (**Tar,** XR, **S8-XR**) structure. (e) Distortion energy distribution (kcal/mol), distortion contribution from (f) bond and (g) angle using **M3** scheme. (h) Scheme for binding site interaction with neighbouring residues taken from the XR and QR structures. The drug structures are colored based on their relative distortion to that optimized structure in gas phase using **M1** scheme. These **M3** results did not include contribution from the key **D1** dihedral. All energies were computed using DFT (ωB97X-D/6-31G(d)) method.

Figure S32 Distortion energy distribution (kcal/mol in log scale) of Imatinib-Spleen tyrosine kinase. The X-ray structure and the optimized structure in the gas phase are set as the target and reference forms. Total distortion distribution using **M1** scheme, total contribution and its contribution from bond, angle and dihedral terms using **M2** scheme. Energy calculations were computed using DFT (ωB97X-D/6-31G(d)), GFN2-xTB, ANI-2x and AIQM1 methods.
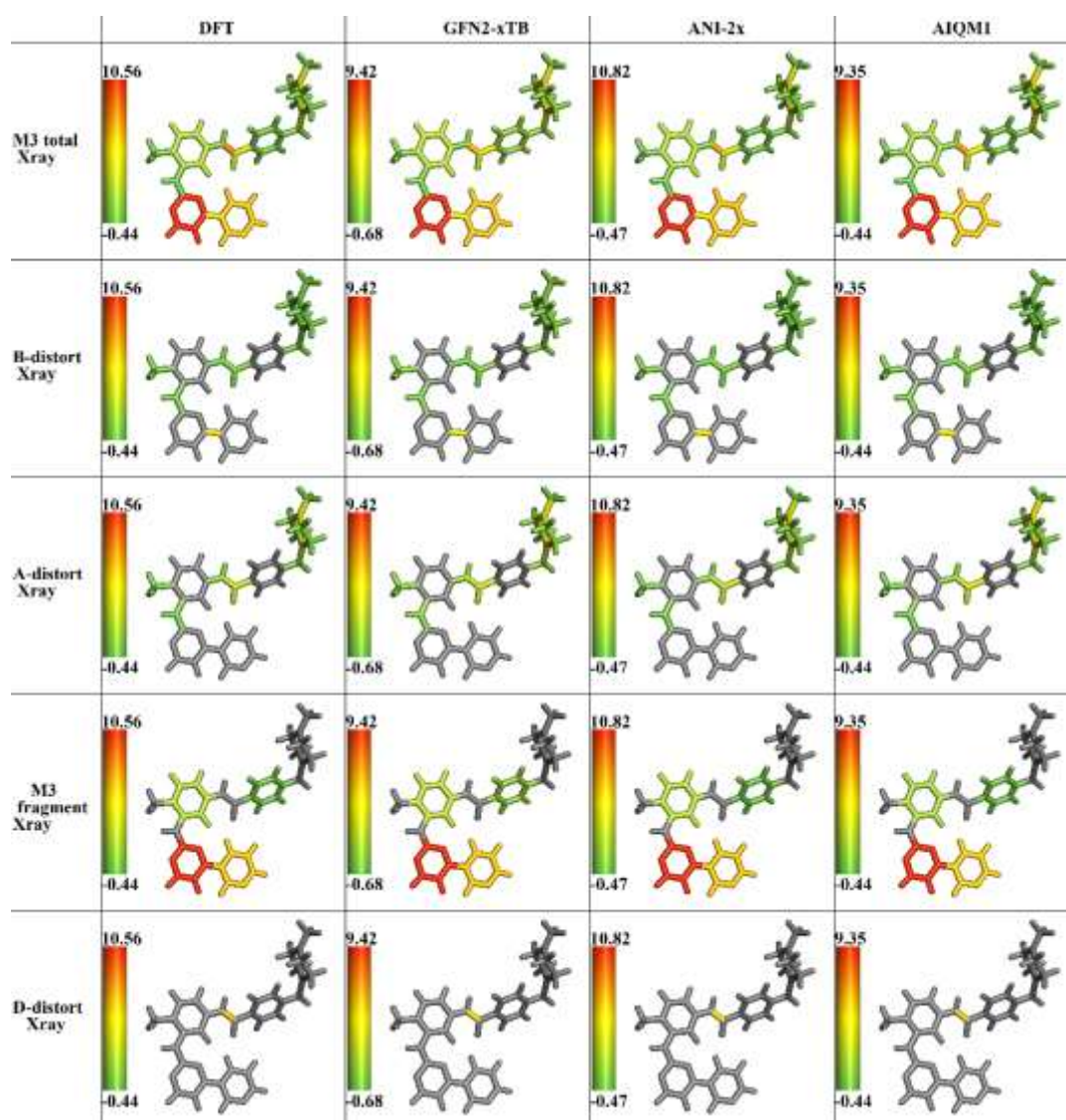
Figure S33 Distortion energy distribution (kcal/mol in log scale) of Imatinib-Spleen tyrosine kinase. The X-ray structure and the optimized structure in the gas phase are set as the target and reference forms. Total contribution and its contribution from bond, angle and dihedral terms using **M3** scheme. Energy calculations were computed using DFT (ωB97X-D/6-31G(d)), GFN2-xTB, ANI-2x and AIQM1 methods.
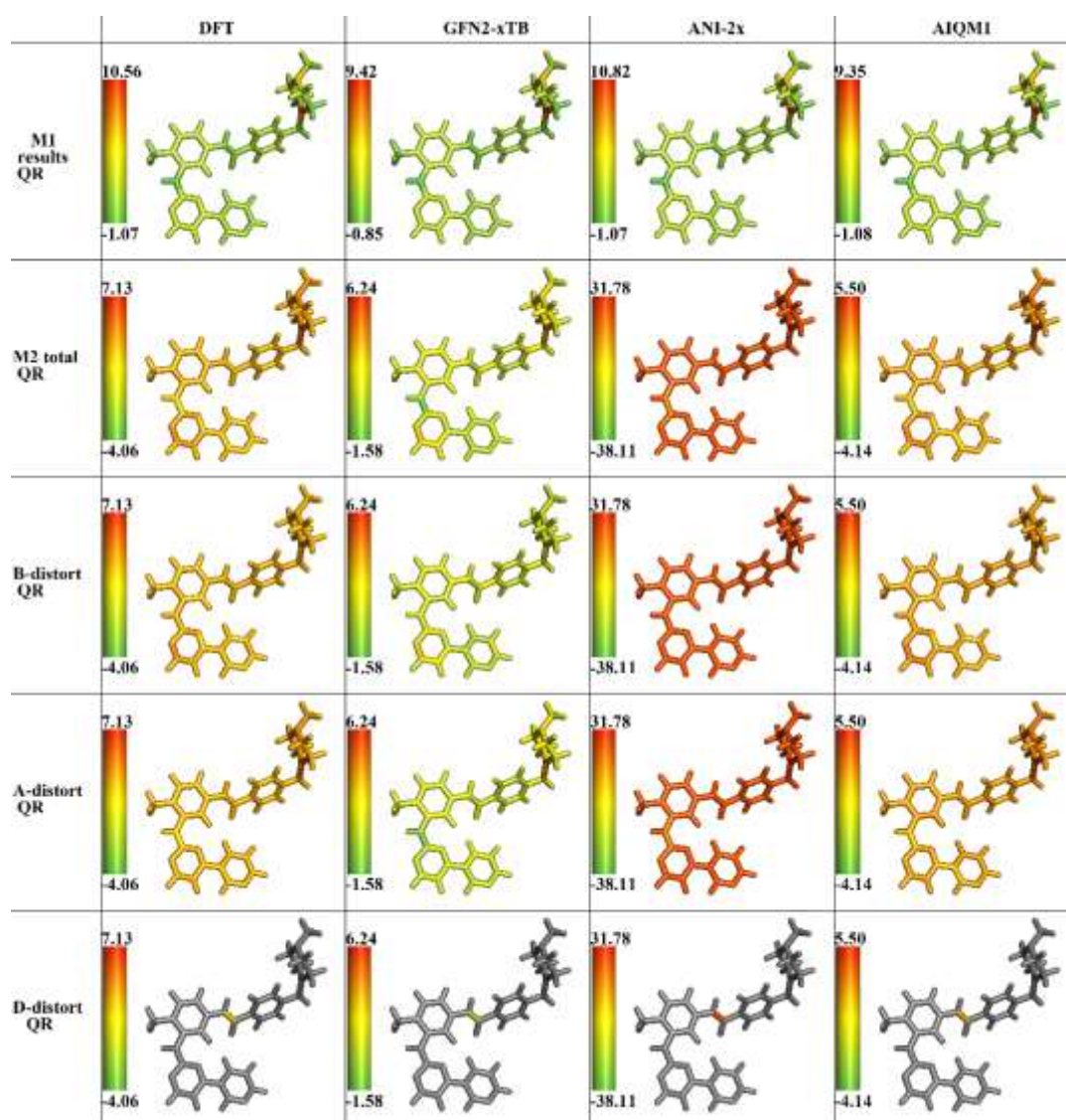
Figure S34 Distortion energy distribution (kcal/mol in log scale) of Imatinib-Spleen tyrosine kinase. The QR refined structure and the optimized structure in the gas phase are set as the target and reference forms. Total distortion distribution using **M1** scheme, total contribution and its contribution from bond, angle and dihedral terms using **M2** scheme. Energy calculations were computed using DFT (ωB97X-D/6-31G(d)), GFN2-xTB, ANI-2x and AIQM1 methods.

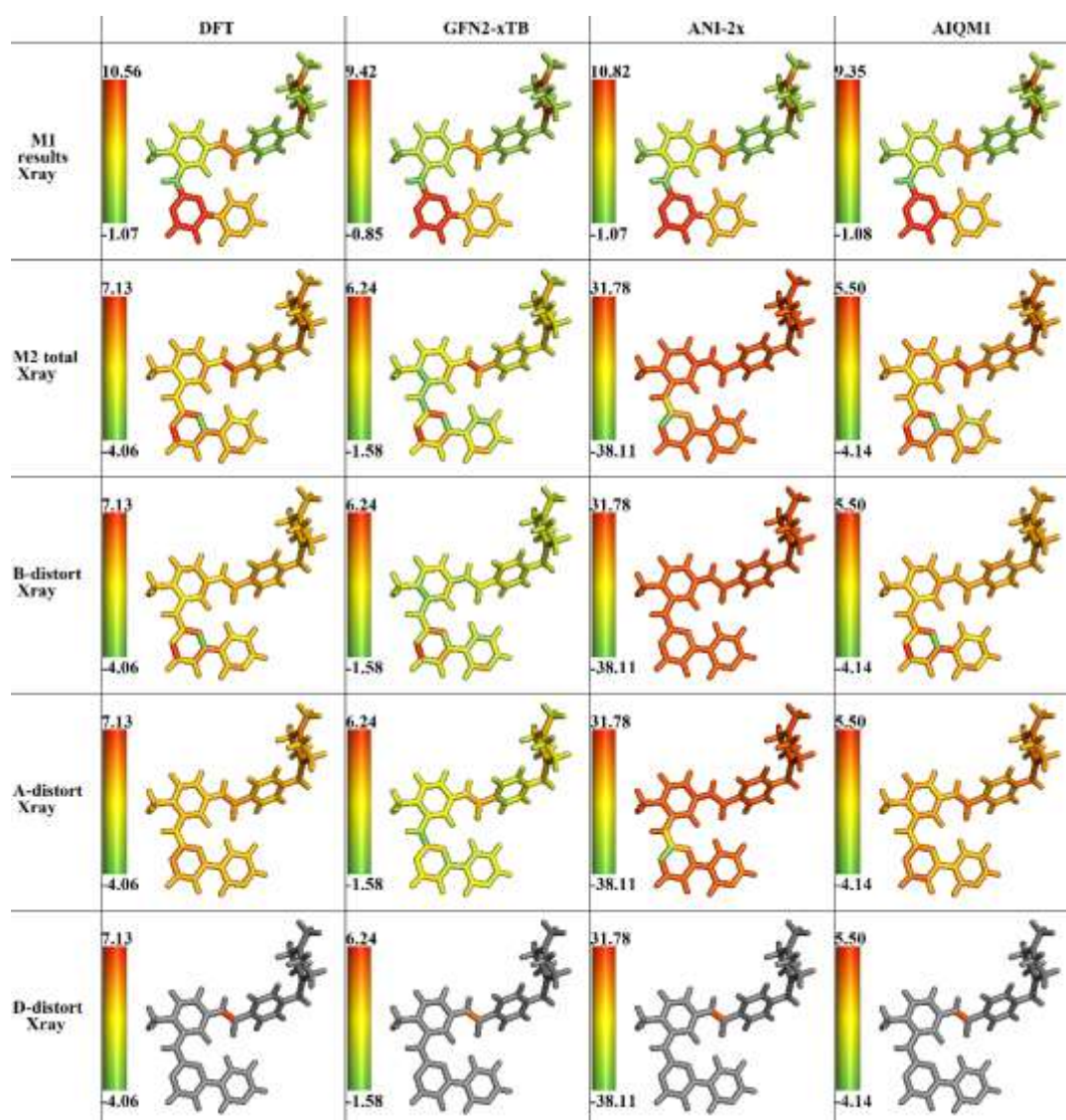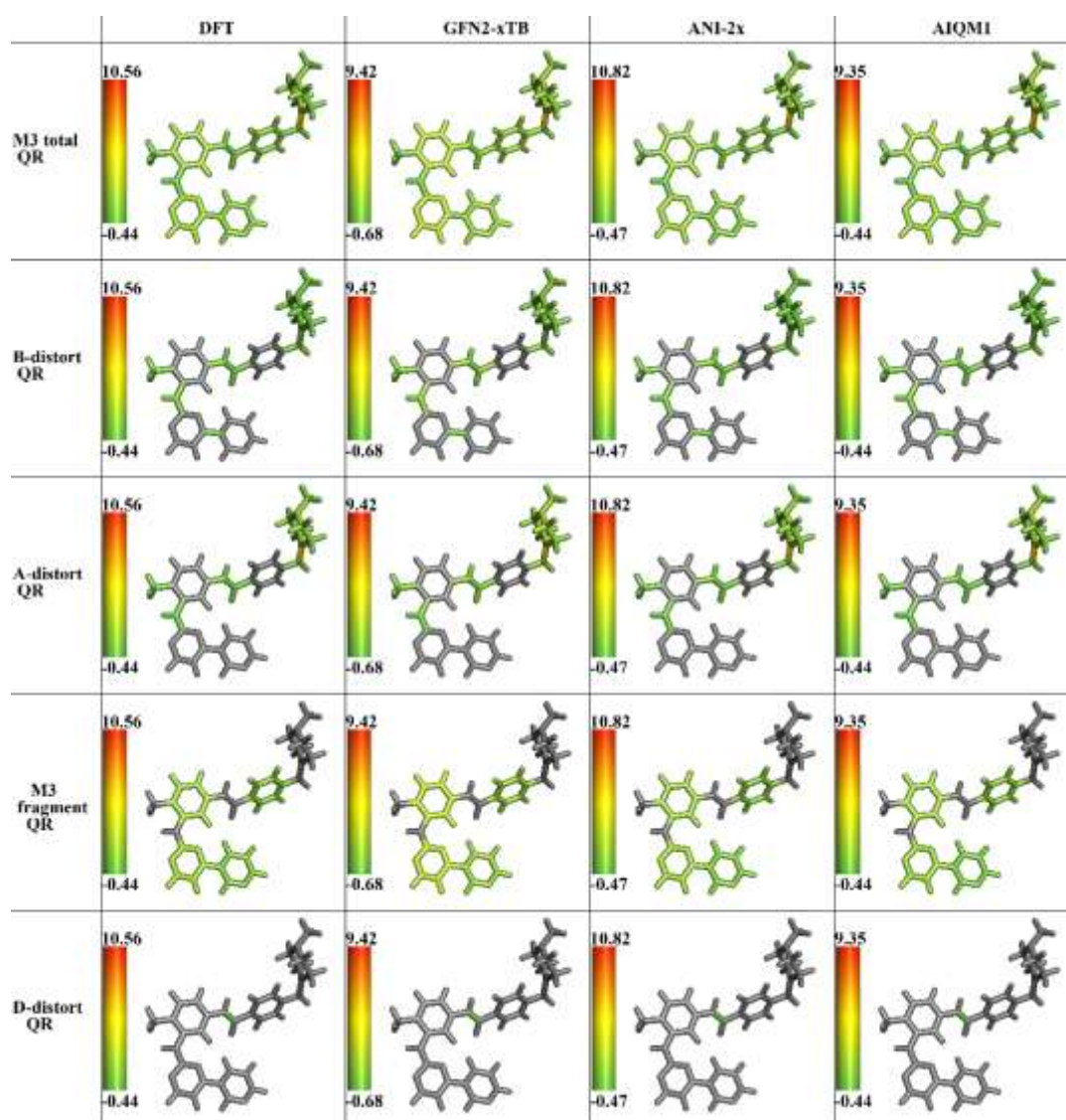**Figure S35** Distortion energy distribution (kcal/mol in log scale) of Imatinib-Spleen tyrosine kinase. The QR refined structure and the optimized structure in the gas phase are set as the target and reference forms. Total contribution and its contribution from bond, angle and dihedral terms using **M3** scheme. Energy calculations were computed using DFT (ωB97X-D/6-31G(d)), GFN2-xTB, ANI-2x and AIQM1 methods.

bond	X-ray	QR
b1	1.37	1.34
b2	1.37	1.35
b3	1.41	1.48
b4	1.37	1.42
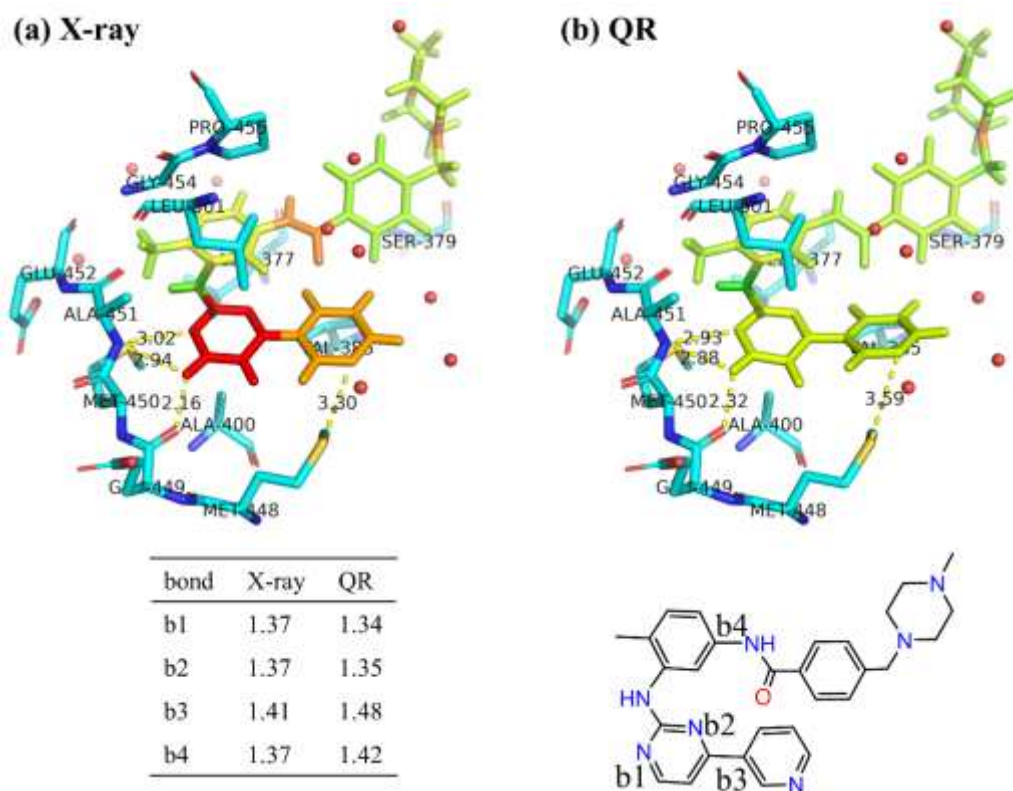
Figure S36 The key distances between Imatinib and neighboring residues in Imatinib-Spleen tyrosine kinase from the X-ray and QR refined structures. The key bond lengths (Å) are listed. Imatinib is colored based on distortion distribution using **M1** scheme while the optimized structure in the gas phase was set as the reference form.

