**Supporting Information**

# Dynamic Intermediates in the Radical Cation Diels-Alder Cycloaddition: Lifetime and Suprafacial Stereoselectivity

Jacqueline S. J. Tan,[a] Viivi Hirvonen[a] and Robert S. Paton[a,b]*

[a] Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK
[b] Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, USA
robert.paton@colostate.edu



## I. Computational Methods

Density Functional Theory (DFT) calculations were performed with *Gaussian09*[1] using Truhlar's M06-2X hybrid meta GGA functional[2] and Pople's 6-31G(d) split valence basis set.[3] The M06-2X functional was proven to yield accurate energetics specifically for cycloaddition reactions from benchmarking done by Houk *et al.*[4,5] and hence chosen for use. M06-2X methods were found to give small errors for activation enthalpies when compared against G3B3[6], a standard method derived for predicting activation energies with high accuracies. A comparison of M06-2X against benchmark values for the binding energies of 1744 non-covalent dimers gives an RMSD of 0.43 kcal/mol. A similar comparison using the barrier heights of 206 reactions, including pericyclic transformations and hydrogen transfers, gives an RMSD of 2.57 kcal/mol.[7] The B3LYP functional has been used previously to study quasi-classical trajectories of Diels-Alder reactions, although due to systematic errors in the treatment of π *vs*. σ bonds[8] and a lack of dispersion interactions B3LYP activation barriers were found to be several kcal/mol higher than the M06-2X with the same basis set.

Stationary points were optimized with default optimization criteria, tight SCF convergence and ultrafine grid for numerical integration of the exchange-correlation energy and potential. Solvation by dichloromethane ($CH_2Cl_2$) was described by a Solvent Model Based on Density Model (SMD).[9] Additionally, high-precision harmonic frequencies for vibrational modes were computed from the Hessian. All transition structures (TSs) were verified by intrinsic reaction coordinate (IRC) calculations connecting them to relevant ground state structures. Reported energies are zero-point corrected electronic energies at 298.15K, using a quasi-RRHO treatment of vibrational entropies introduced by Grimme,[10] unless otherwise stated. All spin density plots have been generated at 0.05 isovalue using *IQmol* [11]and molecular images produced using *Cylview*.[12]

TS frequencies were processed with *Progdyn*, an external script by Singleton,[13] at 298.15K, to give a randomized set of 150 starting structures with initial velocities, where the displacements represent that of a Boltzmann distribution. Trajectories were then propagated from these displaced starting structures using Born-Oppenheimer molecular dynamics[14,15] (BOMD) in Gaussian09 and Gaussian16, with the SMD/M06-2X gradient being calculated on the fly at each time-step.[16] A step size of 2.5 fs was used for 1200 steps (step lengths of 0.5 amu½ ·Bohr), with force constants updated every 99 steps, alongside the ReadStop option to stop the run as soon as a C-C bond formation criterion is reached (forming C-C bond lengths < 1.6 Å). These settings were chosen to ensure conservation of total energy and momentum. These parameters were justified through comparison against trajectories propagated with shorter step lengths of 0.1 amu½ ·Bohr (approximately 0.7 fs per step) and force constants updated every 10 steps. These calculations took more than 5 times longer to run, while the results were quantitatively indistinguishable from those described above (Figures *S7 and S8*). Some trajectories were also ran using the random seed generator keyword *IOp(1/44)* with

and without solvent, which gave rise to the formation of the 4-membered ring structure and the bond rotation structure, which are not observed when running trajectories generated from *Progdyn*.

All preparation and analysis was automated through custom Python and Bash scripts provided in section VI.

## II. Analysis of the reaction mechanism

We obtained the Gibbs energy profile for reaction **R1**, the neutral Diels-Alder cycloaddition of cylopentadiene **1** with ethylene **2** to form norbornene **3**. In this case, the activation barrier, relative to the complex, is 22.3 kcal/mol, and 29.9 kcal/mol relative to the separated reactants. The M06-2X functional performs well here: compared to the DFT Gibbs energy profile, we found the CBS-QB3 activation barrier to be within 0.9 kJ/mol.

