

Machine Learning Analysis of Direct Dynamics Trajectory Outcomes for Thermal Deazetization of 2,3-Diazabicyclo[2.2.1]hept-2-ene

*Nick Rollins[‡], Samuel L. Pugh[‡], Steven M. Maley, Benjamin O. Grant, R. Spencer Hamilton,
Matthew S. Teynor, Ryan Carlsen, Jordan R. Jenkins, and Daniel H. Ess**

Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT 84602

KEYWORDS: Direct dynamics, molecular dynamics, machine learning, deazetization, nonstatistical.

ABSTRACT

Experimentally, the thermal gas-phase deazetization of 2,3-diazabicyclo[2.2.1]hept-2-ene (**1**) results in N₂ loss and formation of bicyclo products **3** (exo) and **4** (endo) in a nonstatistical ratio with preference for the exo product. Here we report unrestricted M06-2X quasiclassical trajectories initialized from the concerted N₂ ejection transition state that were able to replicate the experimental preference to form **3**. We found that the **3**:**4** ratio results from the relative amounts of very fast (ballistic) exo-type trajectories versus trajectories that lead to the 1,3-diradical intermediate **2**. These quasiclassical trajectories provided a set of transition-state vibrational,

velocity, momenta, and geometric features for machine learning analysis. A selection of popular supervised classification algorithms (e.g. Random Forest) provided poor prediction of trajectory outcomes based on only transition-state vibrational quanta and energy features. However, these machine learning models provided more accurate predictions using atomic velocities and atomic positions, attaining ~70% accuracy using initial conditions, and between 85-95% accuracy at later reaction time steps. This increased accuracy allowed feature importance analysis to reveal that at the later time analysis the methylene bridge out-of-plane bending is correlated with trajectory outcomes as either formation of the exo product or towards the diradical intermediate. Possible reasons for the struggle of machine learning algorithms to classify trajectories based on transition-state features is the heavily overlapping feature values, finite, but very large possible vibrational mode combinations, and the possibility of chaos as trajectories propagate. We examined chaos by comparing a set of nearly identical trajectories that differed by only a very small scaling of the kinetic energies resulting from the transition-state reaction coordinate.

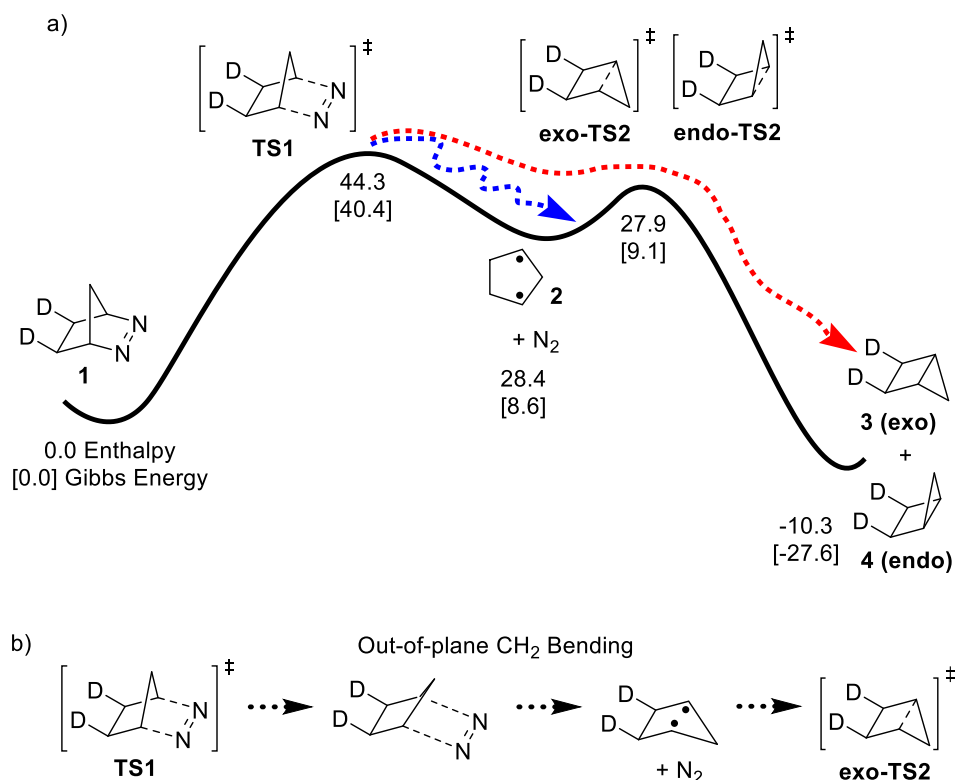
Introduction

Several organic reactions have been discovered that show dynamical reaction selectivity.^{1,2,3,4,5,6,7,8,9,10,11,12} For many of these reactions, transition-state and RRKM statistical theories do not provide adequate quantitative or qualitative modeling of selectivity. One approach to understanding dynamical selectivity effects is to use quasiclassical direct dynamics simulations that provide timing of geometry changes and lifetimes of structures during reactive trajectories,^{13,14,15,16,17,18,19,20,21,22} allowing selectivity modeling beyond statistical theories through revealing non-minimum energy pathway (non-IRC) motion or incomplete intramolecular vibrational redistribution (IVR).^{23,24,25,26}