4.9. S_N2 (MD)

$$Cl^- + CH_3Cl \longrightarrow \left[\begin{array}{c} H \quad H \\ Cl\text{---}C\text{--}Cl \\ | \\ H \end{array} \right]^{\ddagger} \longrightarrow ClCH_3 + Cl^-$$

R   TS   P
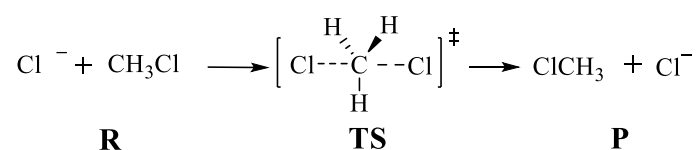
Figure S37 SN2 reaction (**S9**)

Computational details: the reactant, transition state and product for this reaction were first optimized by HF/3-21G. Then, on-the-fly quasi-classical MD simulations were initiated by normal mode sampling at 298.15 K from this optimized transition state at the HF method using Progdyn code.[S20] The trajectory was propagated in both the forward and backward directions for 150 MD steps with a time step of 1 fs. The energy and force of all structures at each step were calculated on-the-fly by the HF method.
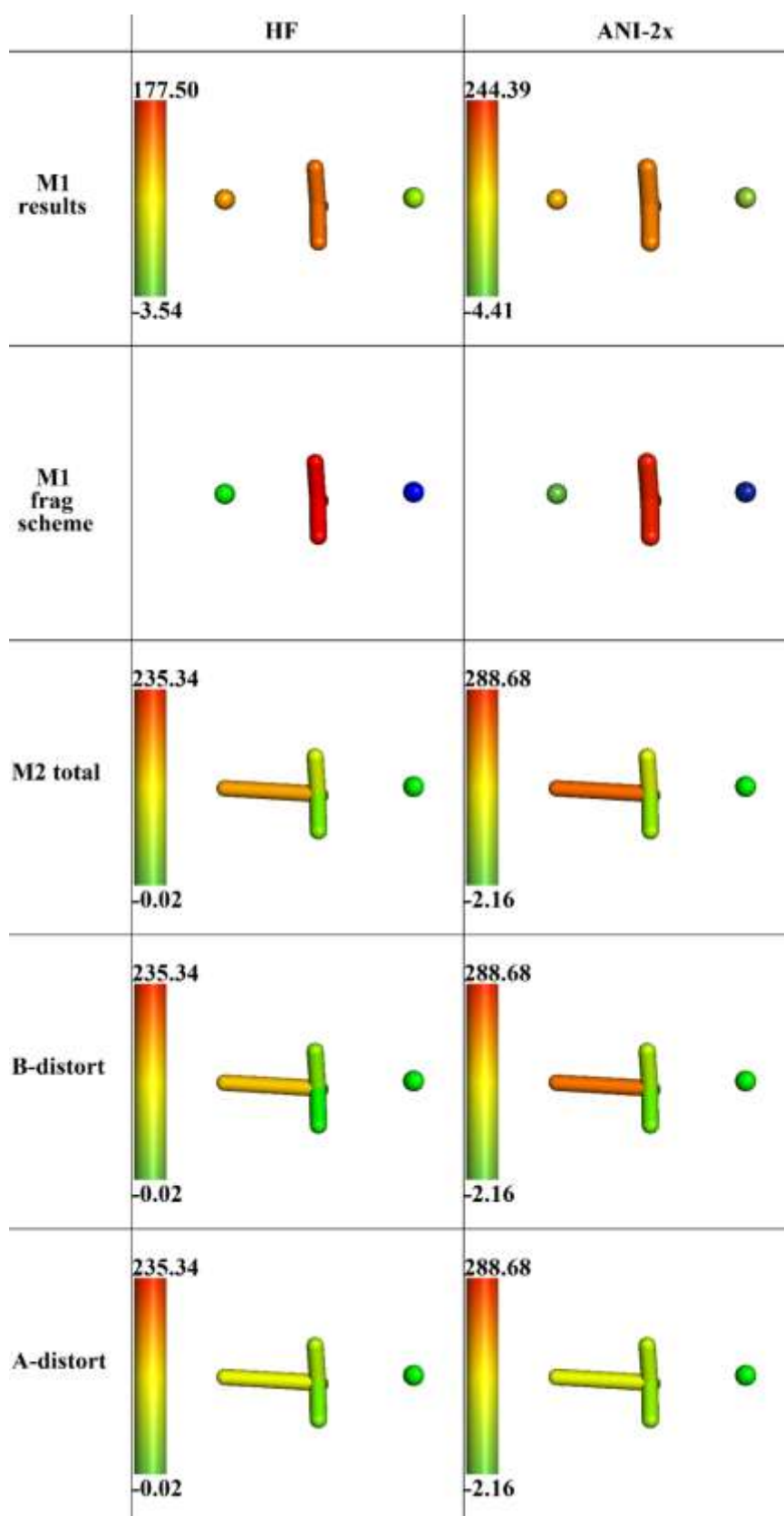
Figure S38 Distortion energy distribution (kcal/mol in log scale) for the transition state (TS) for a S_N2 reaction between CH_3Cl and a Cl^- using **M1** scheme, total distortion and its bond and angle contributions using **M2** scheme. All energies were computed using HF/3-21G and ANI-2x methods.
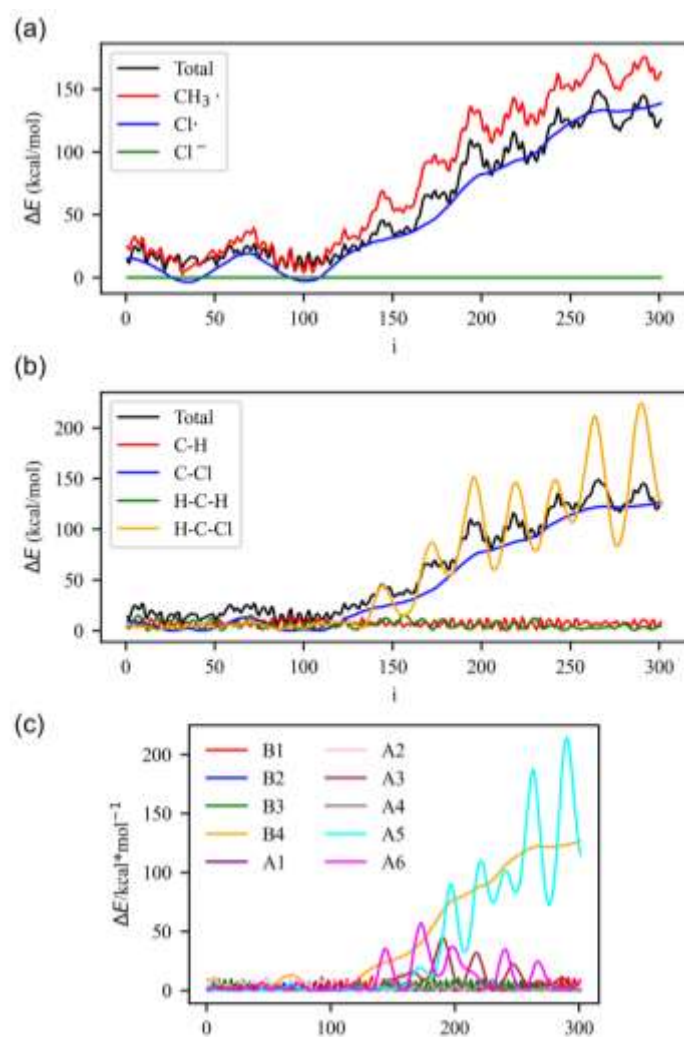
Figure S39 Distortion energy distribution (kcal/mol in log scale) for a MD trajectory of the S_N2 reaction between CH_3Cl and a Cl^-. (a) Distortion distribution using **M1** scheme. (b) Distortion distribution using **M2** scheme (with summing distortion energy belonging to the same kind of chemical bonds/angles) (c) Distortion distribution using **M2** scheme (B1, B2, B3: C-H; B4: C-Cl; A1, A2, A4: H-C-H; A3, A5, A6: H-C-Cl).

5. Supplementary application results

The additional testing systems can be categorized into different types: inorganic systems, organic systems, excited-state organic systems, supramolecular systems, coordination systems, biochemical systems and multiple structures.

5.1. Inorganic systems

5.1.1. Frustrated Lewis-Pair Catalysis

This inorganic system is a geminal aminoborane-based frustrated Lewis pair,[S21] which activates H_2. All structures in Figure S40 were taken from ref. S21 for our D2AF analysis using DFT (M06-2X/def2-TZVPP) and GFN2-xTB methods.
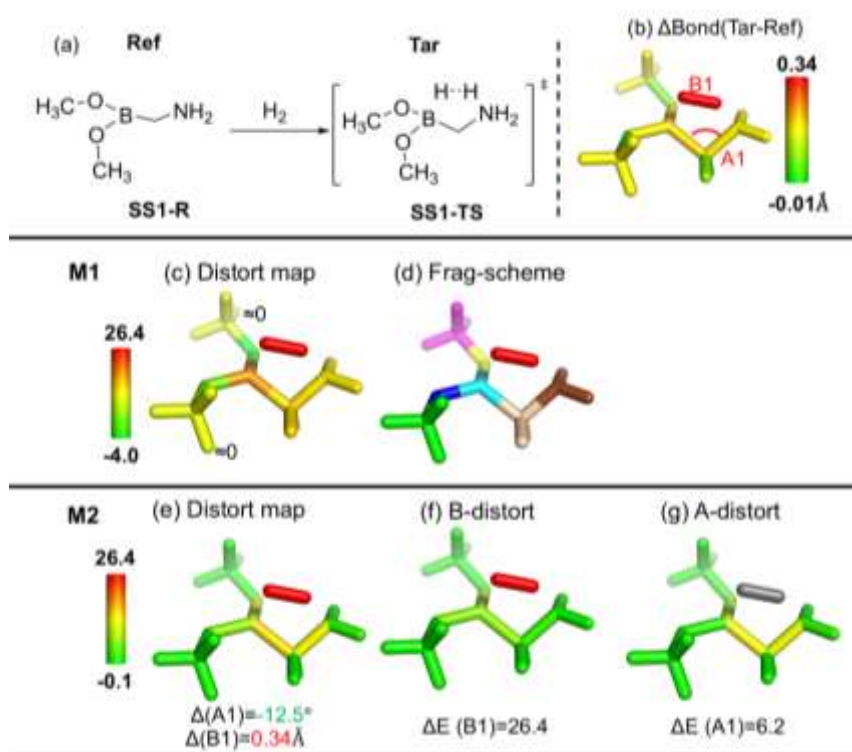
Figure S40 (a) Overview of the system of hydrogen activation with a frustrated Lewis pair. (b) Bond length change (Δbond in Å) from the reference form (**Ref**; R, **SS1-R**) to the target form (**Tar**; TS, **SS1-TS**). (c) Distortion energy distribution (kcal/mol) and (d) fragmentation using **M1** scheme. (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M2** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at M06-2X/def2-TZVPP level.

	DFT	GFN2-xTB

Figure S41 Distortion energy distribution (kcal/mol in log scale) for hydrogen activation by the frustrated Lewis pair using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. All energies were computed using DFT (M06-2X/def2-TZVPP) and GFN2-xTB methods.

5.1.2. Triphosphate Hydrolysis via S_N2

This inorganic system is a Mg-catalyzed triphosphate hydrolysis, with MeO[-] as the nucleophile and pyrophosphate as the leaving group.[S22] All structures in Figure S42 were taken from ref. S22 for our D2AF analysis using OLYP/def2-TZVP and GFN2-xTB methods.
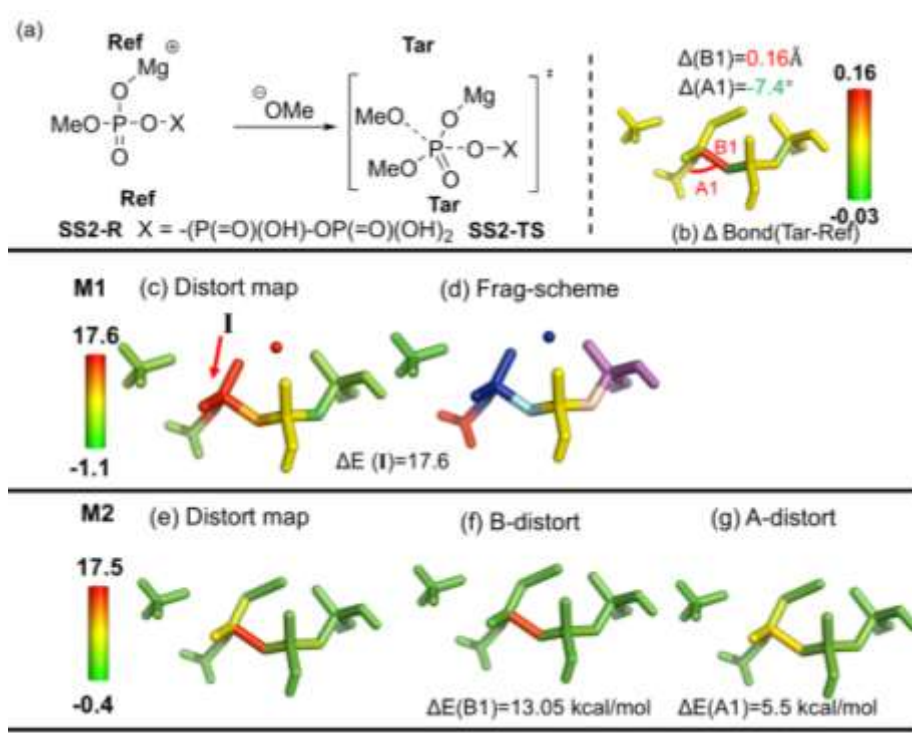
Figure S42 (a) Overview of the Mg-catalyzed triphosphate hydrolysis. (b) Bond length change (Δbond in Å) from the reference form (**Ref**; R, **SS2-R**) to the target form (**Tar**; TS, **SS2-TS**). (c) Distortion energy distribution (kcal/mol) and (d) fragmentation using **M1** scheme. (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M2** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at OLYP/def2-TZVP level

Figure S43 Distortion energy distribution (kcal/mol in log scale) of Mg-catalyzed triphosphate hydrolysis using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. All energies were computed using DFT (OLYP/def2-TZVP) and GFN2-xTB methods.