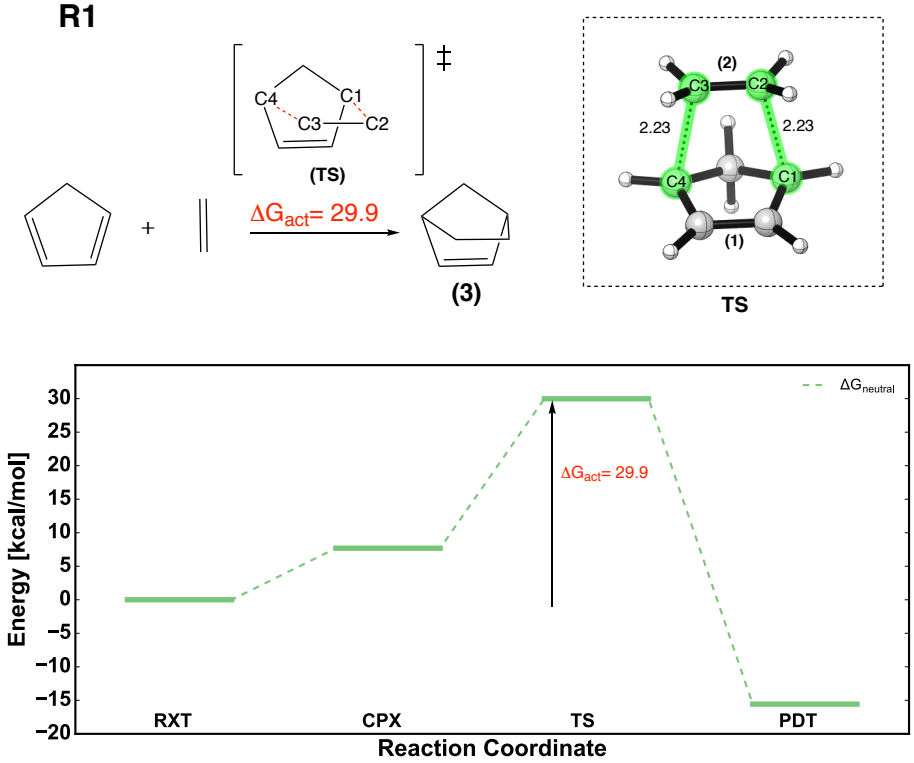
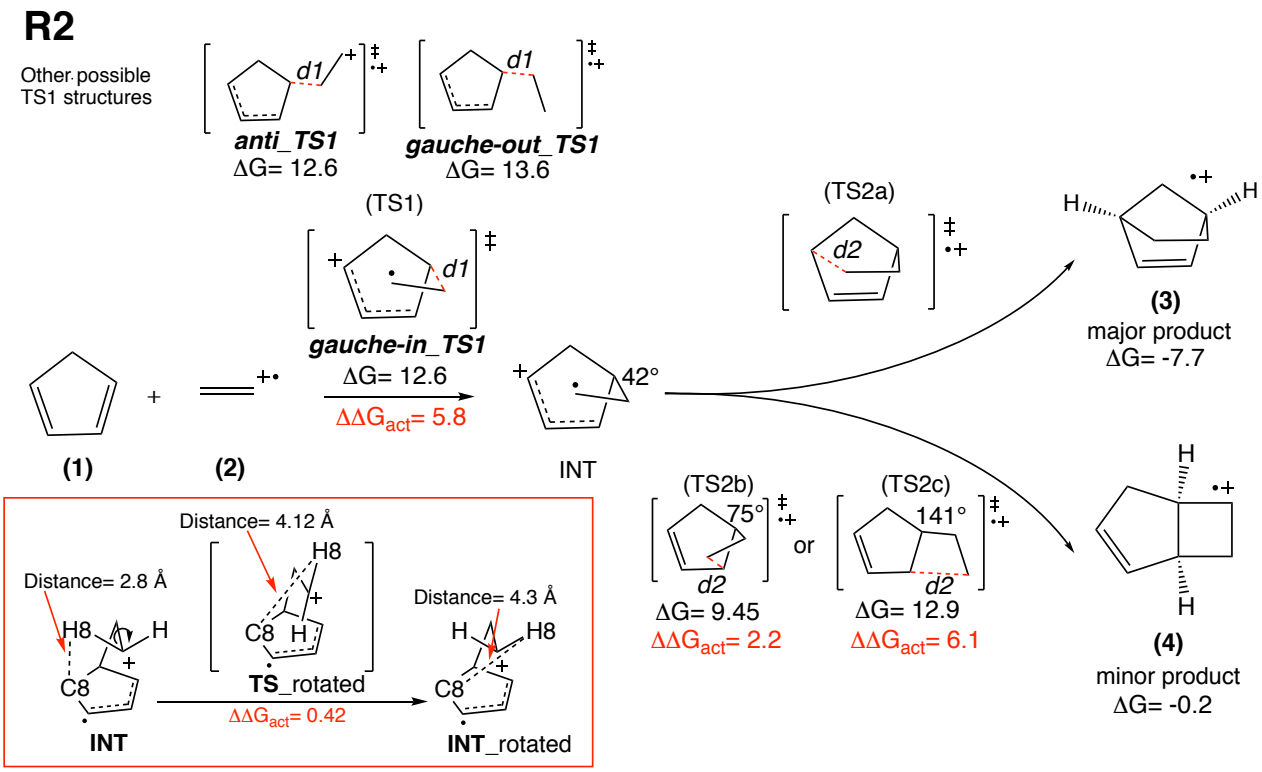



**Figure S1.** M06-2X/6-31G(d)(SMD=dichloromethane) quasiharmonic Gibbs energy profile (298.15K) for **R1** leading to **3**.

We also obtained the Gibbs energy profiles for the radical cationic DA reaction of cylopentadiene **1** with ethylene radical cation **2**, forming norbornene **3** and bicyclo[3.2.0]hept-2-ene **4**. In this case, there are 2 steps to forming the products. There are also 3 TS structures found (*gauche-in*, *anti*, *gauche-out*) for the first step, leading to three slightly different intermediates. We found the first activation energy to be around 6 kcal/mol. Following from the intermediates, two products could be formed from the second step: **3** or **4**. For the formation of **3**, **TS2a** in the second step was not found, as the PES very flat. For the formation of **4**, two TS structures are found (**TS2b and TS2c**), with activation energies of 2.2 and 6.1 kcal/mol respectively. The first C–C forming TS is in every case the highest point along the intramolecular reaction coordinate.


R2
Other possible TS1 structures
d1
anti_TS1
ΔG= 12.6
d1
gauche-out_TS1
ΔG= 13.6
(TS1)
d1
gauche-in_TS1
ΔG= 12.6
ΔΔG_act= 5.8
42°
INT
(1)
(2)
(TS2a)
d2
(3)
major product
ΔG= -7.7
(TS2b)
75°
d2
ΔG= 9.45
ΔΔG_act= 2.2
or
(TS2c)
141°
d2
ΔG= 12.9
ΔΔG_act= 6.1
(4)
minor product
ΔG= -0.2
Distance= 2.8 Å
Distance= 4.12 Å
Distance= 4.3 Å
INT
TS_rotated
INT_rotated
ΔΔG_act= 0.42