While quasiclassical trajectories model and rationalize nonstatistical dynamical reaction effects, understanding and predicting the outcome of deterministic trajectories based on initial conditions is nontrivial due to highly complex multi-dimensional energy landscapes with forces at the transition state and forces along the descent to intermediates or products.^{27,28,29} This difficulty of predicting the outcome of a trajectory can inhibit quantitative analysis of qualitative chemical theories proposed to control reaction selectivity. Therefore, in this work, we set out to determine if popular machine learning algorithms could predict quasiclassical trajectory outcomes using vibrational, velocity, and geometric features. Related to our work, previously, Vacher and co-workers successfully used Bayesian neural networks to predict the outcome of ab initio molecular dynamics simulations for the timescale of 1,2-dioxetane decomposition.³⁰

For several reasons we chose to examine the thermal deazitization of 2,3-diazabicyclo[2.2.1]hept-2-ene (**1**, Scheme 1). First, experiments showed that N₂ loss results in bicyclo products **3** and **4** with a nonstatistical preference for the exo product in a 4.7:1 ± 0.9

(exo:endo) ratio.^{31,32} The statistical ratio would be 1:1 if the intervening 1,3-singlet diradical intermediate **2** were to undergo complete IVR prior to C-C bond formation. Second, several qualitative theories for the origin of selectivity have been proposed, beginning with Roth and Martin,^{31,32} Allred and Smith,³³ and culminating with Carpenter's in-depth description that the nonstatistical exo:endo ratio results from a relatively low barrier (~2 kcal/mol) for transition state **TS2** (**endo-TS2** and **exo-TS2**, see Scheme 1) that results in C-C bond formation without IVR.^{34,35} More specifically, Carpenter's description focused on the out-of-plane bending motion of the methylene (CH₂) bridge group that leads to the unexpected excess of exo product **3**, and this rationale was qualitatively supported by classical trajectory analysis and experiments in supercritical propane.³⁶ Third, density functional theory (DFT) quasiclassical trajectories have not been reported.



Scheme 1. a) Outline of static DFT energy surface for the thermal deazetization of **1**. Enthalpies and Gibbs energies reported in kcal/mol. **TS1** is a symmetrical, concerted transition state with simultaneous cleavage of both C-N bonds. Diradical **2** can lead to either **exo-TS2** or **endo-TS2**. The blue dotted arrow represents the possibility of trajectories leading to diradical intermediate **2** with subsequent IVR. The red dotted arrow represents the idea of a ballistic trajectory leading directly from **TS1** to the bicyclo product **3**. b) Snapshot depiction of Carpenter's proposal that non-IVR, out-of-plane methylene bending results in dynamical access to **exo-TS2** from **TS1**.^{34,36}

Here, we report that unrestricted M06-2X quasiclassical trajectories initialized from **TS1** were able to replicate the experimental nonstatistical preference for forming **exo 3** in excess to **endo 4**. We found that most trajectories resulted in intermediate **2** prior to the formation of either bicyclo product, and therefore the **3:4** ratio results from the relative amounts of ballistic-type trajectories (red arrow in Scheme 1a) versus trajectories that lead to **2** (blue arrow in Scheme 1a). These quasiclassical trajectories provided a set of transition-state vibrational, velocity, and geometric features for machine learning analysis. A selection of popular supervised classification

algorithms (Random Forest, MultiLayer Perceptron, Stochastic Gradient Descent, and Logistic Regression Classifier) provided only moderate accuracy for prediction of trajectory outcomes based on transition-state (or near transition state) vibrational, velocity, and geometric features, which is likely due to heavily overlapping feature values. However, these classification algorithms provided significantly more accurate predictions using feature data at later time steps. We also examined chaos by comparing a set of nearly identical trajectories that differed by only a very small scaling of the kinetic energies.

Computational Methods

All structures were optimized with unrestricted M06-2X/6-31G** using Gaussian 09³⁷ or Gaussian 16.³⁸ **TS1**, **2**, **exo-TS2**, and **endo-TS2** have unrestricted SCF solutions lower in energy than restricted solutions with $\langle S^2 \rangle$ values ranging from 0.6 for **TS1** to 1.0 for **2**. For one set of trajectories, using our program DynSuite,³⁹ trajectories were initiated from **TS1** by creating a vibrationally averaged velocity distribution that includes zero-point energy (ZPE) at 413 K. Vibrational motion associated with the transition-state structure imaginary frequency was assigned a specific direction to progress towards one of the minima identified by intrinsic reaction coordinate (IRC) analysis. Translational energy was added to this reaction coordinate mode using a Boltzmann distribution, treating it as a vibrational mode with a frequency of 2 cm⁻¹. Each trajectory was propagated using a Verlet integration algorithm with a one femtosecond (fs) time step. Trajectories were run for 500-3000 fs. DynSuite trajectories were run in both forward and reverse directions to confirm connected pathways. Energies and forces were calculated using Gaussian 09³⁷ with unrestricted M06-2X/6-31G** with the keyword “guess=mix” at each step to

allow for a lower energy broken symmetry solution. Unrestricted DFT does have the possibility of rough transitions between open-shell and closed-shell regions of the potential energy surface.⁴⁰ For this reaction the rough region is likely to be very close to bicyclo product formation and thus have minimal impact to trajectory selectivity.