5.2. Organic systems
5.2.1. HCl-Mediated Prins Reaction (Organocatalysis)

In this case, a transition state (TS) structure as **Tar** and its preceding complex as **Ref** in a Prins reaction were taken from ref. S23 for our D2AF analysis using DFT (CPCM, B3LYP/6-31G(d)), GFN2-xTB and ANI-2x methods.
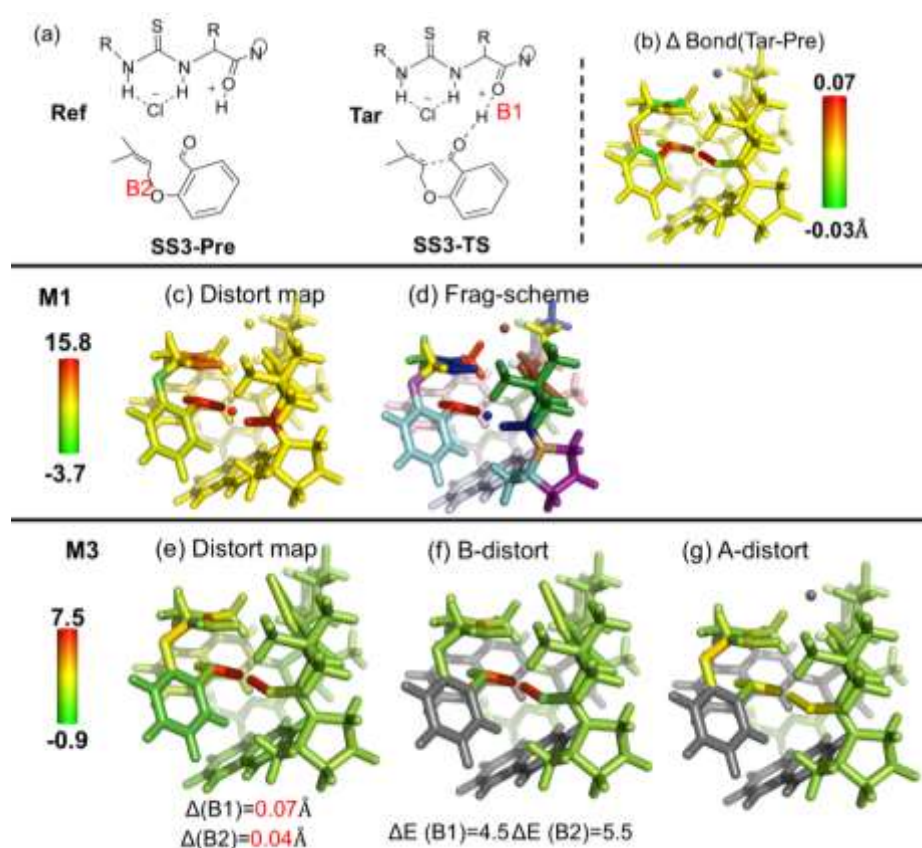
Figure S44 (a) Overview of the HCl-mediated Prins reaction. (b) Bond length change (Δbond in Å) from the reference form (**Ref**; Pre, **SS3-R**) to the target form (**Tar**; TS, **SS3-TS**). (c) Distortion energy distribution (kcal/mol) and (d) fragmentation using **M1** scheme. (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M3** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at CPCM B3LYP/6-31G(d) level.
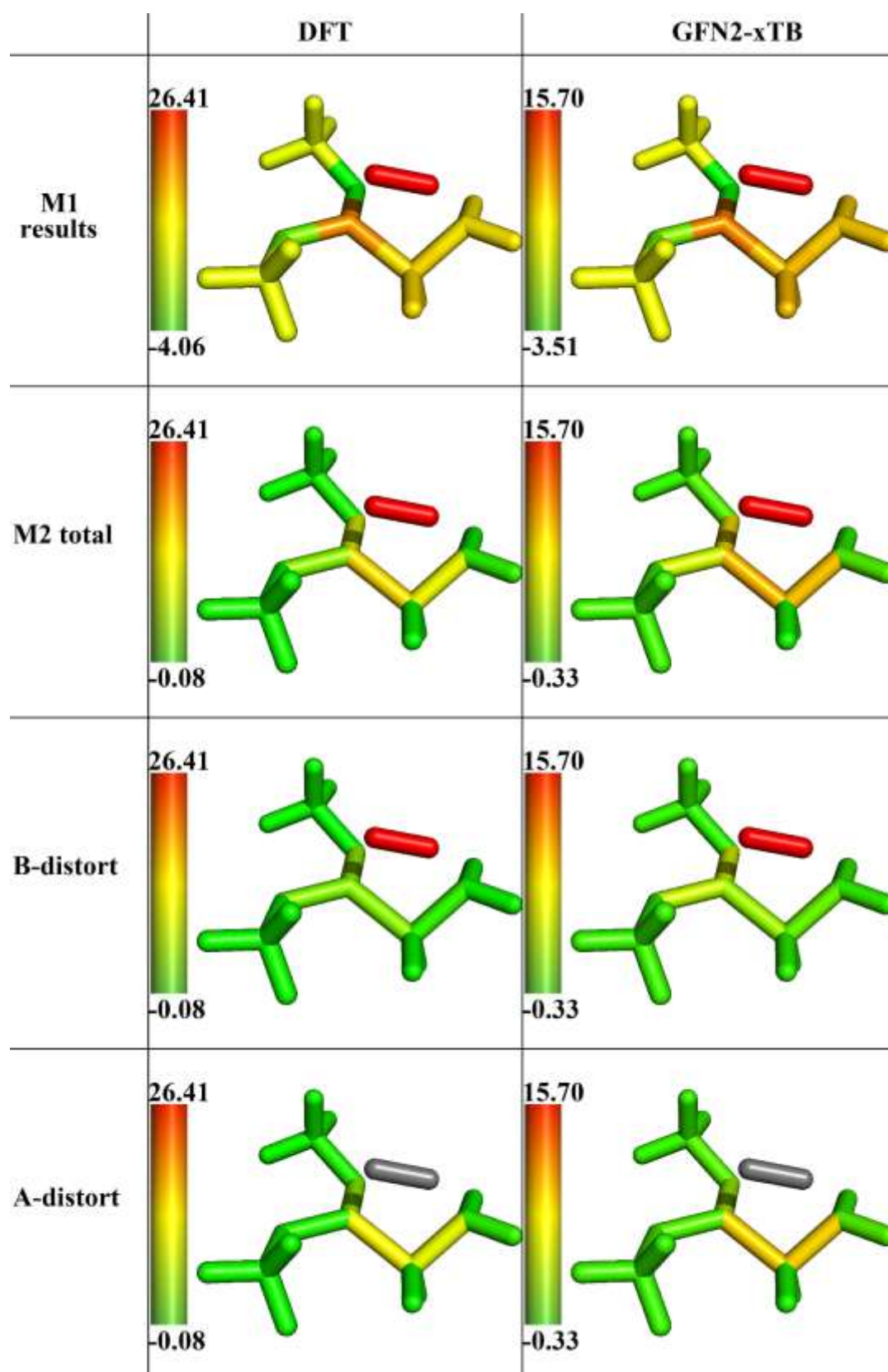
Figure S45 Distortion energy distribution (kcal/mol in log scale) of HCl-mediated Prins reaction using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. All energies were computed using DFT (CPCM B3LYP/6-31G(d)), GFN2-xTB and ANI-2x methods.

Figure S46 Distortion energy distribution (kcal/mol in log scale) of HCl-mediated Prins reaction using **M3** scheme and its contribution from bond, angle and fragment terms. All energies were computed using DFT (CPCM B3LYP/6-31G(d)), GFN2-xTB and ANI-2x methods

5.2.2. Lewis-base Catalyzed Sulfeno-functionalization

In this system, the transition state (TS) as **Tar** and its precursor as **Ref** for the Lewis-base catalyzed sulfeno-functionalization of alkenes[S24] were taken from ref. S24 for our D2AF analysis using DFT (SMD B3LYP/6-31G(d)) and GFN2-xTB methods.
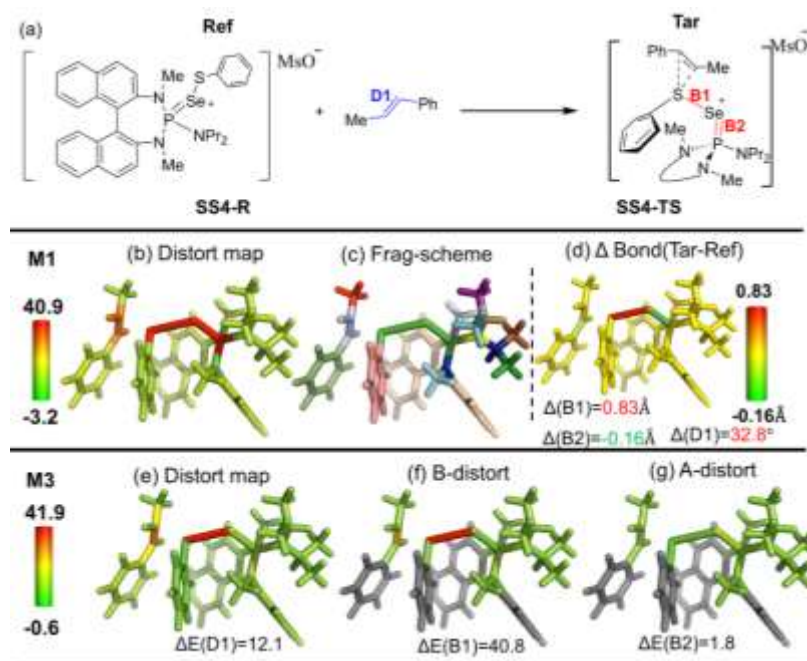
Figure S47 (a) Overview of the Lewis-base catalyzed sulfeno-functionalization. (b) Distortion energy distribution (kcal/mol) and (c) fragmentation using **M1** scheme. (d) Bond length change (Δbond in Å) from the reference form (**Ref**; R, **SS4-R**) to the target form (**Tar**; TS, **SS4-TS**). (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M3** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at SMD B3LYP/6-31G(d) level.
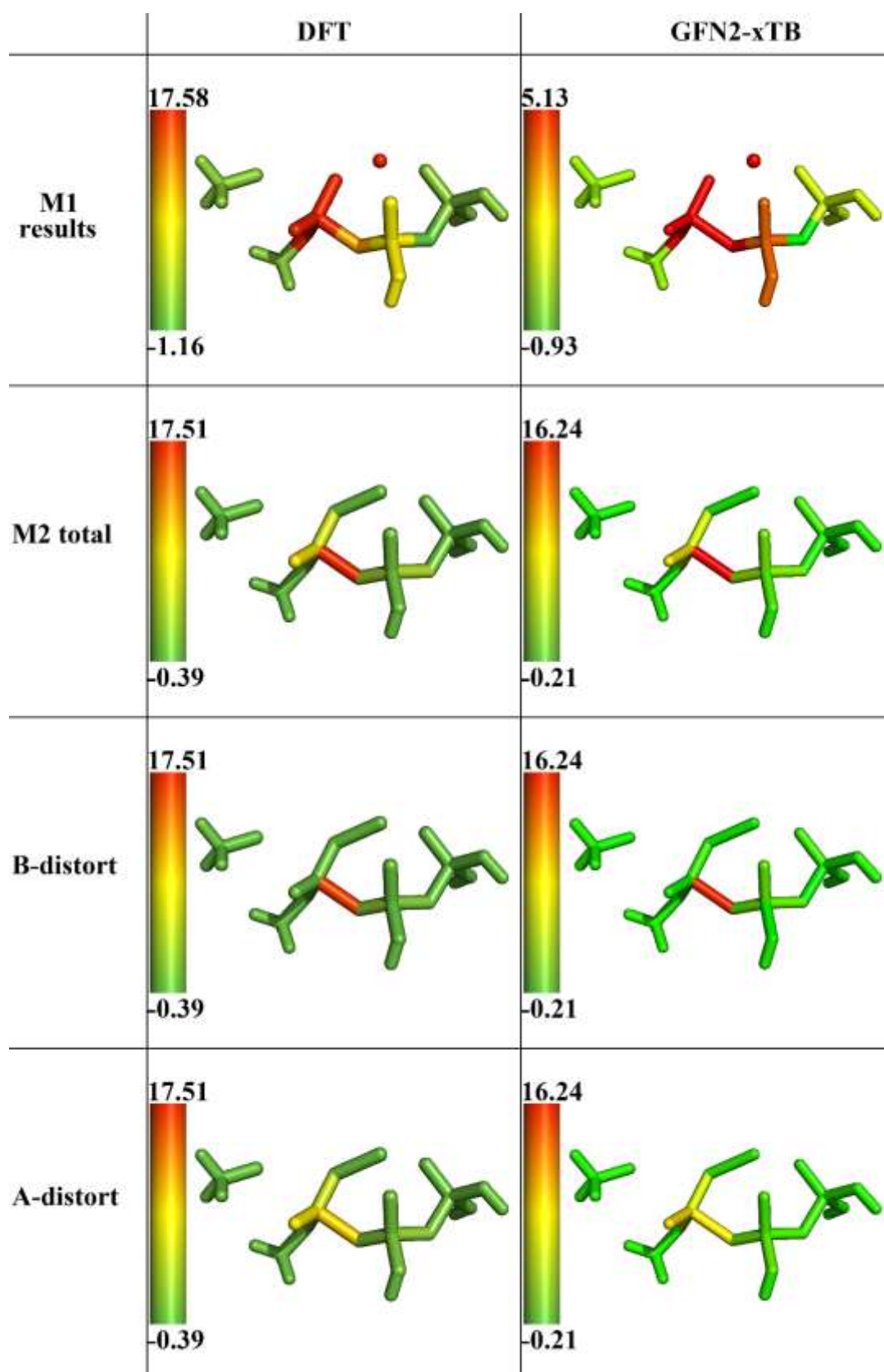
Figure S48 Distortion energy distribution (kcal/mol in log scale) of Lewis-base catalyzed sulfeno-functionalization using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. All energies were computed using DFT (SMD B3LYP/6-31G(d)) and GFN2-xTB methods.

Figure S49 Distortion energy distribution (kcal/mol in log scale) of Lewis-base catalyzed sulfeno-functionalization using **M3** scheme and its contribution from bond, angle and fragment terms. All energies were computed using DFT (SMD B3LYP/6-31G(d)) and GFN2-xTB methods.

58

5.3. Excited system

5.3.1. Radical Sulfinyl-Smiles Rearrangement

The preceding intermediate and its corresponding TS for an asymmetric radical sulfinyl-Smiles rearrangement were taken from ref. S25 for D2AF analysis using DFT (IEFPCM B3LYP/6-31+G(d, p)), GFN2-xTB and ANI-2x methods.