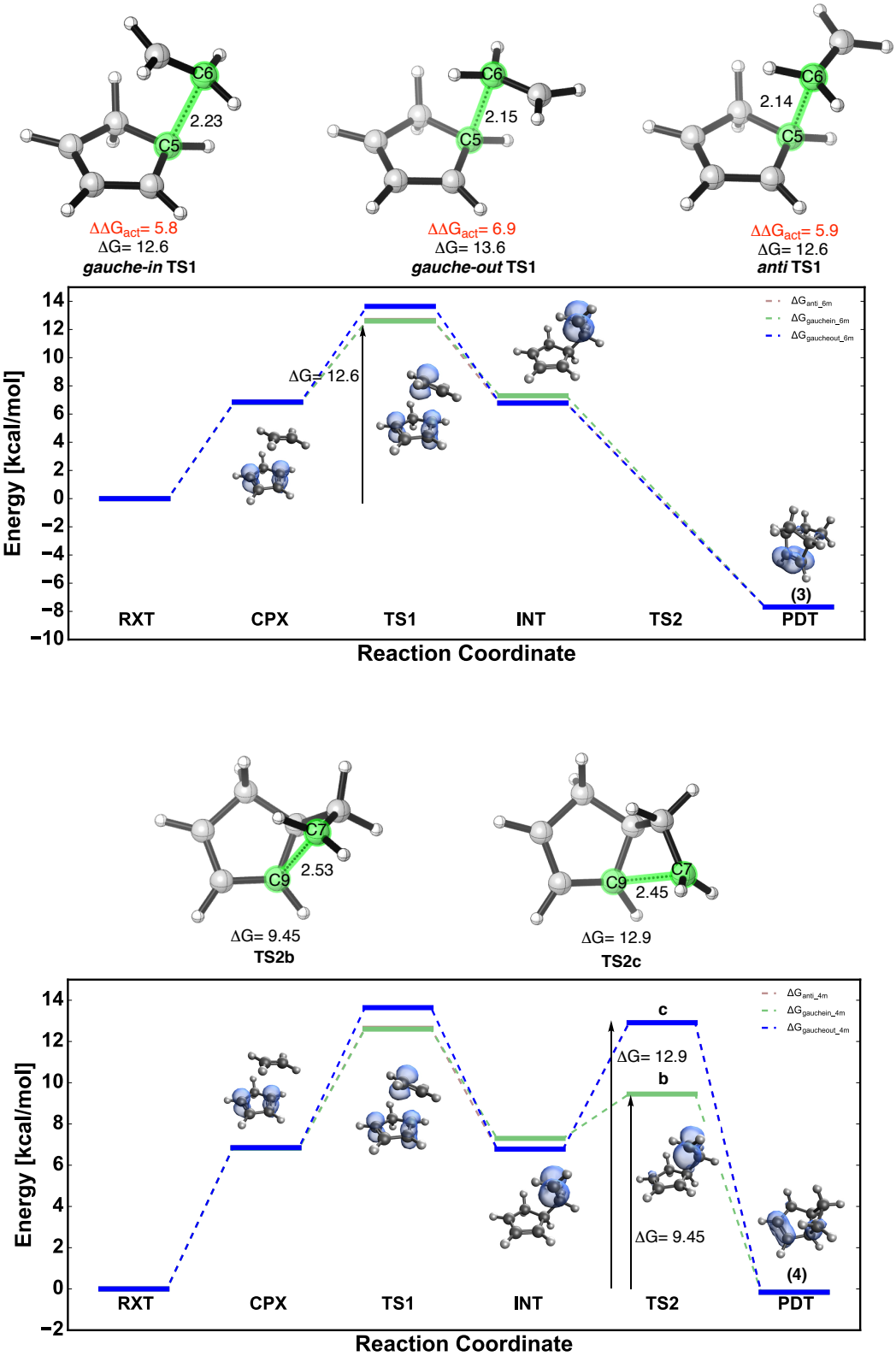



**Figure S2**. M06-2X/6-31g(d)(SMD=dichloromethane) quasiharmonic Gibbs energy profile (298.15K) for **R2** leading to **(3)** and **(4)**.

**Fig. S3** shows the distribution plots of the two forming C–C bond distances for the 142 trajectories of **R1**. For the formation of C3–C4, the average time taken is at 50.3 fs, with a range of bond formation times from 30 to 130 fs from the sampled TS structures. For the formation of C1–C2, the average time taken is 51.7 fs, with a

range of bond formation times from 30 to 130 fs. The degree of concert between the formation of both bonds is high.

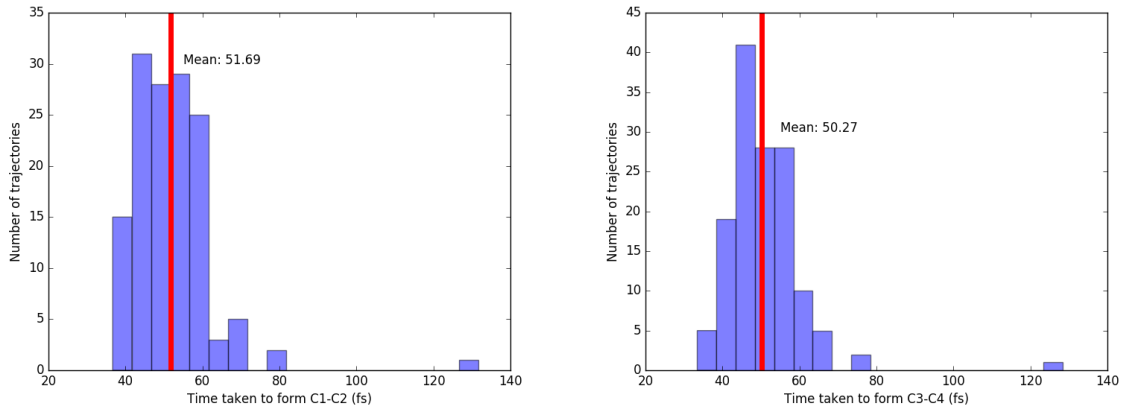



**Figure S3**. C1-C2 and C3-C4 bond forming distances for 142 trajectories

Of the 150 trajectories in **R2**, 40 trajectories recross back the separated reactants in both reverse and forward directions. These trajectories are not analyzed further. **Fig. S4** shows the formation of each C–C bond as a function of time for the 110 trajectories in which product formation was seen. Note that the x-axes now differ by more than an order of magnitude.

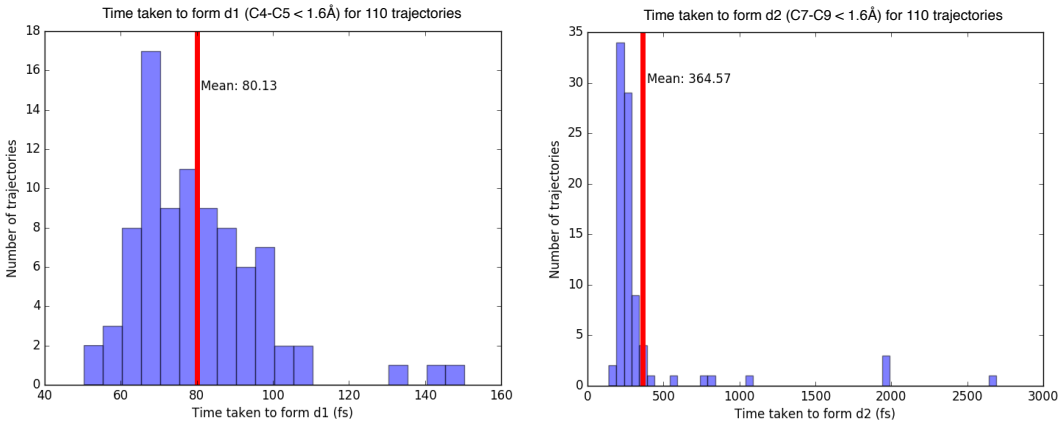



**Figure S4**. C5-C6 and C7-C9 bond forming distances (d1 and d2) for 110 trajectories

The (mean) average time taken for formation of the C5–C6, ($d_1$) bond (i.e. distance < 1.6 Å) is 80.1 fs, with a range of bond formation times from 50 fs to 150 fs from the sampled TS1 structures. The relatively fast formation of the first bond is in contrast to the second (C7–C9 or $d_2$), where the average time taken is at ~365 fs, with a range of bond formation times between 200 fs and 2700 fs.