A second set of quasiclassical trajectories were initialized and propagated in Gaussian 16 using the BOMD procedure. Initialization of trajectories was done using local mode and thermal sampling at 413 K. Trajectories were propagated between 200-500 fs in mass-weighted Cartesian velocities with an approximate average step of 0.25 fs. This second set of trajectories gave a nearly identical exo:endo ratio to the DynSuite set of trajectories. The reason for this second set of trajectories is that the sampling procedure provides a trajectory ensemble initiated as a combination of kinetic and potential energy, which provides the ability to analyze transition-state geometry features.

Results and Discussion

Quasiclassical Trajectory Results

N₂ loss from **1** can occur through either stepwise cleavage of the two C-N bonds or through a one-step concerted cleavage of both C-N bonds. While both mechanisms are plausible, accurate wave function theory (CASPT2) indicates that the concerted mechanism involving transition state **TS1** is ~5 kcal/mol lower in energy than the stepwise pathway.³⁶ Interestingly, DFT generally does not provide a transition state for the first C-N bond cleavage transition state in the stepwise mechanism. Houk has provided an overview of previous quantum-mechanical calculations for deazetization of **1**, and reported a slightly unsymmetrical UB3LYP transition-state structure for

TS1.⁴¹ Because DFT does not provide the first stepwise transition state, and CASPT2 suggests a concerted mechanism that is consistent with experiment, all quasiclassical trajectories were initiated from **TS1** (Figure 1, top).

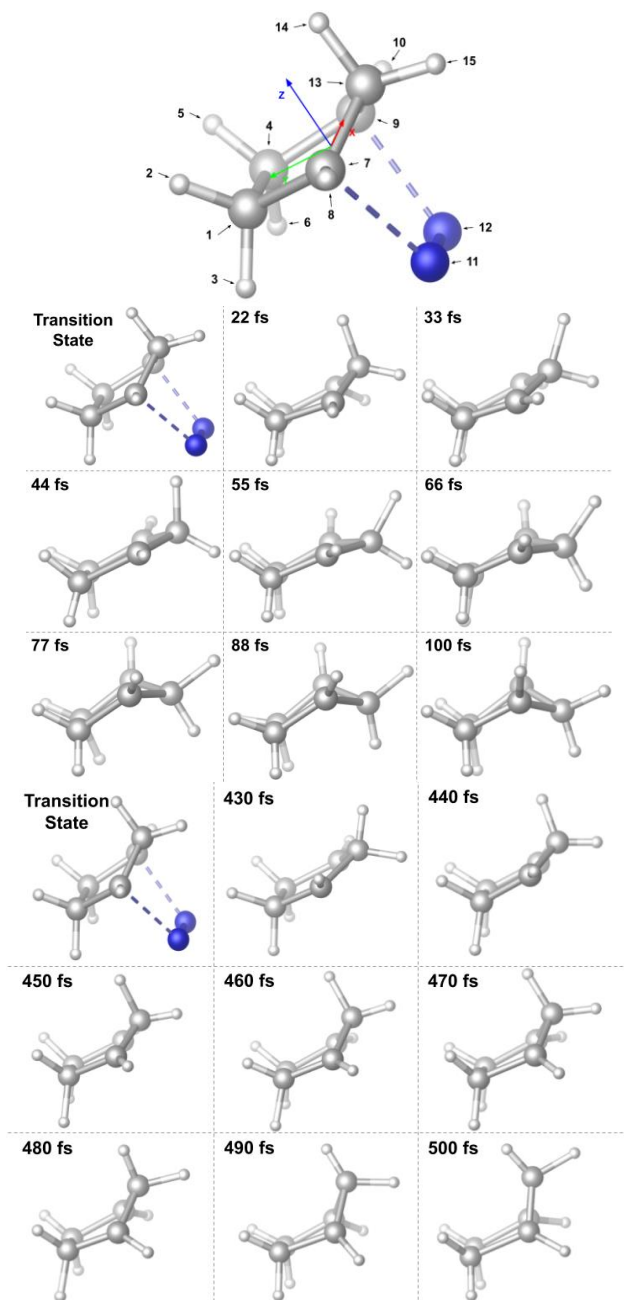


Figure 1. Top: Concerted transition-state structure for N₂ loss. This structure provides atomic number labeling as well as x-, y-, and z-axis definitions for velocities and atomic positions. Middle: Example trajectory leading to exo product **3**. Bottom: Example trajectory leading to intermediate **2** and then to endo product **4**.