Figure S50 (a) Overview of the radical sulfinyl-Smiles rearrangement. (b) Bond length change (Δbond in Å) from the reference form (**Ref**; R, **SS5-R**) to the target form (**Tar**; TS, **SS5-TS**). (c) Distortion energy distribution (kcal/mol) and (d) fragmentation using **M1** scheme. (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M3** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at IEFPCM B3LYP/6-31+G(d, p) level.
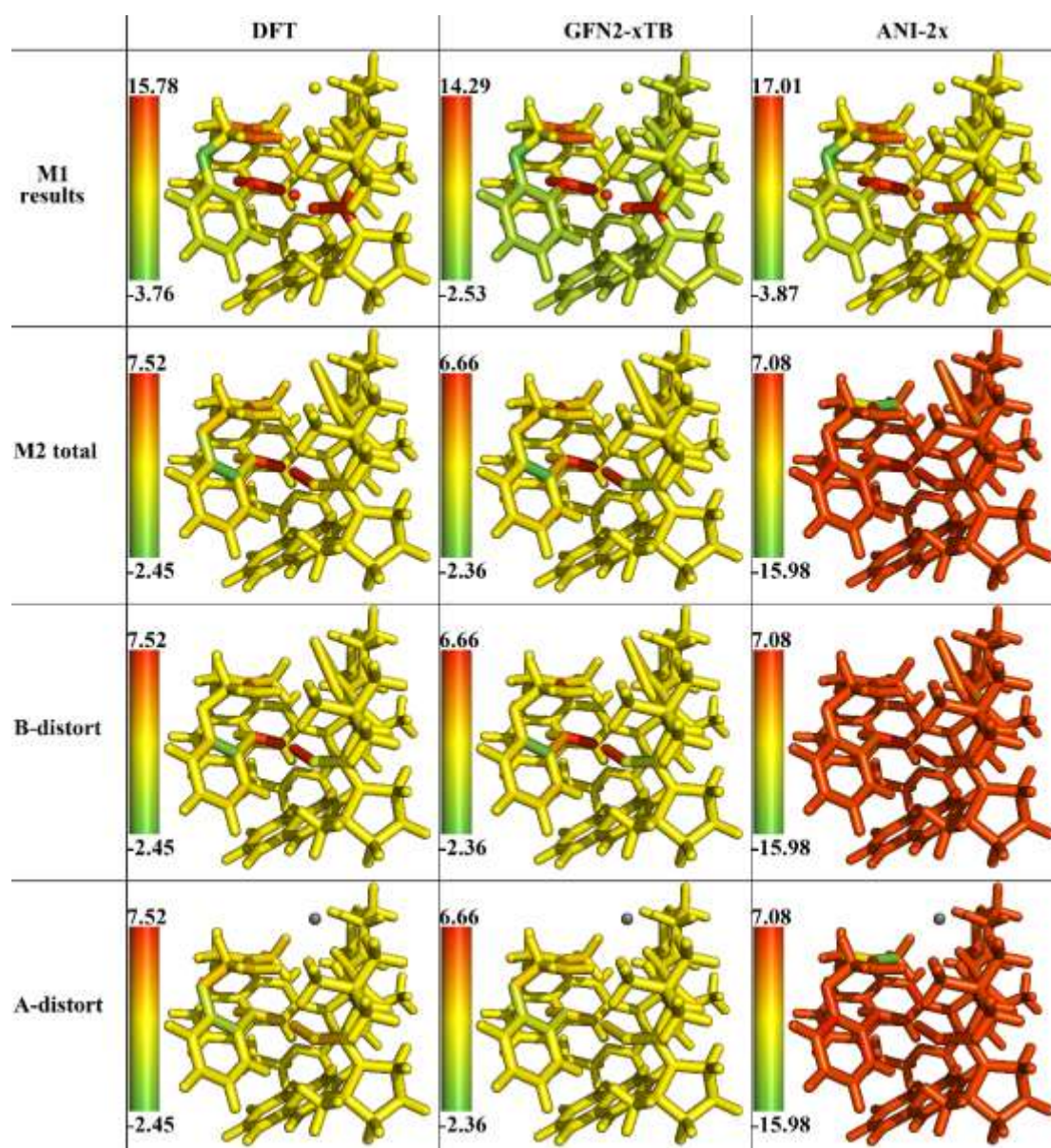
Figure S51 Distortion energy distribution (kcal/mol in log scale) of radical sulfinyl-Smiles rearrangement using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. All energies were computed using DFT (IEFPCM B3LYP/6-31+G(d, p)), GFN2-xTB and ANI-2x methods.
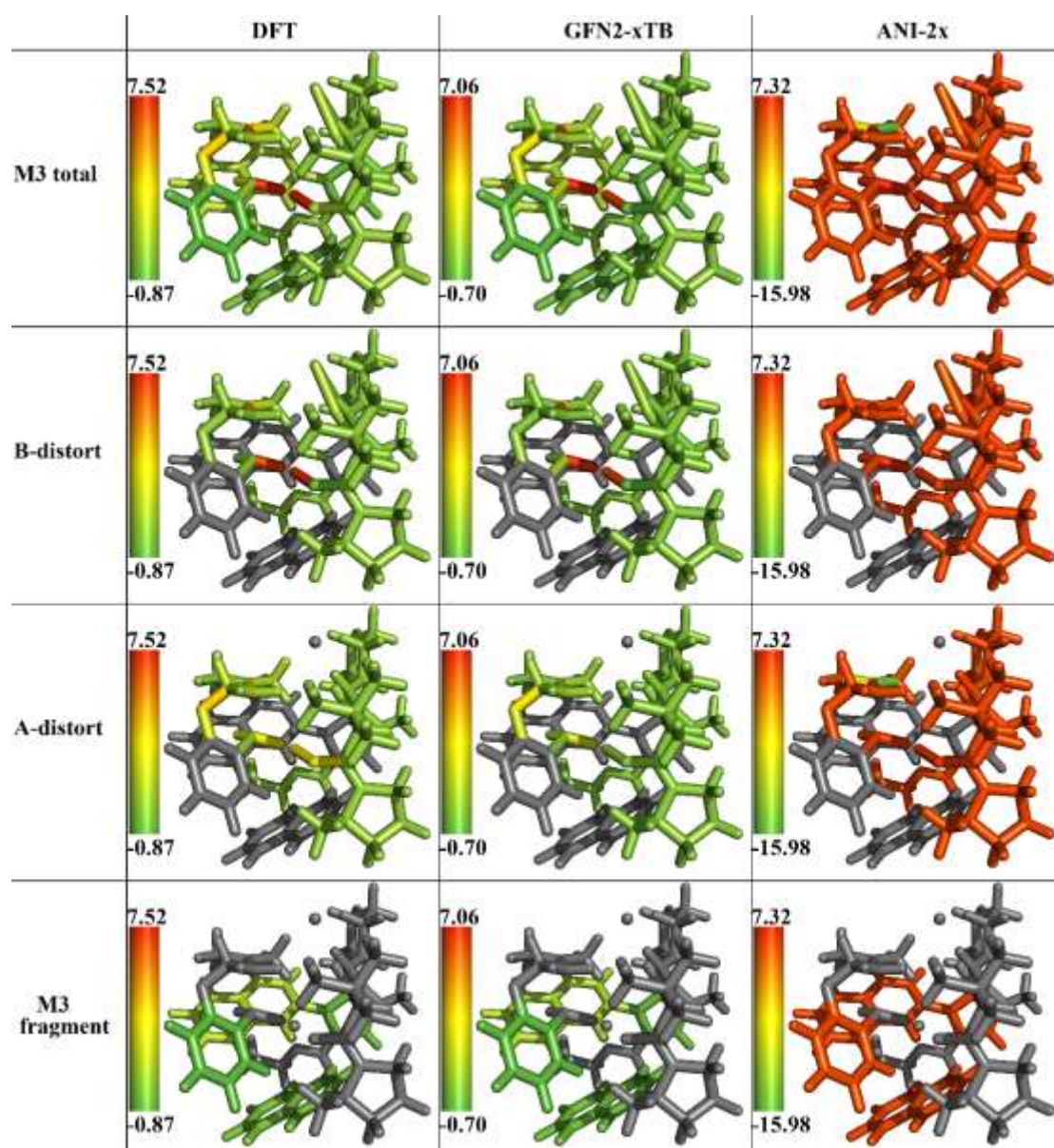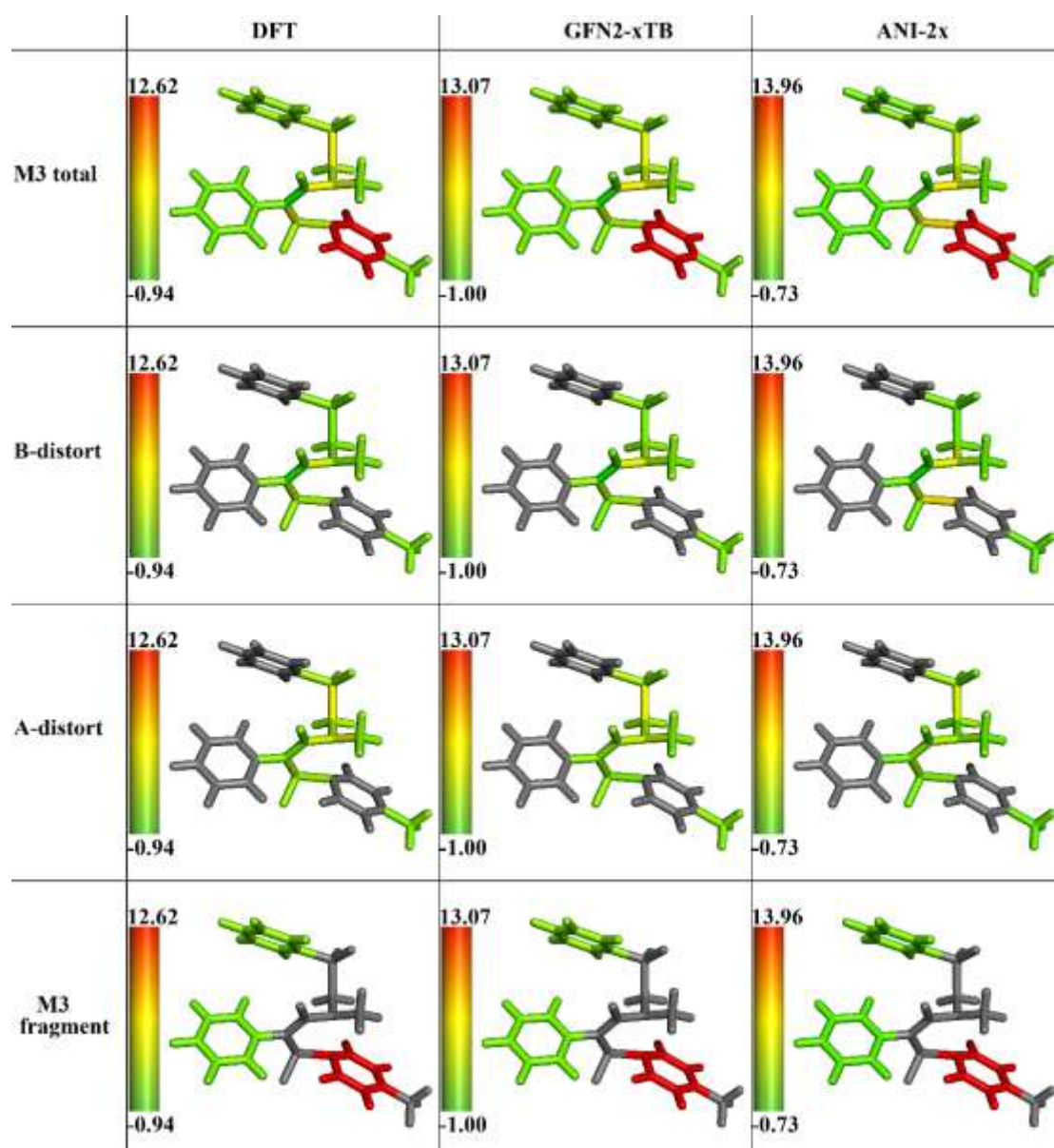
Figure S52 Distortion energy distribution (kcal/mol in log scale) of radical sulfinyl-Smiles rearrangement using **M3** scheme and its contribution from bond, angle and fragment terms. All energies were computed using DFT (IEFPCM B3LYP/6-31G(d)), GFN2-xTB and ANI-2x methods.