The time evolution of $d_1$ and $d_2$ in **Fig. S5** show that $d_1$ is formed reasonably fast for all trajectories, while $d_2$ formation is more erratic. Longer trajectories displayed large changes in the $d_2$ bond, where the carbons associated with forming $d_2$ are actually moving away from each other during the trajectory, before finally consummating the bond forming reaction.

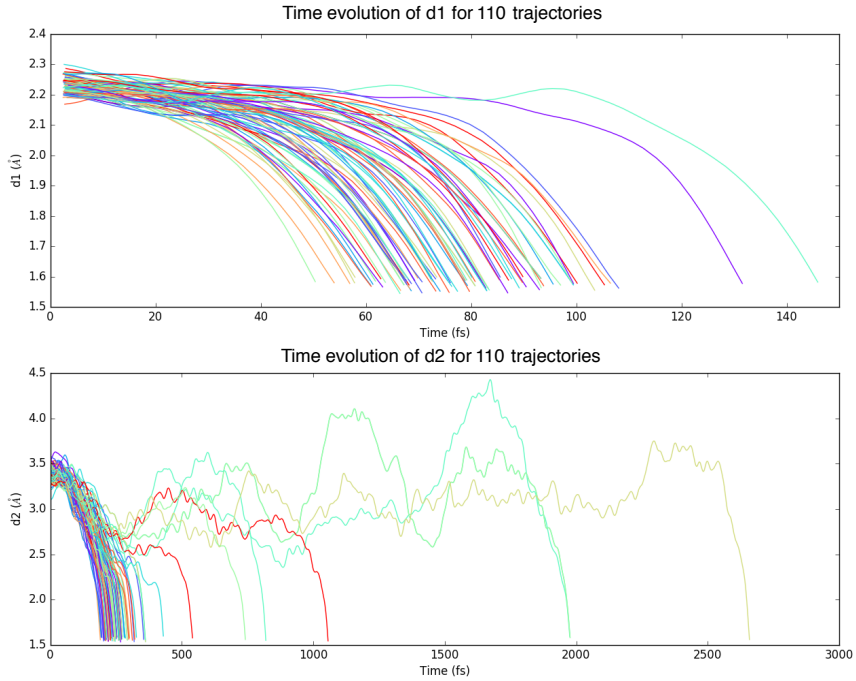



**Figure S5**. *Comparison of time evolutions of d1 and d2*

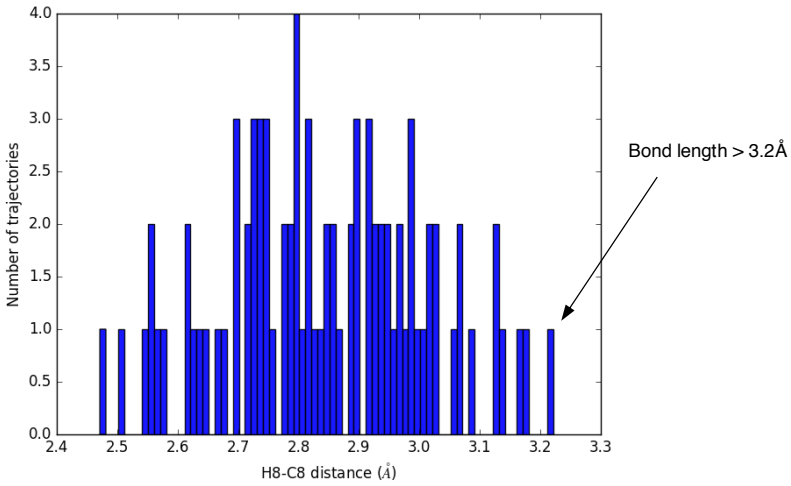



**Figure S6.** Scan of H8-C8 distances, showing a trajectory with bond rotation

As seen from the DFT calculations in **Scheme 2**, one way to differentiate between trajectories that led to bond rotation is to measure the H8-C8 distance of the product. The cut-off for this distance in non-rotated products is hence defined as 3.2Å (**Fig. S6**); if H8-C8 bond length > 3.2Å, it indicates that H8 is now on the other side of the C=C bridge.

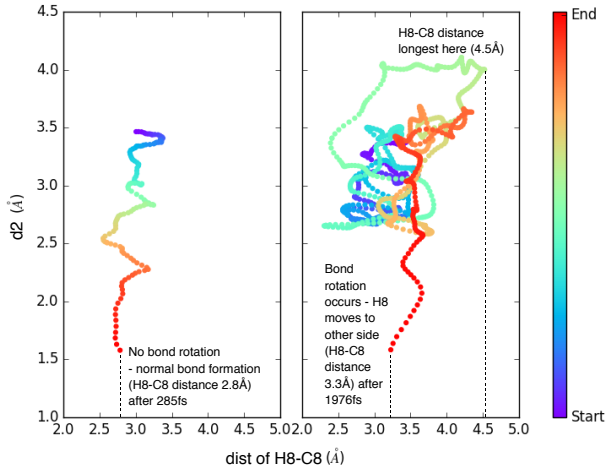



**Figure S7**. Comparison of stereoretentive trajectory with one undergoing alkene rotation.

The normal bond formation trajectory is compared against the long trajectory in **Fig. S7**. On the left figure, the product showed a H8-C8 bond distance of ~2.8 Å, demonstrating that the product retains its stereochemistry due to the normal bond formation of *d2*. The right figure however is more complicated due to bond rotation, resulting in the product having a H8–C8 bond distance of ~3.3 Å, providing credence that H8 has flipped to the other side. Based on Figure S2 as well, the TS structure for bond rotation has a H8-C8 distance of 4.12 Å, which is well-depicted on the right figure as *d2* increases to >4 Å. (See SI Video 4 for example of multiple bond rotations.)

Despite not having any product 4 formation from the quasi-classical dynamics simulations, the long trajectories display a possibility of forming the transition state leading up to product 4, as seen in **Fig. S8** (Video 5). Three trajectories are shown, of the long trajectories, and the red line shows the bond rotation. (See *SI* Video 6 for product 4 formation in gas phase.)