Using UM06-2X, 714 DynSuite reactive trajectories were followed for ~500 fs in the reverse direction towards **1** and between 500-3000 fs in the forward direction towards **3** or **4**. In the forward direction, two types of trajectories were found. The first type of trajectory was a “ballistic” trajectory that showed rapid formation of the exo product **3**, typically within ~200 fs. No trajectories formed the endo product **4** within 200 fs. An example of the ballistic exo trajectory is shown in the middle of Figure 1. These snapshots show that within ~30 fs N₂ has completely disengaged from the hydrocarbon framework and the methylene bridge has already moved to create a flat 1,3-diradical conformation. At ~70 fs the methylene bridge has moved to the exo position and at 100 fs the new C-C bond is beginning to be formed. 230 of the 714 trajectories were “ballistic” exo forming.

The second type of trajectory found involves formation of the 1,3-diradical intermediate. Consistent with statistical expectations and the 1,3-diradical intermediate found on the energy landscape shown in Scheme 1, the majority of trajectories resulted in the 1,3-diradical intermediate. Because these types of trajectories likely have significant IVR, they can be considered to ultimately evolve to an equal mixture of bicyclo products **3** and **4**. Consistent with this assertion, the few non-ballistic trajectories that formed products did so in a nearly 1:1 ratio of exo to endo products. Also consistent with this idea, assuming that the 484 of 714 trajectories that are not ballistic go on to an equal ratio of exo to endo products, our final **3:4** trajectory-based ratio is approximately 2:1, which is consistent with the experimental exo preference.³¹ This 2:1 ratio is also interesting since it is similar to Carpenter’s classical trajectory results using a stepwise mechanism and semiempirical-based energy surface.³⁴ In the future, it will be interesting to compare these DFT results to trajectories based on highly correlated multi-reference methods.

Figure 1 shows an example of one of these trajectories that first results in the 1,3-diradical intermediate and then progresses to the endo product. In this trajectory, which requires 500 fs before C-C bond formation, N₂ loss also occurs rapidly. Different than the ballistic exo trajectory, in this endo trajectory the methylene group bends only enough to result in a flat 1,3-diradical and remains in this position for >400 fs before moving in the endo direction with rapid C-C bond formation.

To visually display the two types of trajectories, ballistic exo and 1,3-diradical that evolve to both exo and endo products, Figure 2 plots trajectory time versus relative change in C1-C7-C13-H14 dihedral angle, which showcases the movement of the methylene bridge. Red trajectories are ballistic exo at 200 fs and blue trajectories result in the 1,3-diradical intermediate at 200 fs.

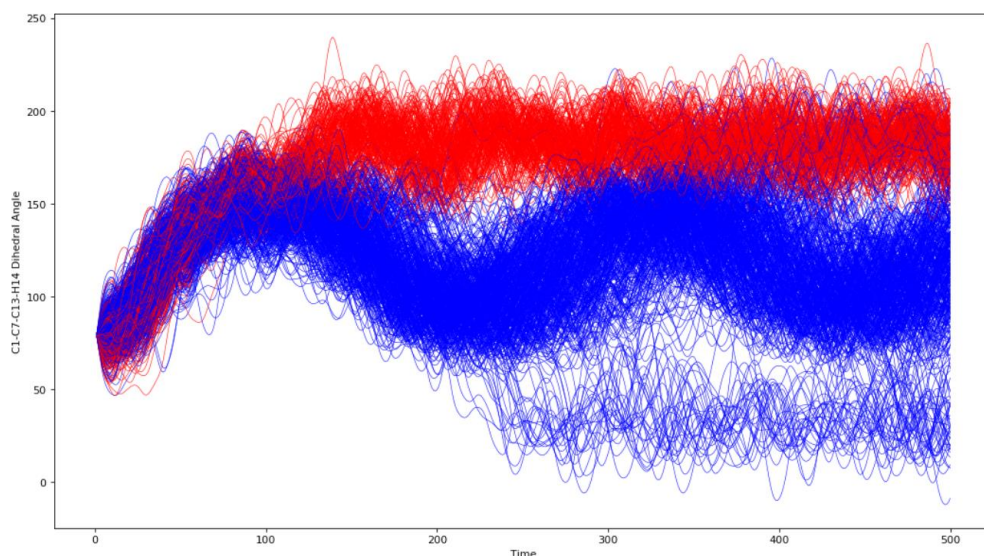


Figure 2. Plot of trajectory time (fs) versus C1-C7-C13-H14 dihedral angle (in degrees). Red trajectories are classified as ballistic exo at 200 fs. Blue trajectories are classified as the 1,3-diradical intermediate at 200 fs. The lower blue band of trajectories corresponds to endo products formed after the intermediate. The upper blue band that overlaps the red band between 300-500 fs shows the formation of exo products after the 1,3-diradical intermediate.

Figure 2 illustrates that in all trajectories (ballistic exo and diradical) during the initial descent down from **TS1** for the first 100 fs there is a relatively large change in the bending of the methylene bridge, and there is essentially no differentiation between the red and blue trajectories. Finally, after 100 fs, this dihedral angle shows segregation into ballistic exo trajectories and 1,3-diradical trajectories. For the ballistic exo trajectories, between 100-150 fs the C1-C7-C13-H14 dihedral angle increases by an additional approximately 35° and then remains at this structure due to C-C bond formation that locks the methylene bridge in the exo position. To show how rapidly C-C bond formation occurs in these ballistic exo trajectories, Figure 3 (top) plots each trajectory time step versus C7-C9 distance. This plot shows that the majority of exo forming trajectories form the C7-C9 bond between 100 and 175 fs. Figure 3 (bottom) also plots the amount of exo and endo products formed versus time. At 100 fs only 1 trajectory has converged to form an exo product and 127 converged on exo products by 130 fs.