5.4. Supramolecular system
5.4.1. Diels-Alder Reaction in Pd-Cage

Self-assembled cages are novel supramolecular platforms to explore certain catalysis.[S26] In this section, a Diels-Alder addition of two dienes (denoted as q1, d1 in the ref. S26) within a $[Pd(pyridine)_4]^{2+}$ cage (denoted as C-1 in the in ref. S26) is considered. The reactant structures were taken from the ref. S26 as **Ref** and **Tar** for D2AF analysis using DFT (M06-2X/def2-TZVP) and GFN2-xTB methods.
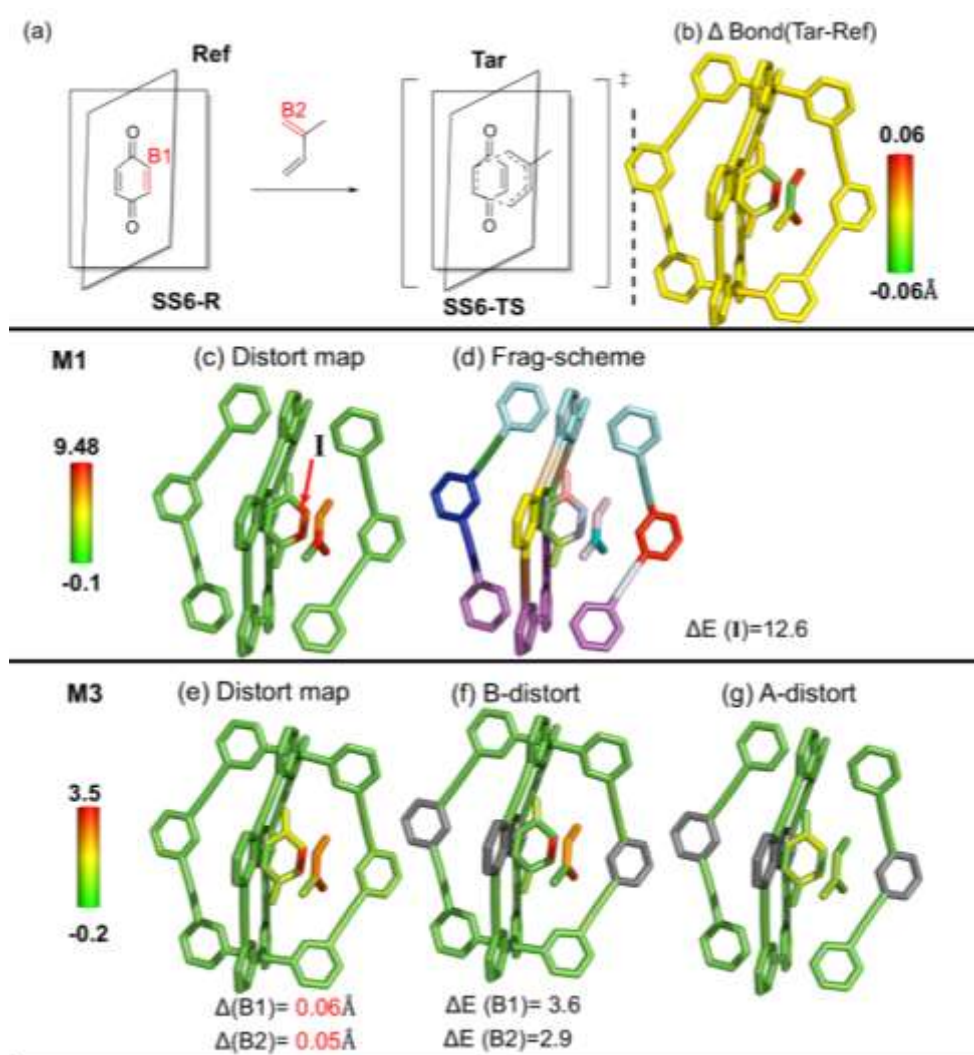
Figure S53 (a) Overview of the Diels-Alder reaction within the Pd-cage. (b) Bond length change (Δbond in Å) from the reference form (**Ref**; R, **SS6-R**) to the target form (**Tar**; TS, **SS6-TS**). (c) Distortion energy distribution (kcal/mol) and (d) fragmentation using **M1** scheme. (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M3** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at M06-2X/def2-TZVP level.
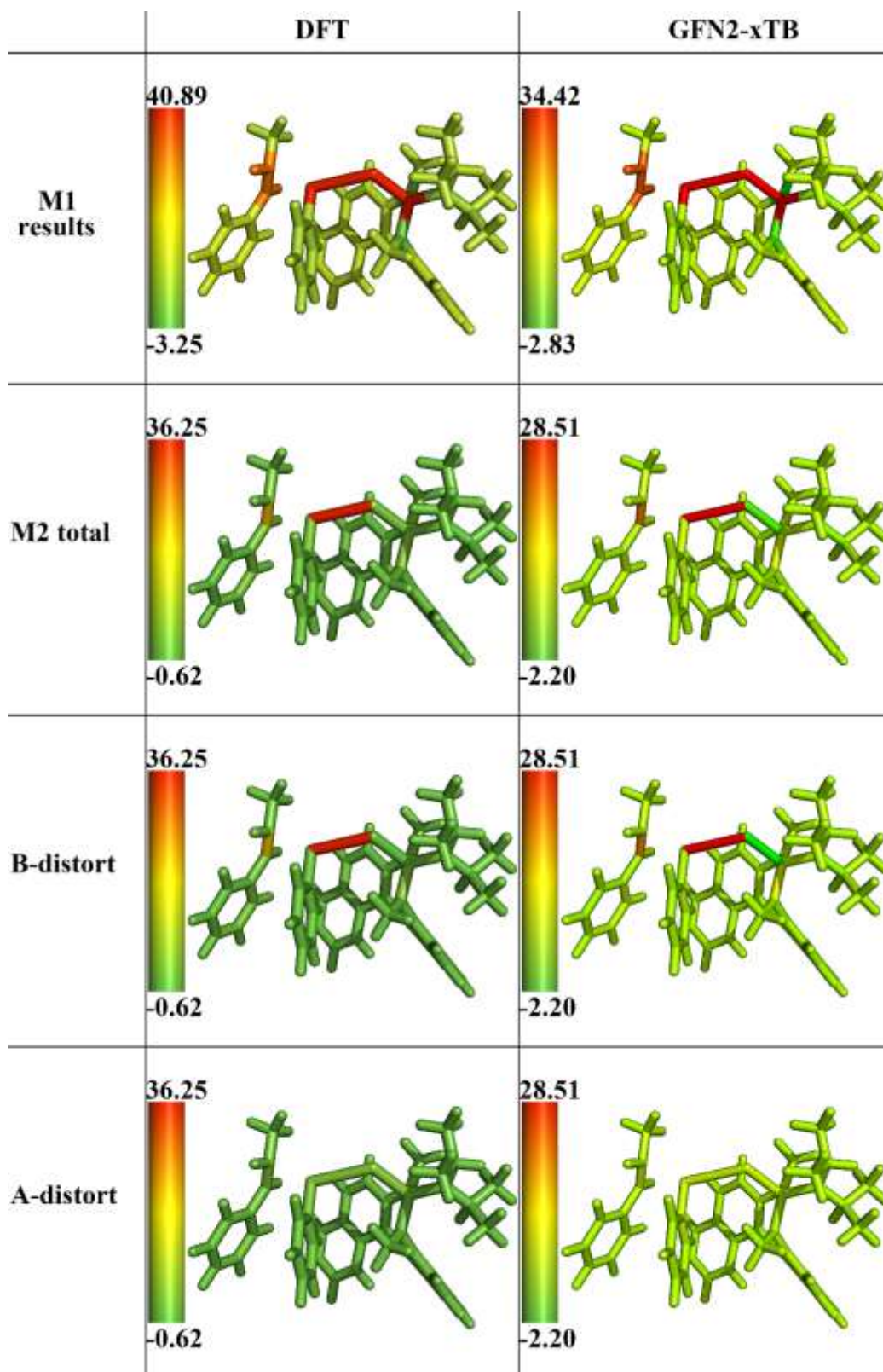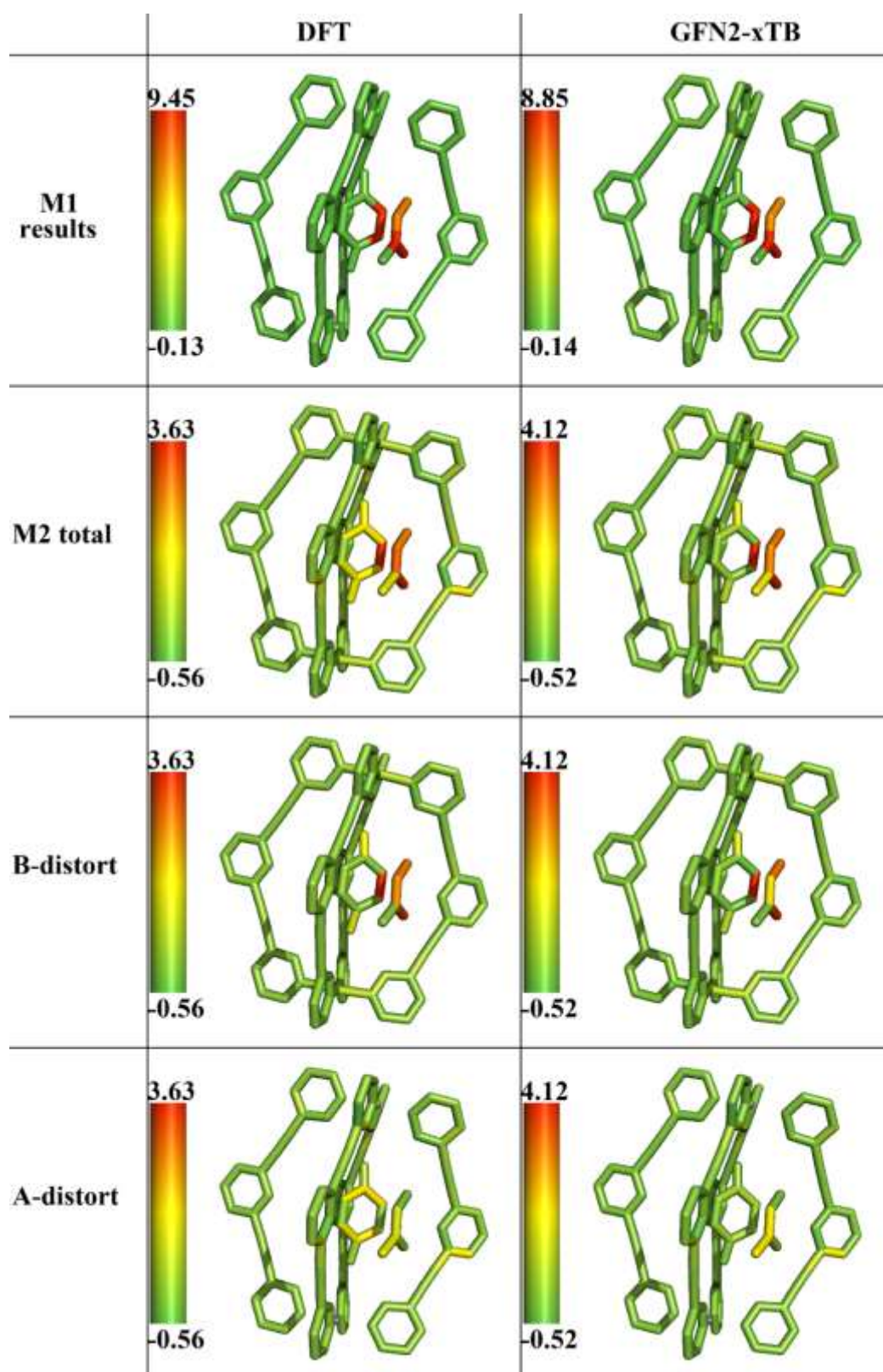
Figure S54 Distortion energy distribution (kcal/mol in log scale) of Diels-Alder reaction inside the Pd-cage using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. All energies were computed using DFT (M06-2X/def2-TZVP) and GFN2-xTB methods.
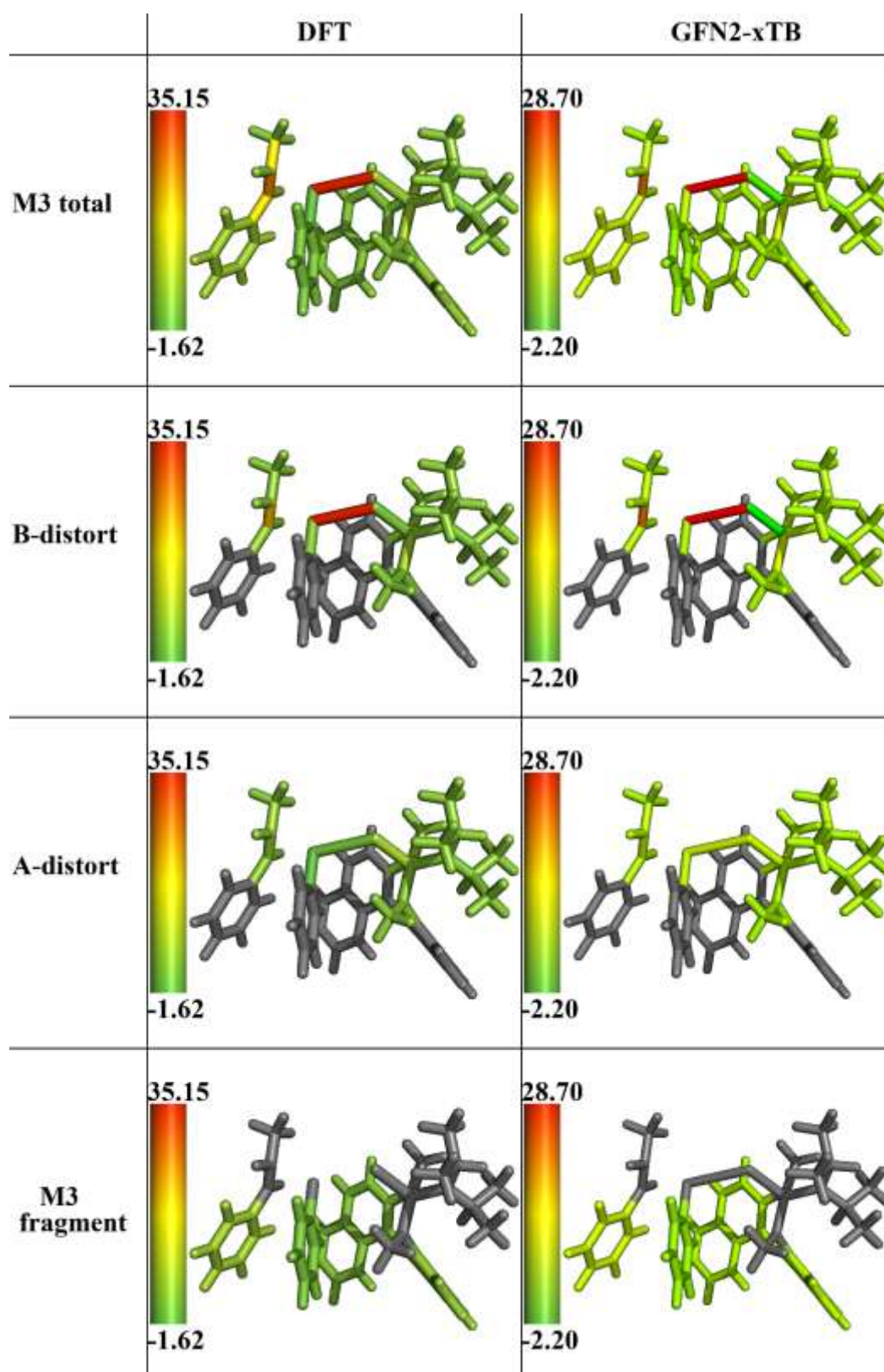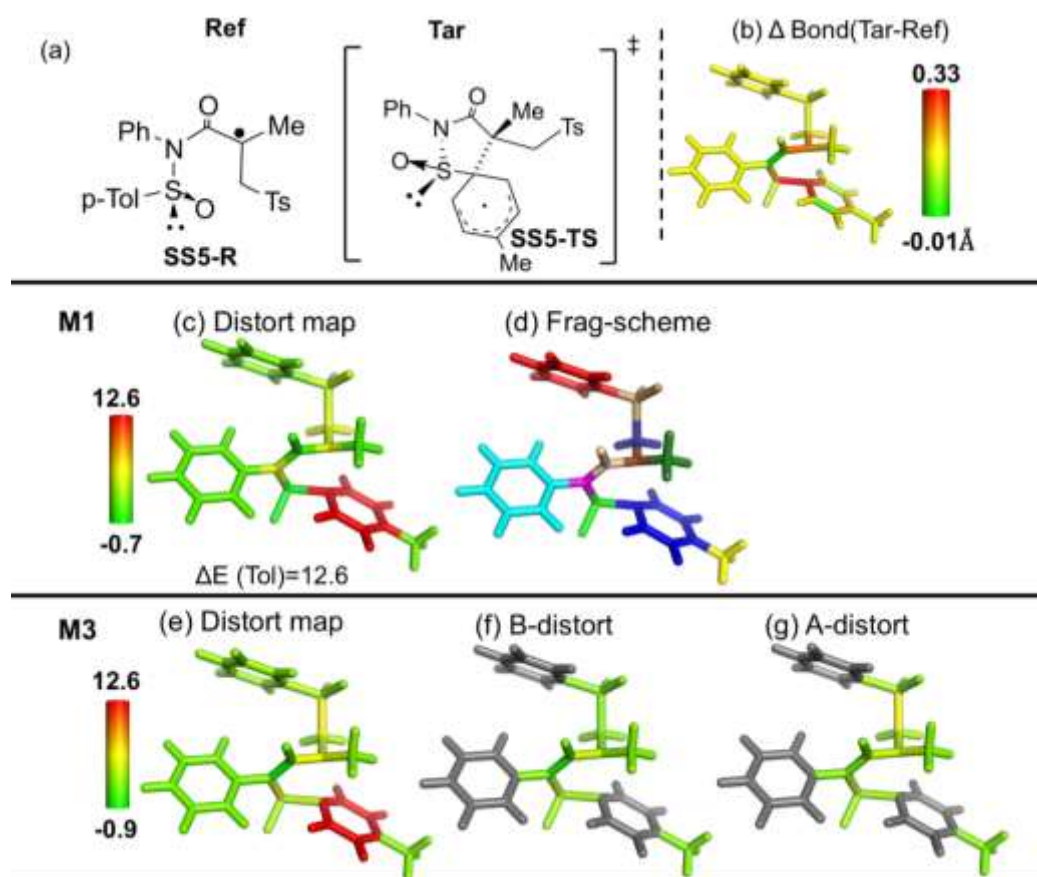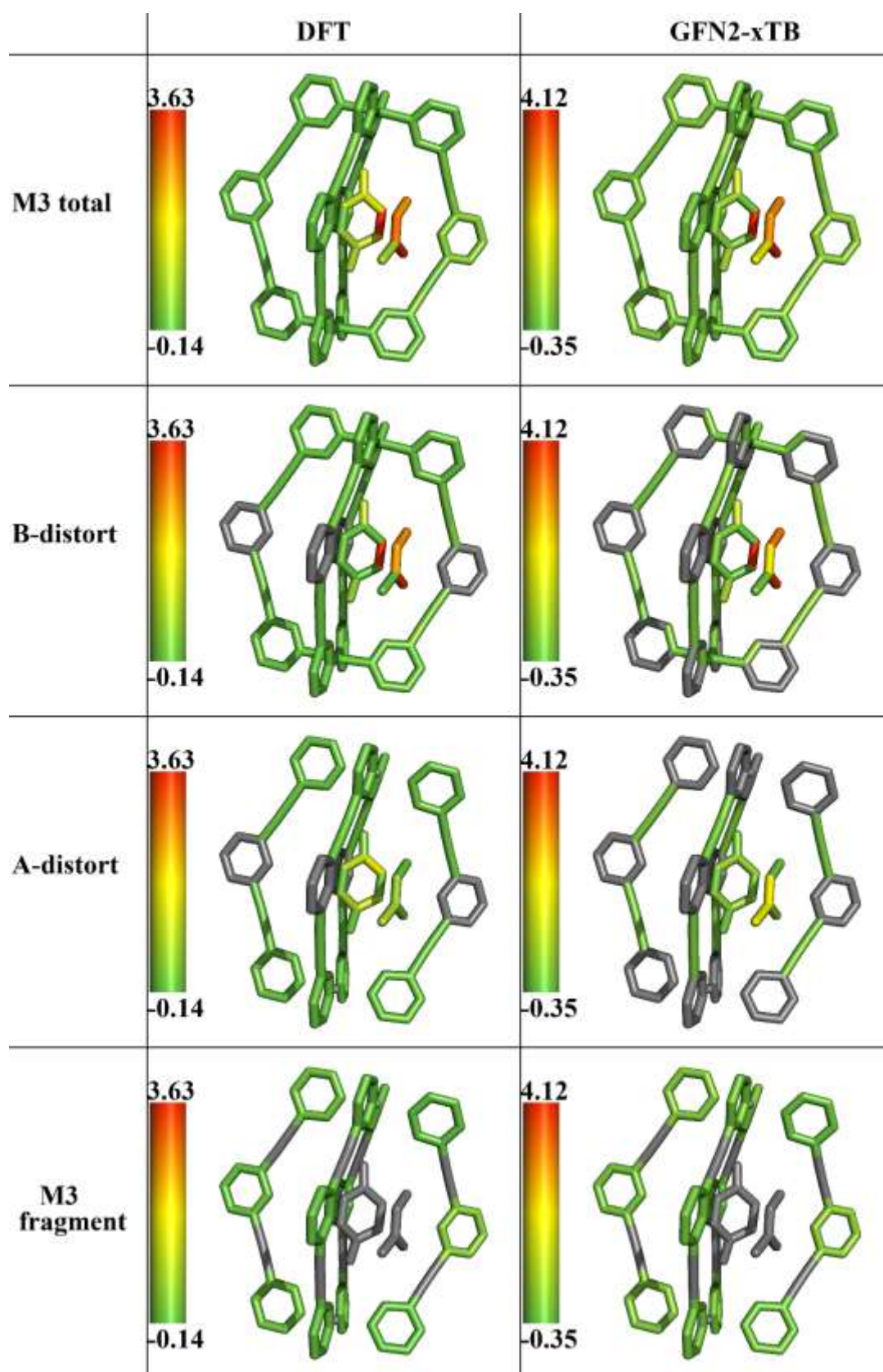
Figure S55 Distortion energy distribution (kcal/mol in log scale) of Diels-Alder reaction in the Pd-cage using **M3** scheme and its contribution from bond, angle and fragment terms. All energies were computed using DFT (M06-2X/def2-TZVP) and GFN2-xTB methods.

5.5. Coordination systems
5.5.1. Palladium-Catalyzed Cross-Coupling Reaction

A famous Pd-catalyzed cross-coupling reaction[S27] was taken as a simple example of our coordination systems. The structures of reactant and catalyst ($C_5H_3NCl_2$ and $Pd(PH_3)_2$) and corresponding transition state taken from ref. S27 were treated as **Ref** and **Tar** for our D2AF analysis using DFT (B3LYP/(Lanl2DZ for Pd and 6-31G(d) for other elements)) method.
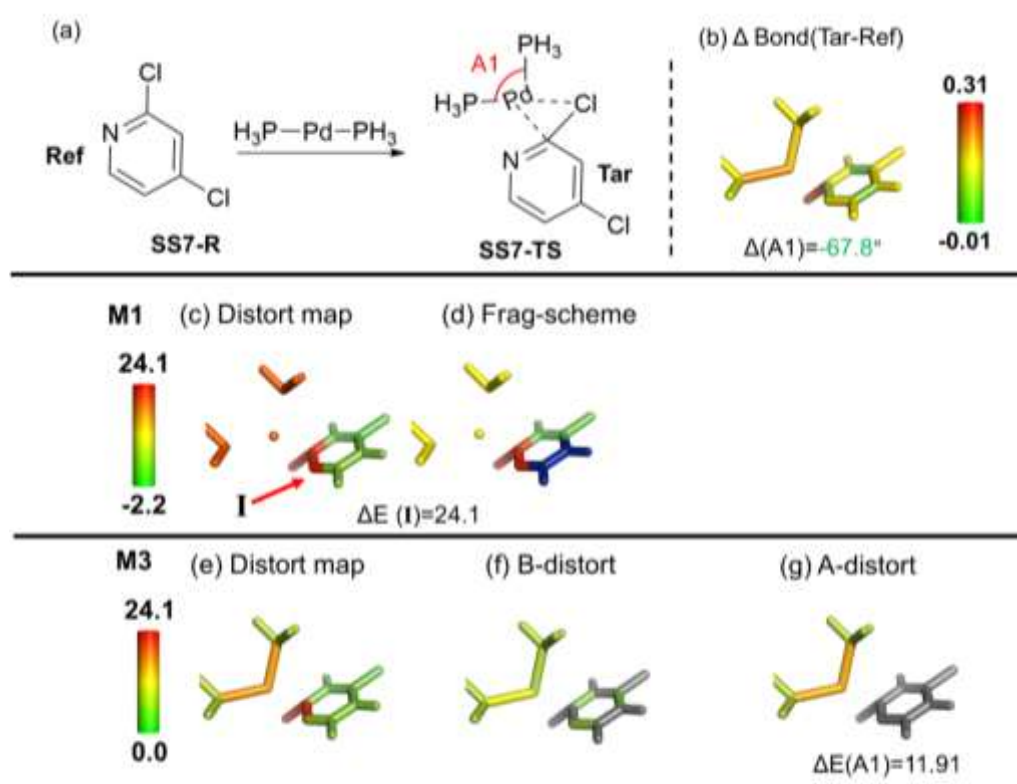
Figure S56 (a) Overview of the Pd-catalyzed cross-coupling reaction. (b) Bond length change (Δbond in Å) from the reference form (**Ref**; R, **SS7-R**) to the target form (**Tar**; TS, **SS7-TS**). (c) Distortion energy distribution (kcal/mol) and (d) fragmentation using **M1** scheme. (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M3** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at B3LYP/(Lanl2DZ for Pd and 6-31G(d) for other elements) level.