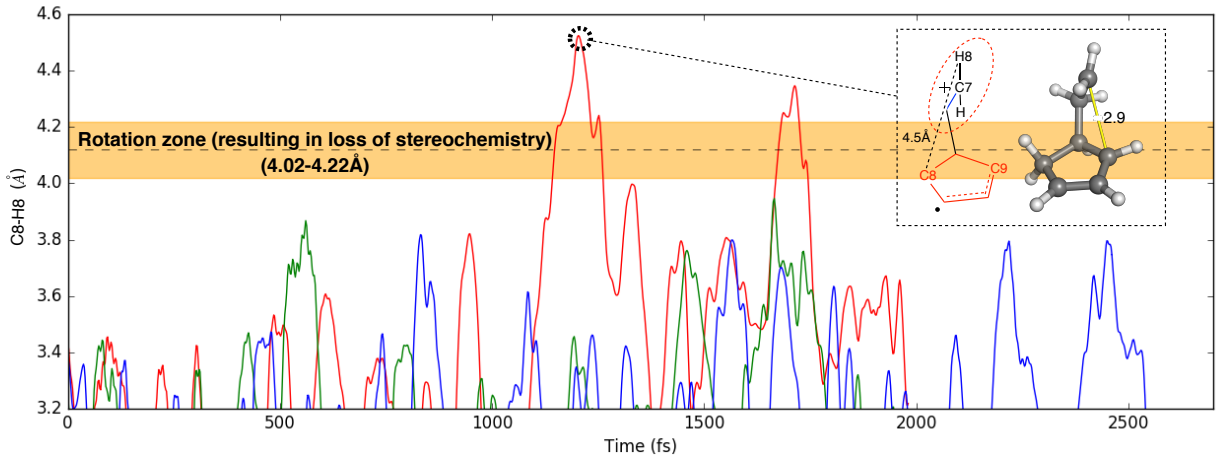



**Figure S8.** H8-C8 distance versus time (rotation zone defined as 4.1 ± 0.1Å)

The rotation zone described in **Fig. S8** is arbitrarily defined as the rotation bond length found from DFT calculations, at 4.1 ± 0.1 Å, where H8 starts to become further from C8. This zone is also a region that includes more than 96% of the initial transition state geometries. Hence, the range of structures formed within this zone is considered as possible TS geometries. It is clear that when the C8–H8 distance reaches a maximum, the intermediate could actually move away from forming product 3; with a possibility for product 4 formations as C7-H9 length decreases.

## III. Absolute energies and Cartesian coordinates

| R1 | M06-2X/6-31G(d) in dichloromethane (SMD) | | | | | |
|---|---|---|---|---|---|---|
| | E/au | ZPE/au | H/au | T.qh-S/au | qh-G(T)/au | imag. v |
| Ethene | -78.539256 | 0.051656 | -78.483622 | 0.025503 | -78.509126 | - |
| Cp | -194.013473 | 0.093663 | -193.914729 | 0.031683 | -193.946412 | - |
| CPX | -272.557058 | 0.146792 | -272.400713 | 0.042567 | -272.443280 | - |
| TS | -272.527576 | 0.149173 | -272.371424 | 0.036340 | -272.407764 | -491.44 |
| PDT | -272.606637 | 0.154786 | -272.445732 | 0.034626 | -272.480358 | - |

| R2 | M06-2X/6-31g* in dichloromethane (SMD) | | | | | |
|---|---|---|---|---|---|---|
| | E/au | ZPE/au | H/au | T.qh-S/au | qh-G(T)/au | imag. v |
| ethene radical | -78.271647 | 0.049609 | -78.218105 | 0.02619 | -78.244295 | - |
| cp | -194.013473 | 0.093663 | -193.914729 | 0.031683 | -193.946412 | - |
| ethene | -78.539256 | 0.051656 | -78.483622 | 0.025503 | -78.509126 | - |
| cp radical | -193.796372 | 0.09344 | -193.697679 | 0.032679 | -193.730358 | - |
| CPX_gauche-in | -272.341202 | 0.146517 | -272.184882 | 0.043731 | -272.228612 | - |
| CPX_anti | -272.341346 | 0.146609 | -272.184989 | 0.043571 | -272.22856 | - |
| CPX_gauche-out | -272.341185 | 0.146509 | -272.184891 | 0.043677 | -272.228568 | - |
| TS1_gauche-in | -272.335328 | 0.147551 | -272.179659 | 0.039741 | -272.2194 | 281.32 |
| TS1_anti | -272.335909 | 0.148078 | -272.179862 | 0.039477 | -272.219339 | -219.63 |
| TS1_gauche-out | -272.334313 | 0.148063 | -272.178278 | 0.03947 | -272.217749 | -260.52 |
| TS_rotated | -272.343657 | 0.147473 | -272.188493 | 0.038694 | -272.227187 | -165.48 |
| INT_gauche-in | -272.343824 | 0.14775 | -272.187641 | 0.040208 | -272.227849 | - |
| INT_anti | -272.344263 | 0.147855 | -272.18796 | 0.040239 | -272.2282 | - |
| INT_gauche-out | -272.344429 | 0.147682 | -272.188204 | 0.040474 | -272.228678 | - |
| INT_rotated | -272.343824 | 0.147735 | -272.187649 | 0.040229 | -272.227878 | - |
| TS2a | - | - | - | - | - | - |
| TS2b | -272.342458 | 0.148575 | -272.186573 | 0.037855 | -272.224428 | -198.00 |
| TS2c | -272.337296 | 0.148869 | -272.181208 | 0.037695 | -272.218903 | -222.21 |
| PDT **(3)** | -272.606637 | 0.154786 | -272.445732 | 0.034626 | -272.480358 | - |
| PDT **(4)** | -272.593386 | 0.153446 | -272.433293 | 0.035886 | -272.46918 | - |

#uM062x/6-31G(d) opt=(ts, noeigen, calcfc) freq SCRF(SMD, solvent=dichloromethane) SCF=(tight, nosym) Int=UltraFine

(neutral DA)

**TS**

| | | | |
|---|---|---|---|
| C | -0.45894 | 1.15406 | 0.36191 |
| C | -1.18026 | 0.70334 | -0.73981 |
| C | -1.18039 | -0.70341 | -0.73962 |
| C | -0.45915 | -1.15404 | 0.36226 |
| H | -0.38931 | 2.19217 | 0.67059 |
| H | -1.55996 | 1.32982 | -1.54022 |
| H | -1.56030 | -1.33004 | -1.53981 |
| H | -0.39011 | -2.19207 | 0.67119 |
| C | -0.36074 | 0.00018 | 1.32366 |
| H | 0.50664 | 0.00021 | 1.98471 |
| H | -1.27049 | 0.00040 | 1.94611 |
| C | 1.61026 | 0.69015 | -0.33759 |
| H | 2.04152 | 1.23868 | 0.49497 |
| H | 1.54620 | 1.24053 | -1.26996 |
| C | 1.61039 | -0.69031 | -0.33774 |
| H | 2.04231 | -1.23916 | 0.49433 |
| H | 1.54643 | -1.24038 | -1.27031 |