In contrast to the ballistic exo trajectories shown in red, all the blue trajectories in Figure 2 show that at about 100 fs there is an inflection point where the C1-C7-C13-H14 dihedral angle rebounds back toward to the starting position from 100-200 fs. It is just beyond 200 fs that some of the trajectories continue this rebound methylene motion and form the endo product **4**, which can be seen by the lower band of blue trajectories. Figure 2 also shows that a small fraction of trajectories that have this rebound motion go onto form nearly equal amounts of endo and exo products. For the 1,3-diradical trajectories shown in blue, the methylene group oscillates above and below the ring plane for the remainder of the trajectory. For many trajectories this oscillation continues for at least 3000 fs.

As another visualization of the two types of trajectories, Figure 4 plots trajectory time both the relative change in C1-C7-C13-H14 dihedral angle and the formation of the C7-C9 bond.

Similar to Figures 2 and 3, the red and blue trajectories are heavily overlapped until about 100 fs when the two types of trajectories begin to diverge.

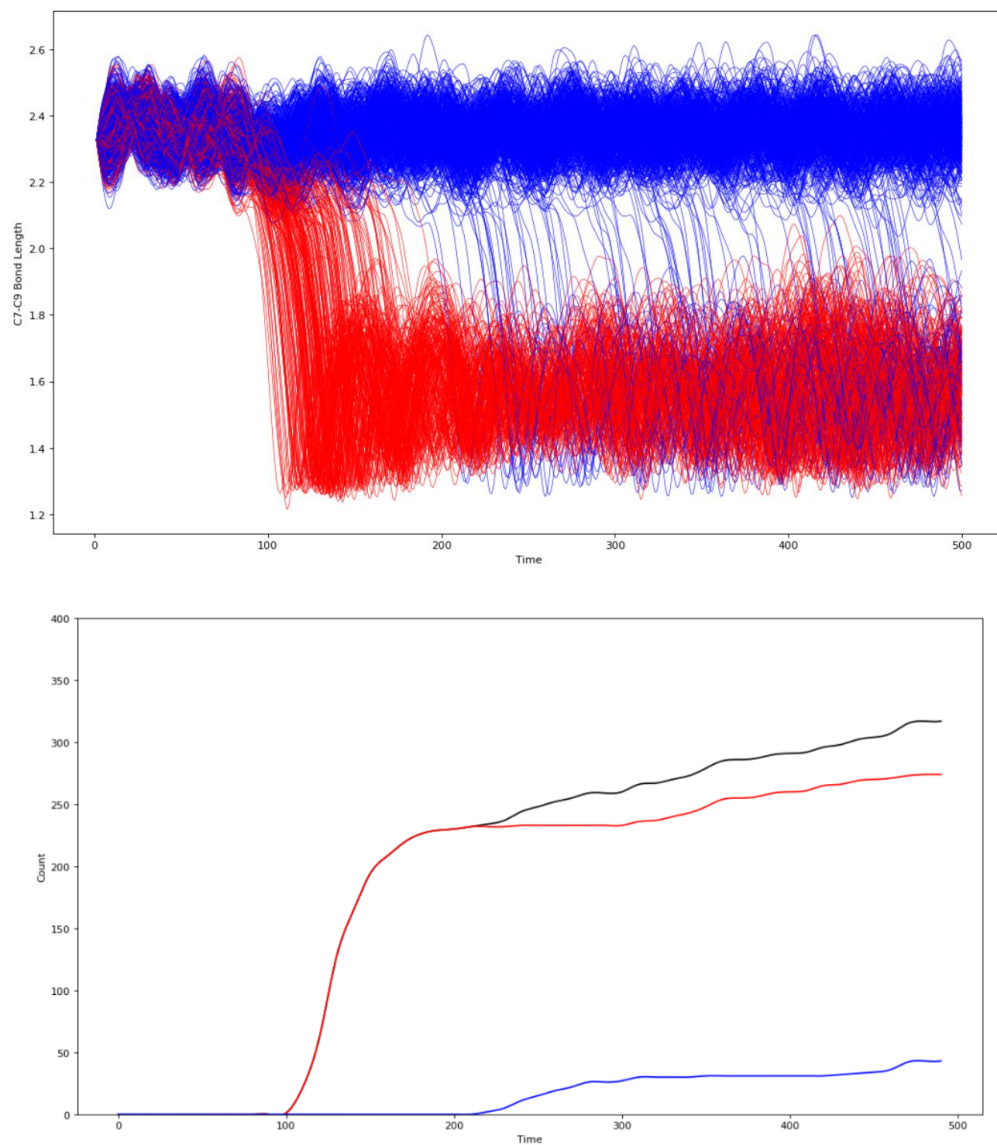


Figure 3. Top: Plot of trajectory time versus forming C7-C9 distance (Å). Red trajectories are classified as ballistic exo at 200 fs. Blue trajectories are classified as the 1,3-diradical intermediate at 200 fs. Bottom: Plot of product count versus trajectory time. Red line shows exo product count, blue line shows endo product count, and black line show the total exo and endo count.