5.5.2. Al-catalyzed Ring-opening Transesterification

The effects of a series of Al complexes which can catalyze ring-opening transesterification polymerization reactions were studied.[S28] Among these Al complexes, one TS and its precomplex were taken from ref. S28 as **Tar** and **Ref** for our D2AF analysis using DFT (SMD M06-L/6-31+G(d, p)) and GFN2-xTB methods.
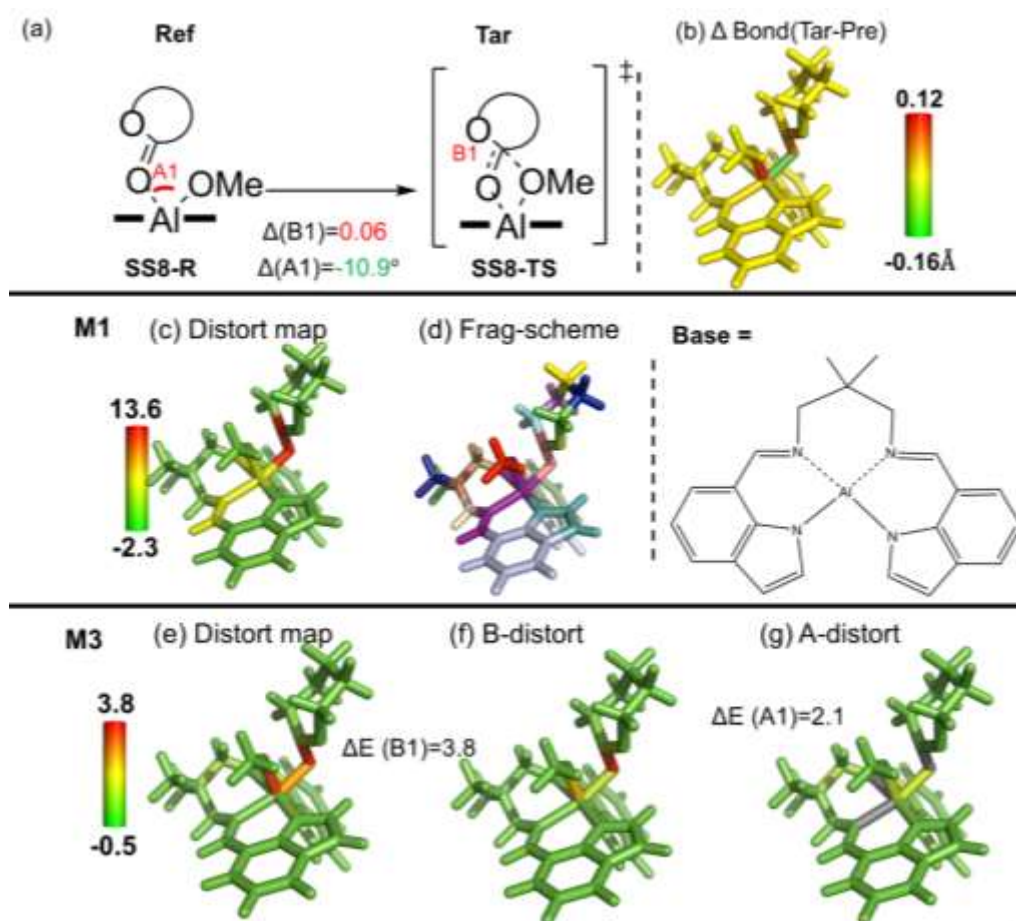
Figure S57 (a) Overview of the Al-catalyzed ring-opening transesterification. (b) Bond length change (Δbond in Å) from the reference form (**Ref**; Pre, **SS8-R**) to the target form (**Tar**; TS, **SS8-TS**). (c) Distortion energy distribution (kcal/mol) and (d) fragmentation using **M1** scheme. (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M3** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed using DFT (SMD M06-L/6-31+G(d, p)) method.
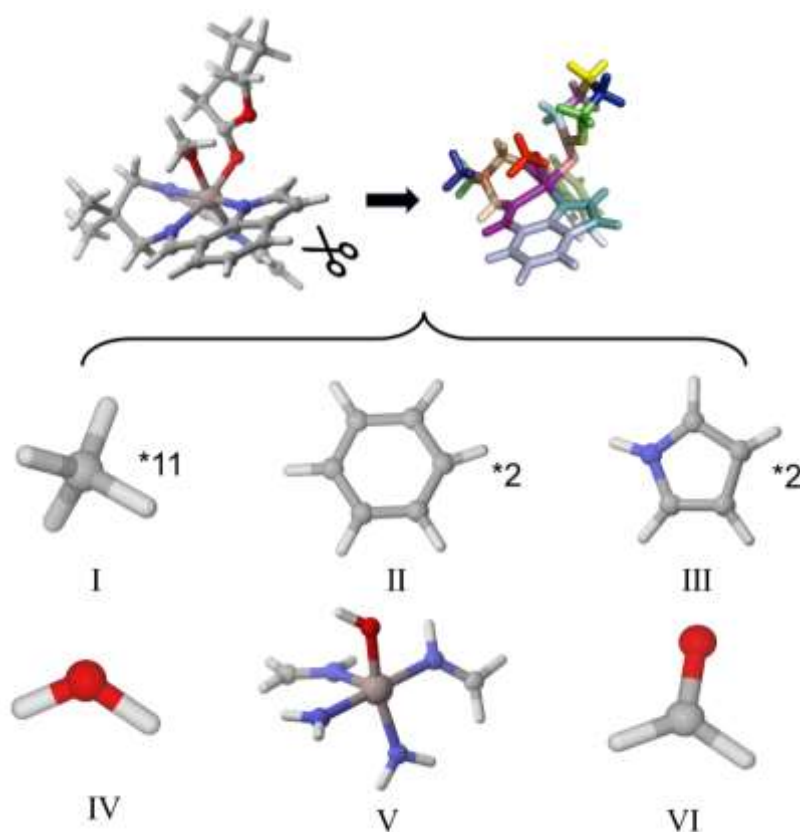
Figure S58 All molecular fragments generated by the **M1** scheme from the Al-catalyzed ring-opening transesterification (**SS8**)

As shown in Figure S58, the whole system of the Al-catalyzed system is divided into six kinds of fragments using the **M1** scheme. Eleven fragments **I** come from eleven sp^3 carbon atoms. Two fragments **II** and two fragments **III** come from separation of two indoles in order to localize the distortion. The fragments **IV** and **VI** come from the sp^3 oxygen and carbonyl within hexano-6-lactone. The specialized fragment **V**, which includes the Al center, along with its coordination nitrogen atoms and N=CH$_2$ are set to capture the crucial coordination interactions.
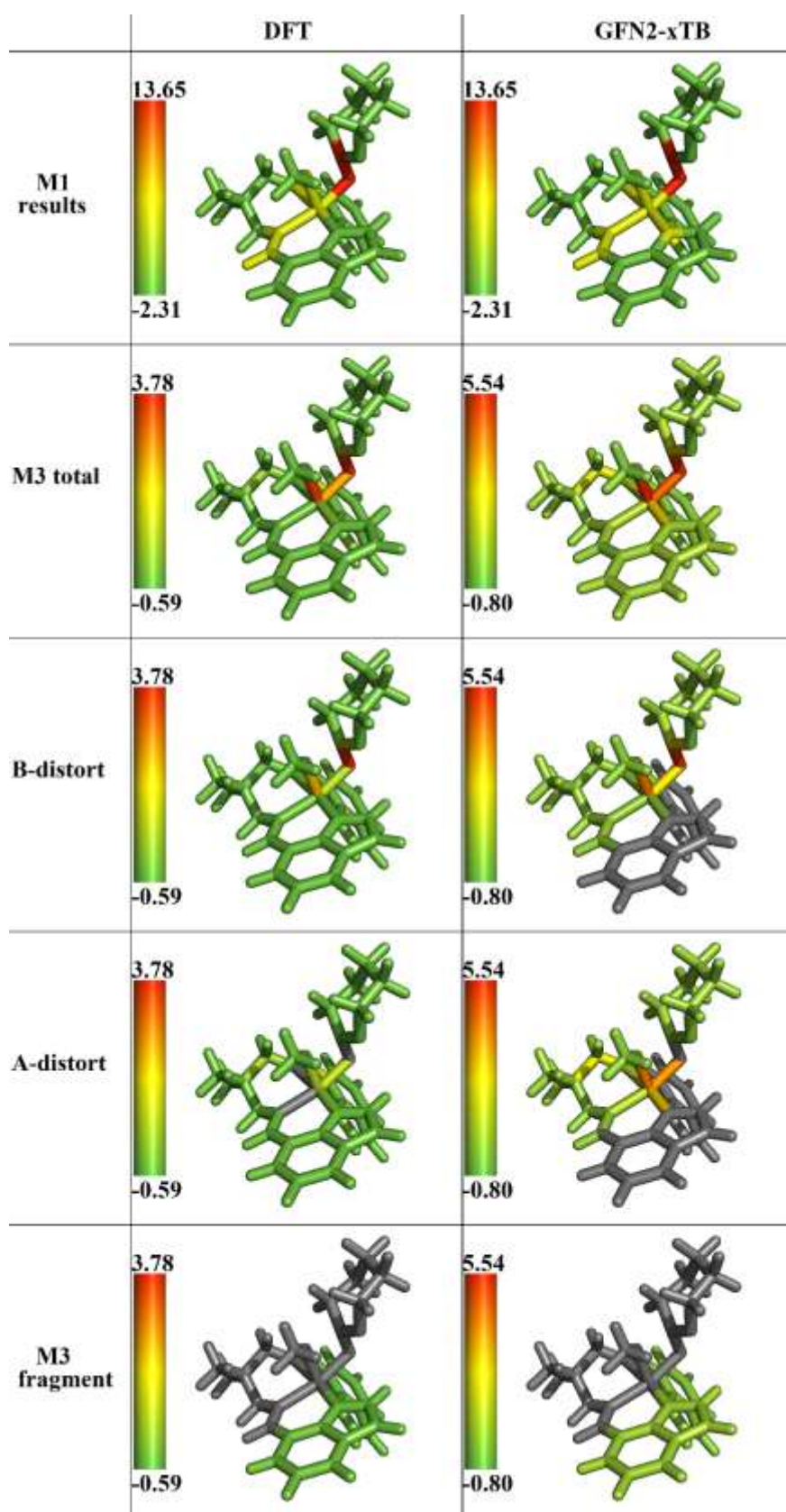
Figure S59 Distortion energy distribution (kcal/mol in log scale) of Al-catalyzed ring-opening transesterification using **M1** scheme and total distortion and its contribution from bond, angle and fragment terms using **M3** scheme. All energies were using DFT (SMD M06-L/6-31+G(d, p)) and GFN2-xTB methods.

5.6. Biochemical systems

5.6.1. Ni-Dependent Lactate Racemase

A Ni-dependent lactate racemase (LarA) and its unique tethered NAD-like cofactor was studied.[S29] The structure of TS and its corresponding reactant (denoted TS_{C4-LS} and RC_{LS} in the ref. S29) were taken from the ref. S29 as **Tar** and **Ref** for our D2AF analysis using DFT (SMD M06/6-311G+(d, p)) and GFN2-xTB methods.
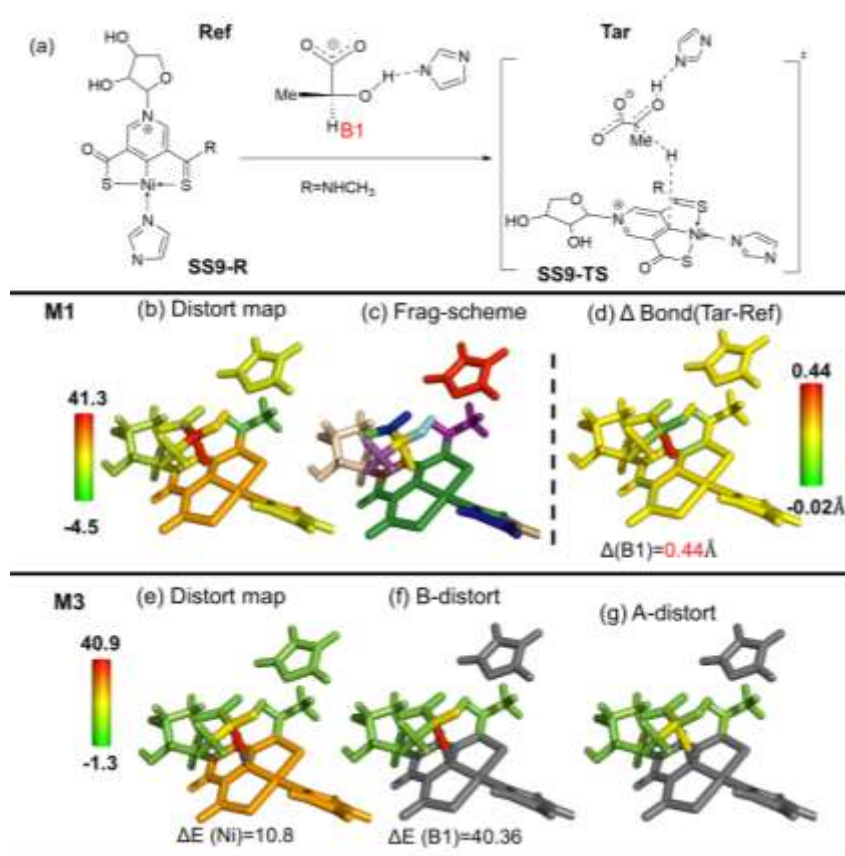
Figure S60 (a) Overview of the Ni-dependent lactate racemase. (b) Distortion energy distribution (kcal/mol) and (c) fragmentation using **M1** scheme. (d) Bond length change (Δbond in Å) from the reference form (**Ref**; R, **SS9-R**) to the target form (**Tar**; TS, **SS9-TS**). (e) Total distortion distribution (kcal/mol), its distortion contribution from (f) bond and (g) angle terms using **M3** scheme. The key geometrical changes and their corresponding distortion energy (ΔE) are also given. All energies were computed at SMD M06/6-311G+(d, p) level.
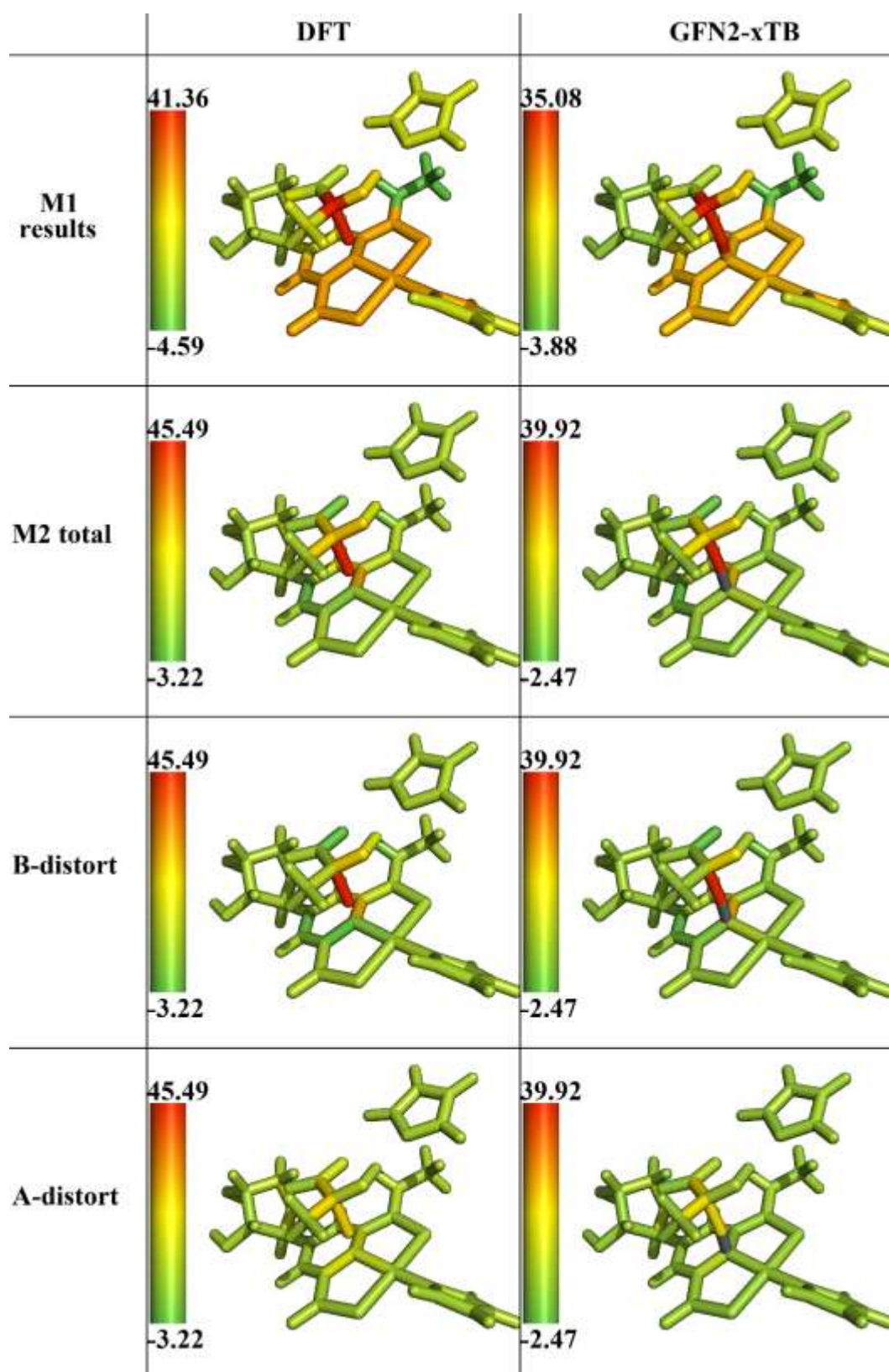
Figure S61 Distortion energy distribution (kcal/mol in log scale) of Ni-dependent lactate racemase using **M1** scheme and total distortion and its contribution from bond and angle terms using **M2** scheme. All energies were computed using DFT (SMD M06/6-311G+(d, p)) and GFN2-xTB methods.
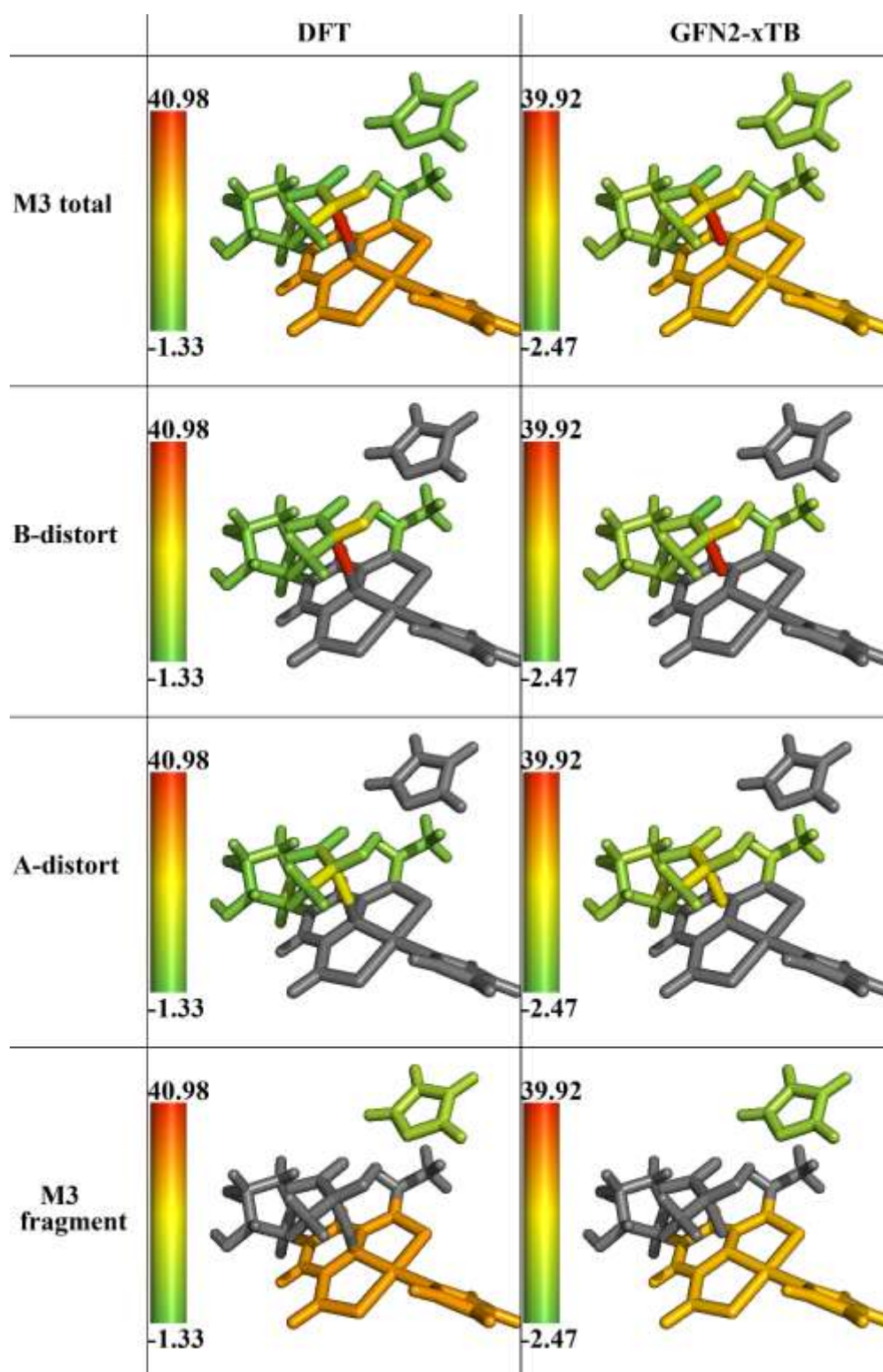
Figure S62 Distortion energy distribution (kcal/mol in log scale) of Ni-dependent lactate racemase using **M3** scheme and its contribution from bond, angle and fragment terms. All energies were computed using DFT (SMD M06/6-311G+(d, p)) and GFN2-xTB methods.