(radical cation DA)

**TS1_gauche-in**

| | | | |
|---|---|---|---|
| C | 0.02159 | 0.26963 | -1.02967 |
| C | 0.63930 | 1.24673 | -0.22023 |
| C | 1.49569 | 0.62014 | 0.70954 |
| C | 1.49369 | -0.73633 | 0.47330 |
| H | -0.46098 | 0.49141 | -1.97454 |
| H | 0.45763 | 2.31271 | -0.29534 |

| | | | |
|---|---|---|---|
| H | 2.06168 | 1.12929 | 1.47925 |
| H | 2.04384 | -1.47984 | 1.03808 |
| C | 0.66056 | -1.05374 | -0.72047 |
| H | -0.02646 | -1.89281 | -0.58658 |
| H | 1.32137 | -1.32305 | -1.55995 |
| C | -2.05382 | 0.18268 | -0.20738 |
| H | -2.37663 | -0.45648 | -1.02460 |
| H | -2.20824 | 1.25091 | -0.32986 |
| C | -1.82231 | -0.34607 | 1.02544 |
| H | -1.80798 | -1.41928 | 1.18940 |
| H | -1.61241 | 0.28889 | 1.88091 |

**TS1_anti**

| | | | |
|---|---|---|---|
| C | -0.10996 | -0.06660 | 0.83114 |
| C | 0.76221 | -1.12091 | 0.47902 |
| C | 1.80202 | -0.63067 | -0.34783 |
| C | 1.67087 | 0.72952 | -0.46886 |
| H | -0.75617 | -0.11053 | 1.69903 |
| H | 0.63575 | -2.15240 | 0.78731 |
| H | 2.58006 | -1.23794 | -0.79283 |
| H | 2.31384 | 1.38200 | -1.04765 |
| C | 0.51354 | 1.21466 | 0.33797 |
| H | -0.15632 | 1.88078 | -0.21373 |
| H | 0.87709 | 1.78736 | 1.20452 |
| C | -1.58143 | -0.55252 | -0.64974 |
| H | -1.65387 | -1.62304 | -0.48216 |
| H | -0.96230 | -0.22690 | -1.48217 |
| C | -2.55775 | 0.27122 | -0.15292 |
| H | -3.28308 | -0.09224 | 0.56792 |
| H | -2.59198 | 1.32466 | -0.41293 |

**TS1_gauche-out**

| | | | |
|---|---|---|---|
| C | 0.05793 | -0.21942 | 0.82407 |
| C | -0.38137 | 1.08600 | 0.51677 |
| C | -1.51971 | 1.03144 | -0.31932 |
| C | -1.87533 | -0.28296 | -0.50018 |
| H | 0.67473 | -0.45290 | 1.68342 |
| H | 0.10467 | 1.99335 | 0.85771 |
| H | -2.02550 | 1.89151 | -0.73938 |
| H | -2.69630 | -0.64258 | -1.10927 |
| C | -0.97544 | -1.17578 | 0.28180 |
| H | -0.58836 | -2.02251 | -0.29218 |
| H | -1.52572 | -1.60592 | 1.13303 |
| C | 1.63846 | -0.54273 | -0.58989 |
| H | 1.01258 | -0.30223 | -1.44568 |
| H | 1.71811 | -1.59158 | -0.31889 |
| C | 2.61125 | 0.32787 | -0.17832 |
| H | 2.64675 | 1.34879 | -0.54575 |
| H | 3.34429 | 0.03750 | 0.56737 |

**TS_rotated**

| | | | |
|---|---|---|---|
| C | 0.25705 | 0.18147 | -0.82905 |
| C | -0.46366 | -1.07456 | -0.53913 |
| C | -1.58980 | -0.87872 | 0.24447 |
| C | -1.68044 | 0.47986 | 0.50670 |
| H | 0.24815 | 0.33157 | -1.91875 |
| H | -0.12451 | -2.04078 | -0.90570 |
| H | -2.27304 | -1.64369 | 0.58883 |
| H | -2.46939 | 0.94260 | 1.09354 |
| C | -0.58702 | 1.24246 | -0.11532 |
| H | -0.03540 | 1.80553 | 0.65141 |
| H | -1.00561 | 2.00041 | -0.79314 |
| C | 1.73579 | 0.09534 | -0.38084 |
| H | 2.21239 | 1.05800 | -0.61559 |
| H | 2.24085 | -0.64909 | -1.01032 |
| C | 1.88041 | -0.24590 | 1.06068 |
| H | 2.86812 | -0.32114 | 1.49851 |
| H | 1.02453 | -0.28311 | 1.72607 |

**TS2b**

| | | | |
|---|---|---|---|
| C | -0.28299 | 0.50006 | 0.71987 |
| C | 0.21758 | -0.90634 | 0.70581 |
| C | 1.43100 | -1.02893 | 0.03011 |
| C | 1.76719 | 0.21665 | -0.45703 |
| H | -0.24591 | 0.89800 | 1.74281 |
| H | -0.28496 | -1.71221 | 1.23441 |
| H | 2.00366 | -1.93928 | -0.08964 |
| H | 2.68303 | 0.44001 | -0.99744 |
| C | 0.72988 | 1.23116 | -0.17600 |
| H | 0.28411 | 1.55351 | -1.13071 |
| H | 1.16137 | 2.13318 | 0.27274 |
| C | -1.72883 | 0.42258 | 0.19687 |
| H | -2.06373 | 1.40527 | -0.16349 |
| H | -2.39229 | 0.12091 | 1.01038 |
| C | -1.72462 | -0.61213 | -0.88291 |
| H | -1.10046 | -0.47051 | -1.75995 |
| H | -2.49998 | -1.36716 | -0.93947 |