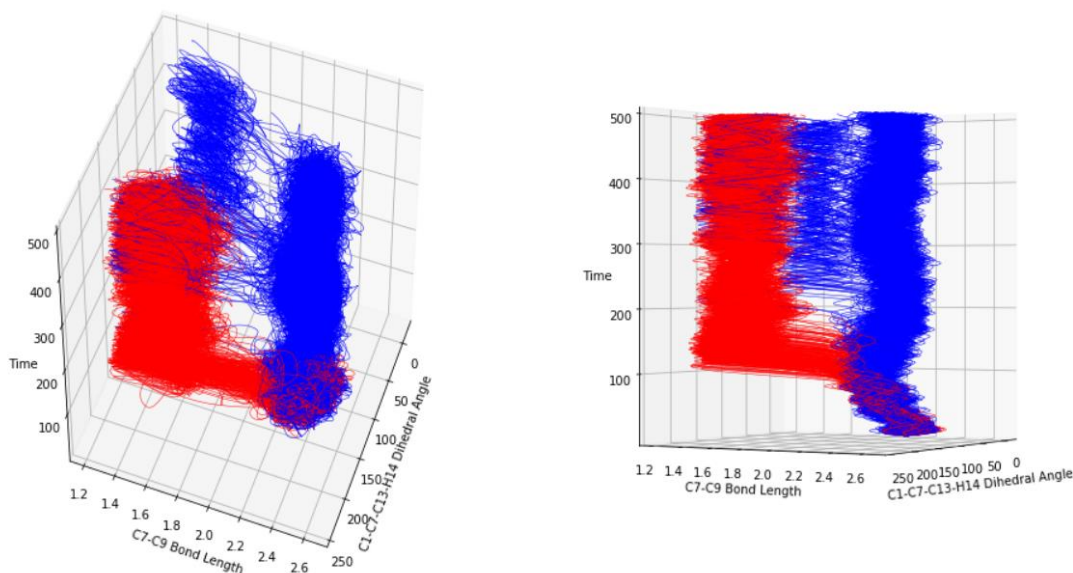


Figure 4. Two perspectives for the plot of trajectory time (fs) versus C1-C7-C13-H14 dihedral angle (in degrees) and C7-C9 distance (Å). Red trajectories are classified as ballistic exo at 200 fs. Blue trajectories are classified as the 1,3-diradical intermediate at 200 fs.

An second set of 439 quasiclassical trajectories were initialized and propagated in Gaussian 16. Different from the DynSuite trajectories where only a vibrationally averaged velocity distribution was generated, these trajectories were initiated with both kinetic and potential (distorted geometries) energy sampling. The geometry sampling is shown at the top left of Figure 5 where all starting atomic positions of trajectories are overlaid. In the top right of Figure 5 these overlaid trajectories are color coded as red and blue for the exo and 1,3-diradical intermediate trajectories, respectively. The bottom of Figure 5 shows the C1-C7-C13-H14 dihedral angle and C-C distances distribution in these trajectories.

Similar to the DynSuite trajectories, we found ballistic trajectories that formed the exo product **3** in ~200 fs. The remainder of the trajectories formed the 1,3-diradical intermediate, then

subsequently formed both endo and exo products in approximately a 1:1 ratio or remained at the intermediate. Following the same assertion that all trajectories that formed the 1,3-diradical intermediate would form bicyclo products **3** and **4** in equal proportion, the trajectory ratio of **3**:**4** is 1.8:1, which is only slightly lower than the ratio found for the DynSuite trajectory set, but again, consistent with the experimental exo preference.