5.6.2. Oseltamivir in Influenza Neuraminidase

The drug Oseltamivir in influenza neuraminidase from the X-ray and QR refined structures were studied, in which the latter one were obtained using ONIOM(ωB97-D3/6-31G(d):GFN2-xTB:Amber) method.[S30] The optimized drug structure in the gas phase using ωB97X-D3/6-31G(d) was set as **Tar**, and X-ray and QR refined structures were set as two **Tar** for D2AF analysis using DFT (ωB97-D3/6-31G(d)) GFN2-xTB, ANI-2x and AIQM1 methods.
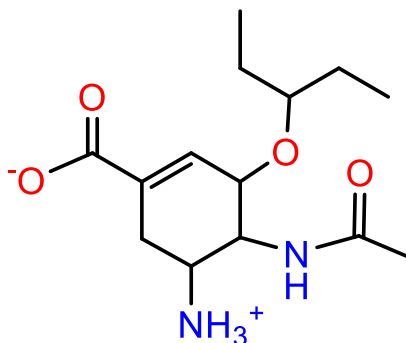
Figure S63 Oseltamivir structure
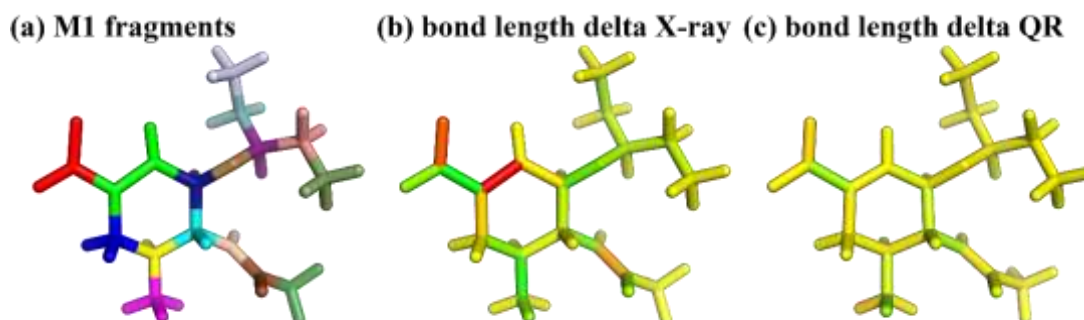
Figure S64 (a) Fragmentation scheme of Oseltamivir using different colors. Structures (optimized structure in the gas phase using ωB97X-D3/6-31G(d), X-ray structure and QR refined structure) taken from the ref. S30. Difference of bond length of (b) X-ray structure and (c) QR-refined structure relative to the optimized structure in the gas phase (color bar green to red: red positive, green negative, yellow: no change).
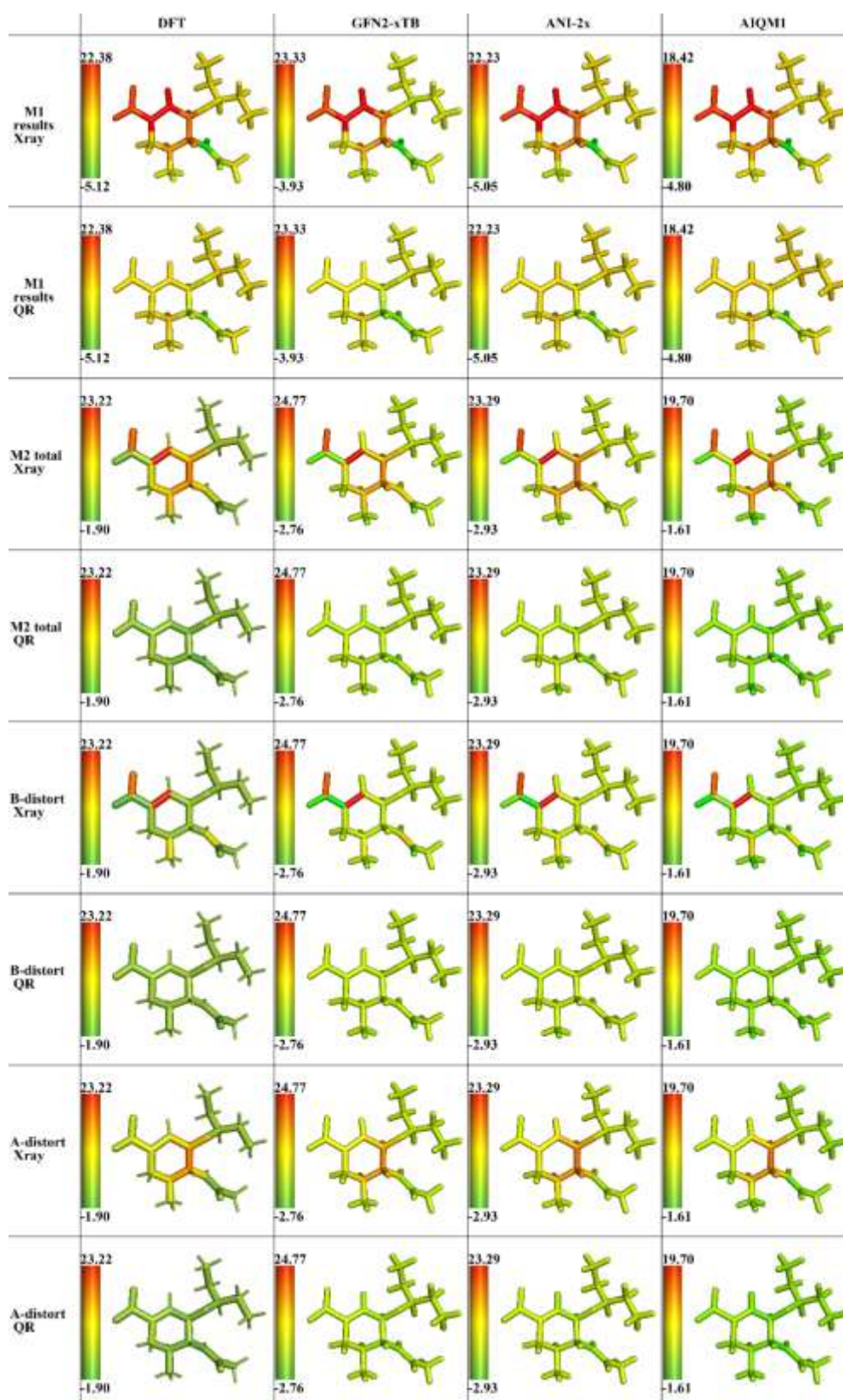
Figure S65 Distortion energy distribution (kcal/mol in log scale) of Oseltamivir's two different **Tar** forms (X-Ray and QR refined structures) including **M1** and **M2** distortion map. Bond and angle contribution of two forms in **M2** are also presented. All energies were computed using DFT (ωB97X-D3/6-31G(d)), GFN2-xTB, ANI-2x and AIQM1 methods.

73

bond	X-ray	QR
b1	1.37	1.27
b2	1.21	1.26
b3	1.47	1.51
b4	1.53	1.34
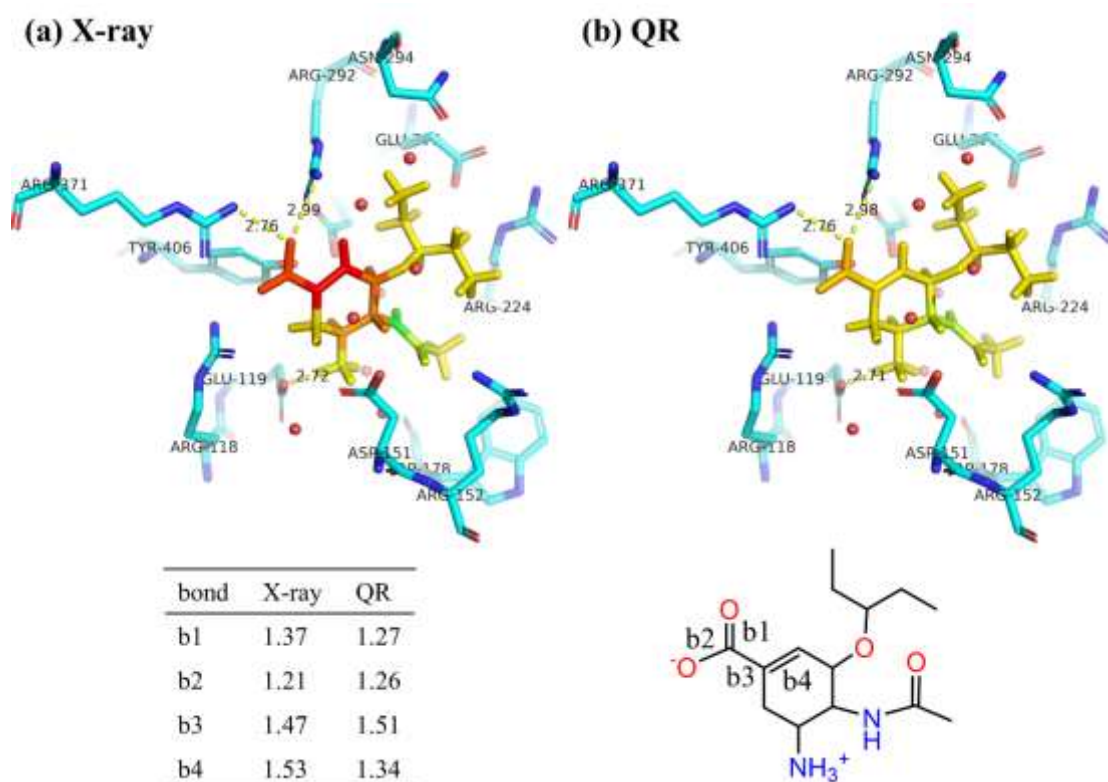
Figure S66 The key distances between Oseltamivir and neighboring residues in influenza neuraminidase from the X-ray and QR refined structures. The key bond lengths (Å) are listed. Oseltamivir is colored based on distortion distribution using **M1** scheme while the optimized structure in the gas phase were set as the reference form.

5.7. Multiple-structures System

5.7.1. S$_N$2 (IRC)

An IRC calculation of one S$_N$2 reaction (NC$^-$ and CH$_3$Cl) was also applied to show the change in (relative) distortion distribution during the reaction process.[S15] The structures (without external electric fields case) were taken from the ref. S15 as the target states for our D2AF analysis throughout the reaction process using DFT(PCM M06-2X/6-31+G(d)), GFN2-xTB and ANI-2x methods.

$$NC^- + CH_3Cl \longrightarrow \left[NC\text{-}\text{-}\overset{\overset{H}{\diagup}\ \overset{H}{}}{\underset{H}{C}}\text{-}\text{-}Cl \right]^{\ddagger} \longrightarrow NCCH_3 + Cl^-$$

$$\textbf{R} \qquad\qquad\qquad \textbf{TS} \qquad\qquad\qquad \textbf{P}$$
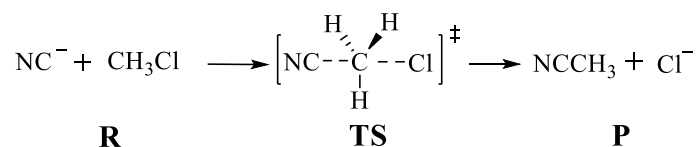
Figure S67 Scheme of a S$_N$2 reaction
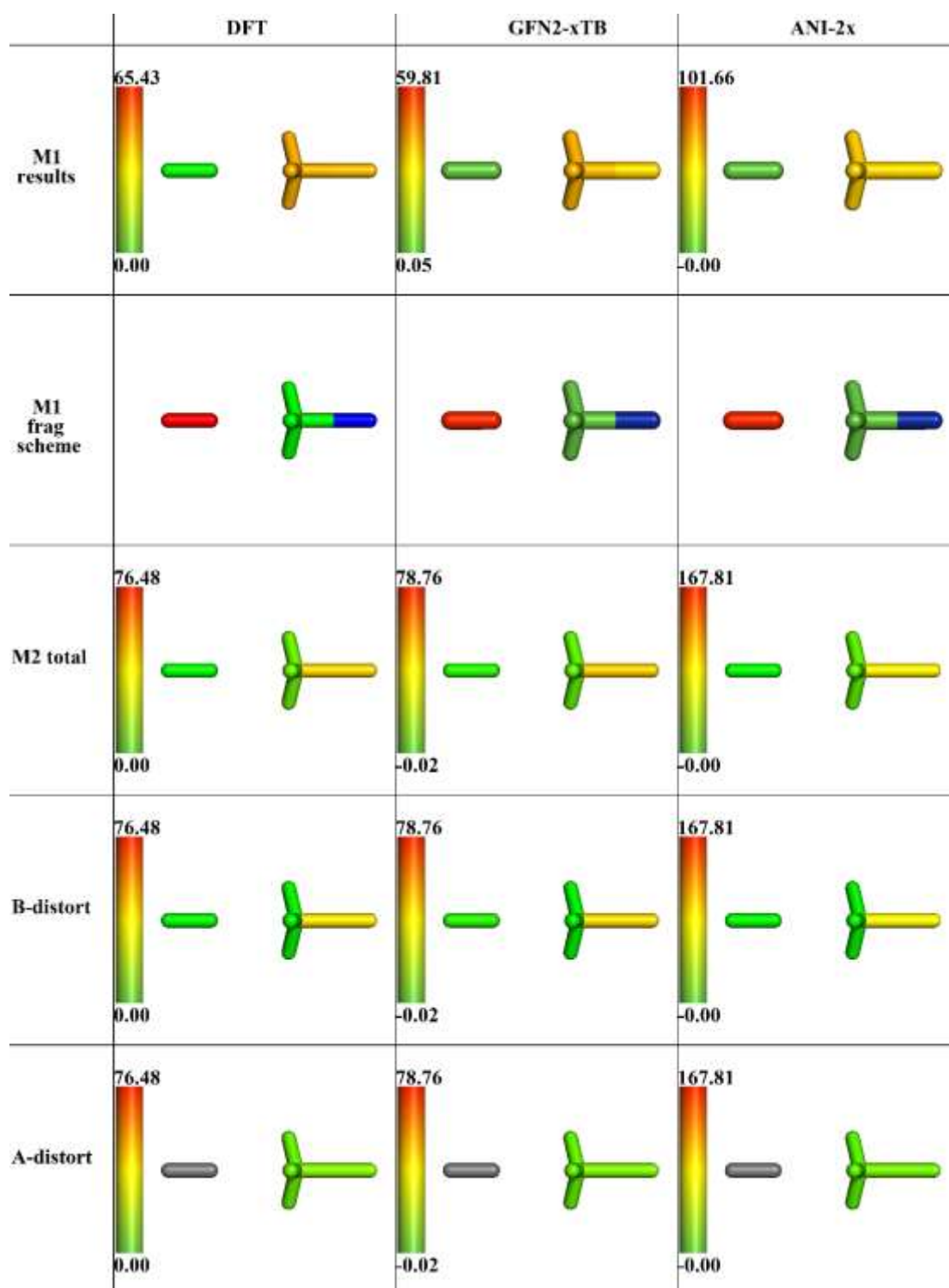
Figure S68 Distortion energy distribution (kcal/mol in log scale) of the TS structure for a S$_N$2 reaction using **M1** scheme, total distortion and its bond and angle contributions using **M2** scheme. All energies were computed using DFT (PCM M06-2X/6-31+G(d)), GFN2-xTB and ANI-2x methods.
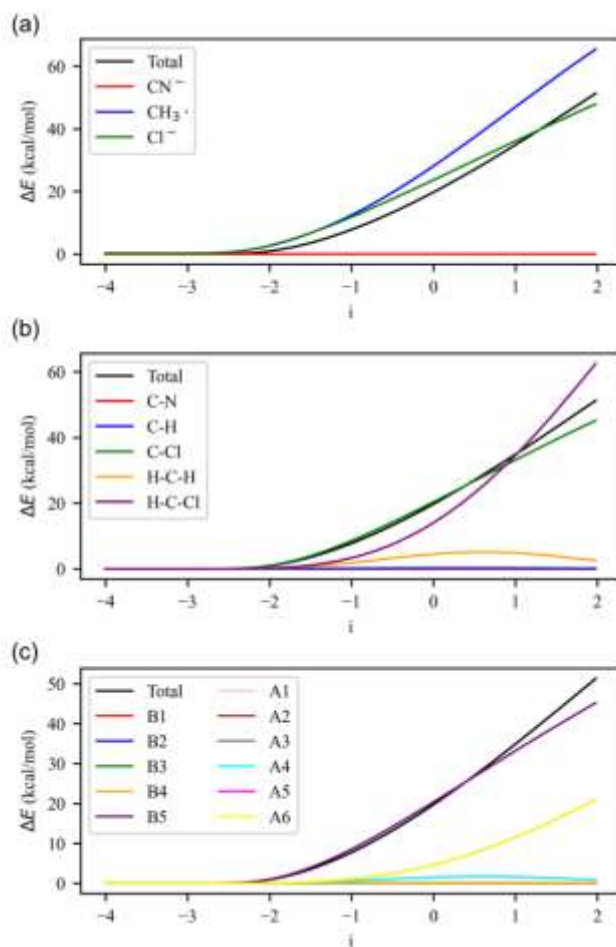
Figure S69 Minimum energy path (MEP) for the S_N2 reaction from the IRC calculations. (a) Distortion distribution changes using **M1** scheme. (b) Distortion distribution changes using **M2** scheme (with summed distortion from the same kind coordinates) (c) Distortion distribution changes using **M2** (B1: C-N; B2, B3, B4: C-H; B5: C-Cl; A1, A2, A4: H-C-H; A3, A5, A6: H-C-Cl). All energies were computed using DFT (PCM M06-2X/6-31+G(d)) method.