**TS2c**

| | | | |
|---|---|---|---|
| C | 0.18617 | -0.57662 | 0.64444 |
| C | -0.02601 | 0.90500 | 0.62509 |
| C | -1.13166 | 1.23917 | -0.16280 |
| C | -1.79638 | 0.07288 | -0.46886 |
| H | 0.46178 | -0.94208 | 1.63653 |
| H | 0.53701 | 1.61324 | 1.22406 |
| H | -1.45352 | 2.24192 | -0.41159 |
| H | -2.73536 | 0.02804 | -1.01333 |
| C | -1.13650 | -1.12884 | 0.09833 |
| H | -1.02729 | -1.93004 | -0.64028 |
| H | -1.77723 | -1.52647 | 0.89934 |
| C | 1.35270 | -0.75619 | -0.36618 |
| H | 0.92153 | -0.84983 | -1.36716 |
| H | 1.91881 | -1.67342 | -0.15399 |
| C | 2.20113 | 0.47142 | -0.30282 |
| H | 2.58005 | 0.93037 | -1.20876 |
| H | 2.67752 | 0.74729 | 0.63193 |

## IV. Sample input BOMD calculation

#uM062X/6-31G(d) SCRF(SMD, solvent=dichloromethane) geom=crowd
BOMD=(readstop, maxpoints=1200, recalcfc=99, rtemp=300, stepsize=5000, readvelocity)

Cartesian coordinates from Progdyn

1 2
C 0.0145210 0.2909864 -1.0524406
C 0.5999745 1.2405723 -0.2479957
C 1.4751708 0.6362996 0.7170793
C 1.4852997 -0.7388993 0.4832679
H -0.4949639 0.5142388 -2.0020842
H 0.4703332 2.3464140 -0.2796783
H 2.2617704 1.0545169 1.4393743
H 2.1406690 -1.4819730 1.0958411
C 0.6286538 -1.0884536 -0.6510167
H -0.0105365 -1.9420100 -0.5633474
H 1.3059457 -1.4597696 -1.4403721
C -2.0706620 0.1591107 -0.2202373
H -2.3271523 -0.5010115 -1.0523393
H -2.2531421 1.2846768 -0.3329432
C -1.7534548 -0.3020682 0.9969079
H -1.6185565 -1.4074336 1.2331893
H -1.4237454 0.3263452 1.7553934

2 #stopping conditions for reaction
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
0, 0, 0, 0, 0, 1, 1, 12, 6.50
0, 0, 0, 0, 0, 0, 4, 15, 2.75
#initial velocities for propagation (from Progdyn)
-813263976000.000000 4448598084000.000000 7342840512000.000000
-2665204668000.000000 -6483938328000.000000 5611869648000.000000
-3441297636000.000000 1138430916000.000000 -7698298860000.000000
-163464588000.000000 -2650778676000.000000 475687296000.000000
20012671728000.000000 -34173460440000.000000 -3220144956000.000000
2394661752000.000000 39887566236000.000000 -10535636412000.000000
-15542932104000.000000 -6184352916000.000000 -35304471972000.000000
-11900522592000.000000 27420074640000.000000 -20743962624000.000000
4159025135999.999512 -1898250984000.000000 -2563921080000.000000
6147298332000.000000 46595922408000.000000 4524099048000.000000
-1923192180000.000000 2996272188000.000000 26427179016000.000000
4440691836000.000000 -4808522880000.000000 8488341540000.000000
-26883545220000.000000 -4147509744000.000000 -32896760148000.000000
14965183296000.000000 3159964332000.000000 3416060088000.000000
805050792000.000000 6157929960000.000000 -5391918252000.000000
-10929562308000.000000 -33499645956000.003906 -7263756864000.000000
-4349711772000.000000 7673294160000.000000 32625179999.999996

## V. Benchmarking BOMD conditions

Distance and energy plots comparing BOMD conditions i.e. stepsize and recalcfc, are shown in Figures S7 and S8. These plots are invariant despite using a smaller stepsize or recalculating force constants every 10 steps in the trajectories. Hence, these simulations are done with recalcfc=99 and stepsize=5000 to optimize the computational time.