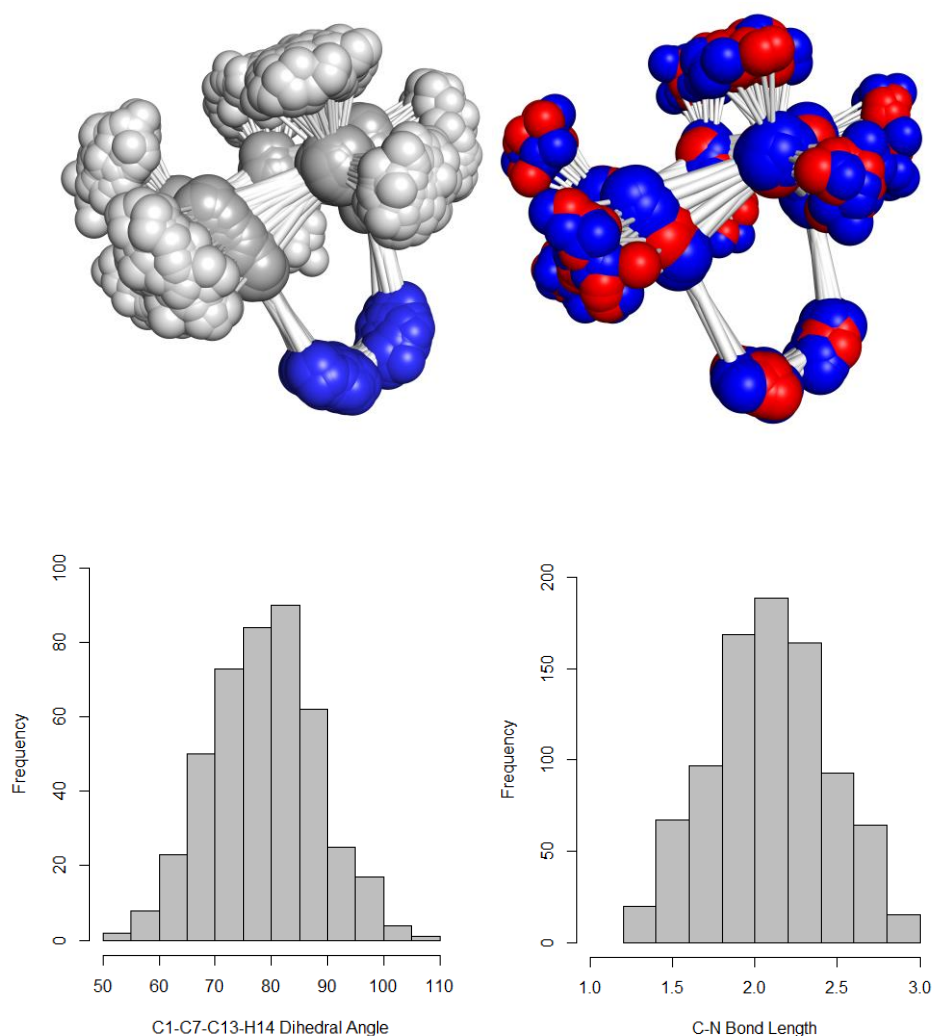


Figure 5. Top left: Overlay of starting geometries for trajectory set 2. Top right: Overlay of starting geometries for trajectory set 2 with all atoms colored as red for ballistic exo trajectories and all atoms colored blue for trajectories that lead to the 1,3-diradical intermediate. Bottom: Histograms of starting C1-C7-C13-H14 dihedral angle (in degrees) and both C-N distances (Å).

Machine Learning Analysis

For machine learning analysis purposes, of the DynSuite trajectories plotted in Figure 2, were categorized as ballistic exo (called class 1) and the remaining were categorized as 1,3-diradical sampling (called class 2). As discussed above, this classification was done at 200 fs. Similarly, the second set of trajectories completely initialized and propagated in Gaussian t were also classified in the same binary fashion. This classification provides a convenient data set to test if common machine learning methods can predict trajectory outcomes using initial vibrational, velocity, or geometric information. While there are many other possible features that could have been used as input for machine learning analysis, vibrational and velocity features naturally represent the result of atomic forces and provide connection to qualitative theories often used to explain nonstatistical effects. Stated another way, this trajectory data set provides the possibility to determine if machine learning methods can identify chemical concepts that provide a rationale for nonstatistical effects.^{42,43,44} The use of machine learning models for analysis is necessary because with 39 vibrational modes, and roughly 20 out of these 39 modes low in energy for excitation with quanta beyond ZPE, there exists an extremely large amount of vibrational mode combinations at the transition state. Also, it is unknown if one or significantly many vibrational mode combinations are the origin of ballistic-exo class 1 trajectories.

For statistical significance, an equal number of class 1 and class 2 trajectories was maintained within the data set by random sampling and iteration, which results in a baseline accuracy of 50%. At each sampling iteration this gave a resulting data set of >400 trajectories with an equal number of class 1 trajectories (>200) and randomly sampled class 2 trajectories. A 20-fold cross validation was used at each iteration to accurately determine the classification accuracy of each model. This entailed dividing the sampled data set into 20 equally-sized subsets, training

the model on 19 of these, and then evaluating the predictive accuracy using the one subset that was withheld (i.e. only withholding ~20 trajectories). This was iterated 20 times, with a different subset not included at each iteration. The model's accuracy is then reported as the mean accuracy of all iterations (200 in total – 20 cross validation iterations for each of 10 random sampling iterations), and is defined by correct predictions divided by total predictions.

The out-of-plane bending motion of the methylene bridge provides a rationale for the excess formation of exo product **3**. Consistent with this key vibrational mode controlling selectivity, Carpenter and coworkers previously reported a histogram of the CASSCF(6,6) vibrational energy transfer in 1,3-diradical **2** that showed that the lowest energy mode is the out-of-plane bending motion and that it showed the largest excitation.⁴⁵ Therefore, for the DynSuite trajectory set, we evaluated thermal energies and vibrational quanta of each vibrational mode to determine if these features alone at the beginning of the trajectory could enable accurate classification.