Reference:

S1.   M. S. Gordon, D. G. Fedorov, S. R. Pruitt and L. V. Slipchenko, *Chem. Rev.*, 2012, **112**, 632-672.

S2.   R. M. Richard, K. U. Lao and J. M. Herbert, *J. Chem. Phys.*, 2014, **141**.

S3.   N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminf.*, 2011, **3**, 33.

S4.   M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

S5.   C. Bannwarth, S. Ehlert and S. Grimme, *J. Chem. Theory Comput.*, 2019, **15**, 1652-1671.

S6.   P. Zheng, R. Zubatyuk, W. Wu, O. Isayev and P. O. Dral, *Nat. Commun.*, 2021, **12**, 7022.

S7.   J. S. Smith, O. Isayev and A. E. Roitberg, *Chem. Sci.*, 2017, **8**, 3192-3203.

S8.   C. Devereux, J. S. Smith, K. K. Davis, K. Barros, R. Zubatyuk, O. Isayev and A. E. Roitberg, *J. Chem. Theory Comput.*, 2020, **16**, 4192-4202.

S9.   F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1606.

S10.   R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.*, 2007, **79**, 291-352.

S11.   L. a. W. D. Schrödinger *PyMOL*, 2.4.0; 2020.

S12.   E. H. Krenske, E. C. Davison, I. T. Forbes, J. A. Warner, A. L. Smith, A. B. Holmes and K. N. Houk, *J. Am. Chem. Soc.*, 2012, **134**, 2434-2441.

S13.   R. S. Paton, S. Kim, A. G. Ross, S. J. Danishefsky and K. N. Houk, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 10366-10368.

S14.   F. Schoenebeck, D. H. Ess, G. O. Jones and K. Houk, *J. Am. Chem. Soc.*, 2009, **131**, 8121-8133.

S15.   Z. Ma, Z. Yan, X. Li and L. W. Chung, *J. Phys. Chem. Lett.*, 2023, **14**, 1124-1132.

S16.   R. Liang, Q. Zhou, X. Li, M. W. Wong and L. W. Chung, *J. Org. Chem.*, 2023, **88**, 10460-10469.

S17.   M. Radon, *J. Chem. Theory Comput.*, 2014, **10**, 2306-2321.

S18.   M. Garcia-Borras, S. B. J. Kan, R. D. Lewis, A. Tang, G. Jimenez-Oses, F. H. Arnold and K. N. Houk, *J. Am. Chem. Soc.*, 2021, **143**, 7114-7123.

S19.   Z. Yan, X. Li and L. W. Chung, *J. Chem. Theory Comput.*, 2021, **17**, 3783-3796.

S20.   K. K. Kelly, J. S. Hirschi and D. A. Singleton, *J. Am. Chem. Soc.*, 2009, **131**, 8382-8383.

S21.   D. Yepes, P. Jaque and I. Fernandez, *Chem. Eur. J.*, 2016, **22**, 18801-18809.

S22.   M. A. van Bochove, G. Roos, C. Fonseca Guerra, T. A. Hamlin and F. M. Bickelhaupt, *Chem. Commun*, 2018, **54**, 3448-3451.

S23. D. A. Kutateladze, C. C. Wagen and E. N. Jacobsen, *J. Am. Chem. Soc.*, 2022, **144**, 15812-15824.

S24. S. E. Denmark, E. Hartmann, D. J. Kornfilt and H. Wang, *Nat. Chem.*, 2014, **6**, 1056-1064.

S25. C. Hervieu, M. S. Kirillova, T. Suárez, M. Müller, E. Merino and C. Nevado, *Nat. Chem.*, 2021, **13**, 327-334.

S26. T. A. Young, V. Marti-Centelles, J. Wang, P. J. Lusby and F. Duarte, *J. Am. Chem. Soc.*, 2020, **142**, 1300-1310.

S27. C. Y. Legault, Y. Garcia, C. A. Merlic and K. N. Houk, *J. Am. Chem. Soc.*, 2007, **129**, 12664-12665.

S28. M. Mandal, A. M. Luke, B. Dereli, C. E. Elwell, T. M. Reineke, W. B. Tolman and C. J. Cramer, *ACS Catal.*, 2018, **9**, 885-889.

S29. X. Zhang and L. W. Chung, *Chem. Eur. J.*, 2017, **23**, 3623-3630.

S30. Z. Yan, D. Wei, X. Li and L. W. Chung, *Nat. Commun.*, 2024, **15**, 4181.

6. Cartesian coordinates and key absolute energy of the new optimized structures

Structures for benchmark:

The following structures are optimized at M06-2X/6-31G(d) level.

C₂H₆(Benchmark, Fig. S1)

E = -79.83222 AU

C	0.000000	0.000000	0.765684
H	-0.510282	0.884734	1.164342
H	1.021343	-0.000450	1.164342
H	-0.511061	-0.884284	1.164342
C	0.000000	0.000000	-0.765684
H	0.510282	0.884734	-1.164342
H	0.511061	-0.884284	-1.164342
H	-1.021343	-0.000450	-1.164342

C₃H₈(Benchmark, Fig. S2, S4)

E = -119.148581 AU

C	1.275119	-0.260775	0.000000
H	2.176122	0.363318	-0.000001
H	1.315335	-0.908297	-0.884815
H	1.315335	-0.908296	0.884816
C	-0.000157	0.589409	0.000000
H	0.000075	1.249746	-0.878170
H	0.000075	1.249747	0.878169
C	-1.274882	-0.260638	0.000000
H	-1.315481	-0.908505	0.884715
H	-2.176462	0.362812	0.000002
H	-1.315483	-0.908503	-0.884716

NH₂OH(Benchmark, Fig. S3)

E = -131.644293 AU

O	-0.714904	-0.000813	-0.141154
H	-1.115593	0.004269	0.738466
N	0.680407	0.000720	0.156000
H	1.036236	-0.812337	-0.347016
H	1.035735	0.809530	-0.354219

H₃P₂O₄(Benchmark, Fig. S5-7)

E = -644.234007 AU

O	-0.013386	-0.030175	-1.571957
P	-0.000734	-0.002017	-0.109174
O	-1.257511	-0.636653	0.598796
O	1.197297	-0.757850	0.581337
O	0.076895	1.427118	0.550237
H	-1.607481	-1.419724	0.154190
H	2.040265	-0.670355	0.117490
H	-0.448134	2.100820	0.098625

C₂H₄(Benchmark, Fig. S8)

E = -78.536838 AU

C	0.000021	-0.663656	0.000000
H	0.923927	-1.234268	0.000000
H	-0.923937	-1.234282	0.000000
C	0.000021	0.663586	0.000000
H	-0.923956	1.234302	0.000000
H	0.923716	1.234670	0.000000

CH₂=NCH₃(Benchmark, Fig. S11, S12)

E = -133.868779 AU

C	-1.173590	0.184240	-0.000007
N	-0.140505	-0.540887	0.000003

C   1.141609   0.135855   -0.000003
H   -2.156790   -0.286757   0.000002
H   -1.144910   1.284022   0.000021
H   1.707547   -0.182489   0.880409
H   1.707640   -0.182646   -0.880296
H   1.061932   1.233512   -0.000100

CH₂=CH-CH=CH₂(Benchmark, Fig. S13,S14)

E = -155.899155 AU

C   1.523110   -0.493068   0.088557
H   1.133676   -1.426160   0.486821
H   2.586688   -0.449542   -0.121140
C   0.726761   0.556885   -0.116569
H   1.161693   1.492798   -0.464433
C   -0.726761   0.556885   0.116569
H   -1.161693   1.492798   0.464433
C   -1.523110   -0.493068   -0.088557
H   -1.133676   -1.426160   -0.486821
H   -2.586688   -0.449542   0.121140

CH₂=C(CH₃)₂ (Benchmark, Fig. S15)

E = -157.133604 AU

C   0.000012   0.124582   -0.000022
C   -1.273105   -0.678192   0.000003
H   -1.320240   -1.330943   0.880022
H   -2.156747   -0.035539   -0.000076
H   -1.320174   -1.331044   -0.879939
C   1.272947   -0.678456   -0.000001
H   1.319909   -1.331212   0.880020
H   1.319923   -1.331286   -0.879958
H   2.156724   -0.035977   -0.000006
C   0.000168   1.456531   0.000015

H   0.926179   2.024493   -0.000025
H   -0.925704   2.024724   -0.000008

C₅H₈(Benchmark, Fig. S16)

E = -194.900537 AU

C   0.000000   0.000000   0.000000
H   0.000000   0.000000   1.070000
H   1.008806   0.000000   -0.356667
C   -0.889119   -1.088945   -0.628702
H   -0.664527   -2.121923   -0.463131
C   -0.889119   1.088944   -0.628703
H   -0.714669   2.118964   -0.397346
C   -1.891735   0.747129   -1.473979
H   -1.668246   0.357601   -2.445174
H   -2.910337   0.867188   -1.169126
C   -1.955024   -0.743396   -1.390943
H   -1.889943   -0.784264   -2.458180
H   -2.864813   -0.428619   -0.923931

1-butyne (Benchmark, Fig. S16)

E = -155.883820 AU

C   -0.169625   0.045234   -0.122600
H   -0.164117   0.046092   0.973172
H   0.879547   0.047460   -0.438750
C   -0.816286   1.271196   -0.599463
C   -1.366620   2.263150   -1.005261
H   -1.849979   3.144881   -1.361538
C   -0.869062   -1.217358   -0.640814
H   -0.859437   -1.240936   -1.733132
H   -0.364043   -2.112743   -0.269010
H   -1.910182   -1.242835   -0.310352

81

Structures for S3:

The following single point energies are calculated at SCS-MP2/6-31G(d) level by Gaussian16.

S3*-Reactant-C₂Me₂ (Fig. 5)

E = -153.297226 AU

C	0.604438	0.000003	-0.000063
C	-0.604450	0.000119	0.000323
C	-2.065989	-0.000029	-0.000091
H	-2.465688	0.819494	-0.610431
H	-2.465447	-0.938400	-0.404886
H	-2.466138	0.118579	1.014666
C	2.065992	-0.000011	-0.000041
H	2.465725	-0.907841	-0.469156
H	2.465814	0.860125	-0.551633
H	2.465792	0.047552	1.020670

S3*-Reactant-PhCH₂N₃ (Fig. 5)

E = -580.920661 AU

C	0.604438	0.000003	-40.000063
C	-0.604450	0.000119	-39.999677
C	-2.065989	-0.000029	-40.000091
H	-2.465688	0.819494	-40.610431
H	-2.465447	-0.938400	-40.404886
H	-2.466138	0.118579	-38.985334
C	2.065992	-0.000011	-40.000041
H	2.465725	-0.907841	-40.469156
H	2.465814	0.860125	-40.551633
H	2.465792	0.047552	-38.979330
N	2.234002	-0.787780	-0.106845
N	2.387654	0.383710	-0.468037
N	2.611074	1.426139	-0.878938
C	1.177048	-0.989589	0.931108

H	1.165338	-2.069343	1.092051
H	1.493087	-0.513219	1.868691
C	-0.180468	-0.482897	0.503640
C	-0.985628	-1.247907	-0.350309
C	-0.623713	0.782693	0.904200
C	-2.215181	-0.759438	-0.788717
H	-0.641184	-2.227372	-0.673653
C	-1.854600	1.274933	0.465594
H	-0.004236	1.384459	1.565543
C	-2.652188	0.504141	-0.380941
H	-2.833711	-1.363254	-1.447241
H	-2.188841	2.257968	0.785807
H	-3.611182	0.884653	-0.721904

S3_model-Reactant- C₂Me₂ (Fig. 5)

E = -153.271224 AU

C	-1.000278	-1.467497	0.059332
C	0.159942	-1.135488	0.153244
C	1.267604	-0.199305	0.317529
H	1.133227	0.357112	1.221544
H	2.188725	-0.741597	0.366072
H	1.292880	0.472881	-0.514594
C	-2.445249	-1.342047	-0.130215
H	-2.887000	-2.316154	-0.159728
H	-2.864383	-0.784784	0.681379
H	-2.641366	-0.833435	-1.050950

S3_model-TS (Fig. 5)

E = -580.879329 AU

C	1.298958	-0.143020	-0.422582
C	1.618059	-1.313337	-0.184712
N	-0.682907	-0.402831	-1.558934
N	-0.705721	-1.643230	-1.360614

N	-0.046953	-2.521043	-0.964919
C	-1.959480	0.334571	-1.605097
H	-1.650165	1.377907	-1.689160
H	-2.498258	0.072890	-2.525295
C	-2.855664	0.136584	-0.398069
C	-2.462304	0.605153	0.863905
C	-4.081736	-0.524469	-0.525465
C	-3.285483	0.413318	1.972387
H	-1.508025	1.114877	0.964911
C	-4.908111	-0.714582	0.584597
H	-4.393737	-0.893144	-1.500556
C	-4.510623	-0.246073	1.836587
H	-2.972206	0.780968	2.946353
H	-5.858456	-1.229447	0.469572
H	-5.150487	-0.393184	2.702821
C	2.519099	-2.333933	0.376089
H	3.570307	-1.959774	0.428031
H	2.482657	-3.259910	-0.209253
H	2.169814	-2.599387	1.384373
C	1.609821	1.291962	-0.479130
H	2.392394	1.525255	0.287523
H	2.004885	1.533473	-1.481410
H	0.641053	1.934767	-0.302558

S3*-TS (Fig. 5)

E = -580.887864 AU

C	2.247220	0.182674	0.929575
C	2.843937	-0.540508	0.118292
N	0.851203	1.310451	-0.401869
N	1.127736	0.660492	-1.453222
N	1.928692	-0.122006	-1.796685
C	-0.528893	1.721043	-0.159810
H	-0.483295	2.267461	0.790180

H	-0.846511	2.452315	-0.915660
C	-1.546852	0.595107	-0.058162
C	-1.166546	-0.708916	0.279984
C	-2.902299	0.873830	-0.271455
C	-2.128980	-1.712774	0.406030
H	-0.116773	-0.940771	0.432649
C	-3.864117	-0.128034	-0.141666
H	-3.207061	1.882310	-0.544713
C	-3.479083	-1.426857	0.198350
H	-1.819801	-2.722307	0.664273
H	-4.912049	0.103474	-0.313827
H	-4.225837	-2.210457	0.294573
C	3.823244	-1.543991	-0.338519
H	4.513165	-1.110514	-1.070949
H	4.409748	-1.935094	0.501371
H	3.319527	-2.382132	-0.832244
C	1.839444	0.809725	2.195784
H	0.783721	0.609882	2.416109
H	2.437593	0.418766	3.028417
H	1.971706	1.896720	2.161840

S3model-P (Fig. 5)

E = -581.039276 AU

C	2.815363	-1.449917	1.203873
C	1.193882	1.210024	-0.323396
C	0.878727	-0.245571	-0.153391
C	1.519588	-1.314724	0.454338
N	-0.257727	-0.801049	-0.676947
N	-0.317094	-2.120182	-0.417583
N	0.754488	-2.432603	0.257598
C	-1.377644	-0.165498	-1.372512
H	-1.061621	0.847511	-1.631702
H	-1.537470	-0.716934	-2.303111

83

C   -2.651853   -0.135942   -0.550793
C   -2.678223   0.483020   0.705822
C   -3.819746   -0.714363   -1.054375
C   -3.861294   0.519854   1.442015
H   -1.767989   0.925011   1.098701
C   -5.006464   -0.671098   -0.319192
H   -3.801067   -1.208941   -2.022781
C   -5.029264   -0.053735   0.931290
H   -3.872460   0.997511   2.418360
H   -5.907358   -1.127296   -0.720755
H   -5.949818   -0.023428   1.508189
H   3.737327   -1.428085   0.569113
H   2.759673   -2.423746   1.701153
H   2.891963   -0.699661   2.000401
H   2.226086   1.481588   -0.022833
H   1.072605   1.468603   -1.391131
H   0.391242   1.812233   0.295391

C   -3.541195   -0.840139   -0.129102
H   -2.352681   -1.658975   1.473245
C   -2.659906   1.149094   -1.180345
H   -0.786764   1.879010   -0.391929
C   -3.652024   0.168304   -1.086588
H   -4.305428   -1.609034   -0.053910
H   -2.740533   1.935361   -1.926056
H   -4.504974   0.189795   -1.759460
H   3.944794   -1.493162   -1.314715
H   4.750176   0.072815   -1.092785
H   3.566886   -0.137527   -2.388460
H   1.814171   -2.590127   -0.815178
H   1.044407   -2.608964   0.778933
H   0.138339   -2.039101   -0.628810

Structures of S_N2 MD:

The following energy is calculated at HF/3-21G level.

S3*-P (Fig. 5)

E = -581.048768 AU

C   3.801952   -0.407883   -1.351129
C   1.129452   -2.045973   -0.159460
C   1.632328   -0.660136   0.071213
C   2.728072   0.035605   -0.408187
N   1.006681   0.239406   0.882112
N   1.663185   1.424861   0.893637
N   2.704473   1.297129   0.118268
C   -0.259524   0.101011   1.586532
H   -0.239533   -0.822882   2.174512
H   -0.296243   0.942731   2.284379
C   -1.449678   0.113638   0.645423
C   -2.440578   -0.868099   0.731230
C   -1.564397   1.122957   -0.319236

S_N2-TS (MD, Fig. 11)

E = -954.059827 AU

C   0.000000   0.000000   0.000000
H   0.000000   1.060526   0.000000
H   0.918442   -0.530263   0.000000
H   -0.918442   -0.530263   0.000000
Cl   0.000000   0.000000   2.395711
Cl   0.000000   0.000000   -2.395711