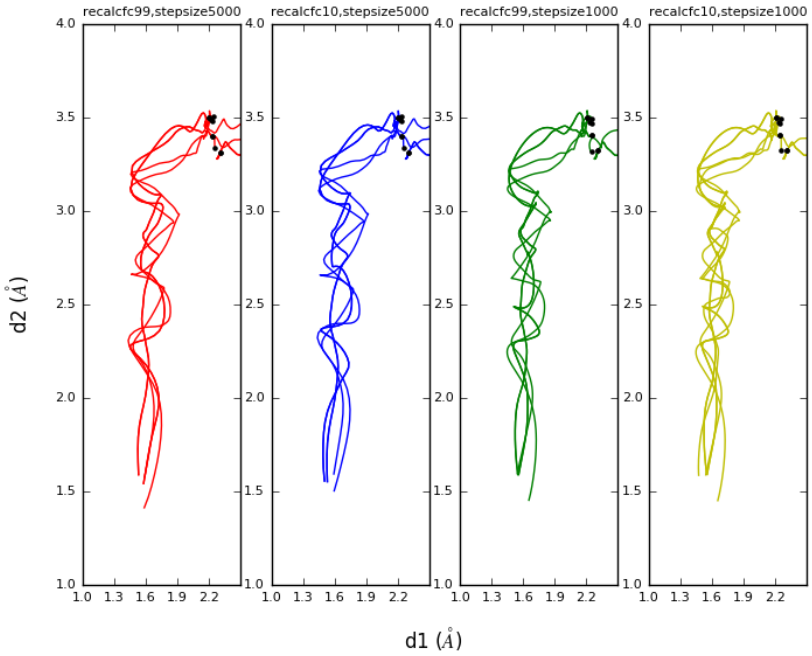



**Figure S7**. Comparison of 10 trajectories with different stepsize and recalcfc conditions (bond lengths d1 and d2)

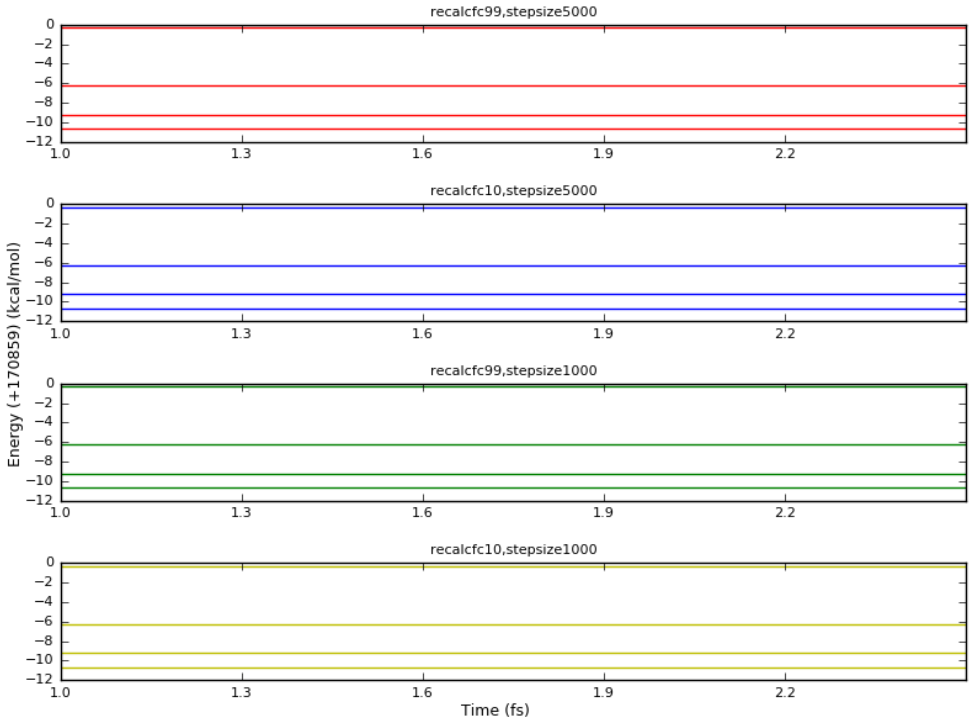



**Figure S8**. Comparison of 10 trajectories with different stepsize and recalcfc conditions (total energy conservation)

## VI. Usage of Progdyn script

First step: **Python extractout.py**. Run this command in the "out.txt" directory, which contains all the Gaussian output files.
Second step: **./sampling.sh**

**extractout.py**:

```
import glob

lookup1 = 'Input orientation:'
lookup2 = 'Harmonic frequencies (cm**-1), IR intensities (KM/Mole), Raman scattering'

filename= ' *.out'
for out in glob.glob(filename):
	NUM_1=1
	NUM_2=1

	with open(out) as myFile:
		for num, line in enumerate(myFile, 1):
			if lookup1 in line: NUM_1=num
			if lookup2 in line: NUM_2=num

	fileout=open(out+'.txt','w')
	with open(out) as myFile:
		for num, line in enumerate(myFile, 1):
			if num >= NUM_1 and num <= NUM_2 : fileout.write(line)
			if num > NUM_2 : break
```

**sampling.sh**:

```
mkdir g09_NC
for i in {1..150}     #to generate 160 geoPlusVel files
do
	if (test -f ./out.txt/TS_NC.out.txt) then
		sh g09.sh ./out.txt/TS_NC.out.txt
		mv geoPlusVel ./g09_NC/geoPlusVel_1$i
	fi
done
```

**g09.sh**:

```
mkdir used
origdir=./
randdir=~/bin
proggramdir=~/program/
#freqfile=$1
# B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1B1 generate geoPlusVel and first input file
echo 3 > runpointnumber
$randdir/randgen > temp811
# the next 8 lines would have to be changed to use low-precision modes
awk '/A       A       A       A       A/,/Harmonic frequencies/ {print}' $freqfile > temp401
awk '/Frequencies --/ {print $3;print $4;print $5;print $6;print $7}' temp401 > tempfreqs
awk '/Reduced masses/ {print $4;print $5;print $6;print $7;print $8}' temp401 > tempredmass
awk '/Force constants/ {print $4;print $5;print $6;print $7;print $8}' temp401 > tempfrc
mv tempfreqs tempfreqs_1
mv tempfrc tempfrc_1
python revisefreq.py
awk '/0/ && ((length($1) < 2) && ($1 < 4)) {print}' temp401 > tempmodes
awk '/has atomic number/ {print}' $freqfile > tempmasses
awk '/Input orientation:/,/tional const/ {if ( $2 < 50 && $2 > 0) print}' $freqfile > tempstangeos
awk -f $proggramdir/proggenHP ' $freqfile > geoPlusVel

if (test -f isomernumber) then
```

```
        cp isomernumber temp533
        awk 'BEGIN {getline;i=$1+1;print i,"----trajectory isomer number----"}' temp533 > isomernumber
        rm temp533
else
        echo "1 ----trajectory isomer number----" > isomernumber
fi
```

**revisefreq.py:**

```
import sys
import os
from numpy import *
import numpy as np
datain=np.genfromtxt('tempfreqs_1')
datain1=np.genfromtxt('tempfrc_1')
fileout=open('./tempfreqs','w')
fileout1=open('./tempfrc','w')
num=len(datain)
for i in range(0,num):
        if datain[i]<100 and datain[i]>0 :
                datain1[i]=datain1[i]*(100/datain[i])*(100/datain[i])
                datain[i]=100
        fileout1.write(str(datain1[i])+'\n')
        fileout.write(str(datain[i])+'\n')
fileout.close()
fileout1.close()
```

## VII. References


(1) 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, H.; 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 Inc.* **2009**, Wallingford CT.

(2) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117* (2), 185–199.

(3) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56* (5), 2257–2261.

(4) Lan, Y.; Zou, L.; Cao, Y.; Houk, K. N. *J. Phys. Chem. A* **2011**, *115* (47), 13906–13920.

(5) Yu, P.; Chen, T. Q.; Yang, Z.; He, C. Q.; Patel, A.; Lam, Y.; Liu, C.-Y.; Houk, K. N. *J. Am. Chem. Soc.* **2017**, *139* (24), 8251–8258.

(6) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **1999**, *110* (16), 7650–7657.

(7) Mardirossian, N.; Head-Gordon, M. *J. Chem. Theory Comput.* **2016**, *12* (9), 4303–4325.

(8) Pieniazek, S. N.; Clemente, F. R.; Houk, K. N. *Angew. Chemie - Int. Ed.* **2008**, *47* (40), 7746–7749.

(9) Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101* (49), 10506–10517.

(10) Grimme, S. *Chem. - A Eur. J.* **2012**.

(11) A. T. B. Gilbert, IQmol molecular viewer, Available at: http://iqmol.org.

(12) Legault, C. Y. *Comput. Programs Biomed.* **2010**, *18* (1–2), 99–108.

(13) Ussing, B. R.; Hang, C.; Singleton, D. A. *J. Am. Chem. Soc.* **2006**, *128* (23), 7594–7607.

(14) Helgaker, T.; Uggerud, E.; Jensen, H. J. A. *Chem. Phys. Lett.* **1990**, *173* (2–3), 145–150.

(15) Uggerud, E.; Helgaker, T. *J. Am. Chem. Soc.* **1992**, *114* (11), 4265–4268.

(16) Gaussian 16, Revision A.03, 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. Mennu, 2016. *Gaussian, Inc. Wallingford CT,* **2016**.