With and without inclusion of zero-point energy (ZPE), for all machine learning algorithms (see list below), vibrational energy alone, vibrational quanta alone, and vibrational energy and quanta provided poor classification, yielding only marginal (<10%) improvement over the baseline accuracy of 50%. This was surprising to us based on our initial assumption that only a few vibrational mode patterns induce exo-ballistic trajectory outcomes. Stated another way, this machine learning result, which can account for correlation of highly complex vibrational mode combinations, suggests that a single vibrational mode excitation or just a few vibrational mode patterns do not provide a direct mapping or correlation to make trajectory outcome predictions. In retrospect this is perhaps not surprising given the approximately 2^{39} possible directionality and excitation combinations of 39 modes.

To confirm that poor classification accuracy is not a failure of the machine learning algorithms, but that class 1 and 2 trajectories are not significantly distinguishable at the transition state due to a single or combination of vibrational modes, we examined the vibrational mode excitation quanta. As an example of a single vibrational mode analysis, for mode 6, Figure 6 compares the percentage within each trajectory class with each type of vibrational quanta (ZPE, first excited state, second excited state, and so on). Mode 6 was selected as the example because it corresponds to out-of-plane motion of the methylene bridge group. Plots of the other vibrational modes are given in the Supporting Information (SI). This plot visually shows, and based on a statistical χ^2 evaluation, that the frequency of excitation of mode 6 is statistically indistinguishable between class 1 and class 2 trajectories. All other vibration modes also showed indistinguishable excitation quanta patterns. Similarly, for our second trajectory set initiated with kinetic and potential energy, there was not ability to distinguish between class 1 and 2 using mode 6 excitations.

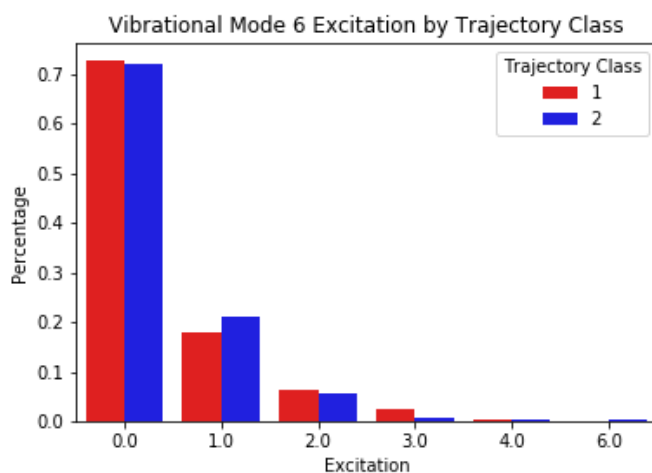


Figure 6. Plot of transition-state vibrational mode 6 excitation quanta for class 1 and class 2 trajectories.

Because transition-state vibrational mode energies and excitation quanta did not provide accurate trajectory classification, we expanded our feature set to include the atomic velocities that were recorded in Cartesian vector components outlined at the top of Figure 1. We chose velocities because these features naturally arise due to transition-state forces, and are directly the result of vibrational modes and excitation, but more importantly velocities can be analyzed at time steps after the transition state. Also, our DynSuite trajectories provide a vibrationally averaged velocity distribution to begin each trajectory. Stated another way, all vibrational energy is added to the transition-state structure as kinetic energy. Momenta can also be used, but for comparison across trajectories, this amounts to a scalar multiplication of velocity features and results in nearly identical results.

For the DynSuite trajectory set, we collected initial atomic velocities as well as velocities at various time steps throughout each trajectory. Correction of velocities after time step 1 for rotation does not significantly alter machine learning results. The Scikit-Learn Python library⁴⁶ was used to set up and train classifiers. Source code illustrating the creation, training, and use of our classifiers can be found in the SI notebooks. We selected four common supervised machine learning classification algorithms: Random Forest, MultiLayer Perceptron, Stochastic Gradient Descent, and Logistic Regression Classifier. Accuracy for each model was evaluated using cross validation, averaged across sampling iterations that was described earlier. With the inclusion of the initial atomic velocities in the feature space the classifiers attained higher predictive accuracy. The best performing model, Logistic Regression, provided 69.5% accuracy. All other models had between 59-67% accuracy (Figure 7). For example, the popular random forest model gave prediction accuracy of 66.8% and 67.1% before and after hyperparameter optimization. As mentioned earlier, the use of momenta rather than velocities, does not provide higher prediction

accuracy. For example, with random forest, the use of momenta gave prediction accuracy of 66.4% before hyperparameter optimization.

While we were at first disappointed with this nearly 70% accuracy, in retrospect, to us, it is actually remarkable given the amount of possible initial starting transition-state configurations and similarity of class 1 and class 2 trajectories (see later pairplots), which is a testament to the ability of the machine learning methods to find predictive correlation patterns.

The GridSearchCV method from the Scikit-Learn library was used to perform hyperparameter optimization for these models. Hyperparameters are keyword options provided to the model algorithm before learning begins. This method tested permutations of different hyperparameters and used five-fold cross validation to determine the set of hyperparameters that maximized model classification accuracy. The optimized hyperparameters returned from the grid search were then used to fit new models on the training data for validation. All four machine learning models showed slight improvement in their classification accuracy with hyperparameter optimization (Figure 7).

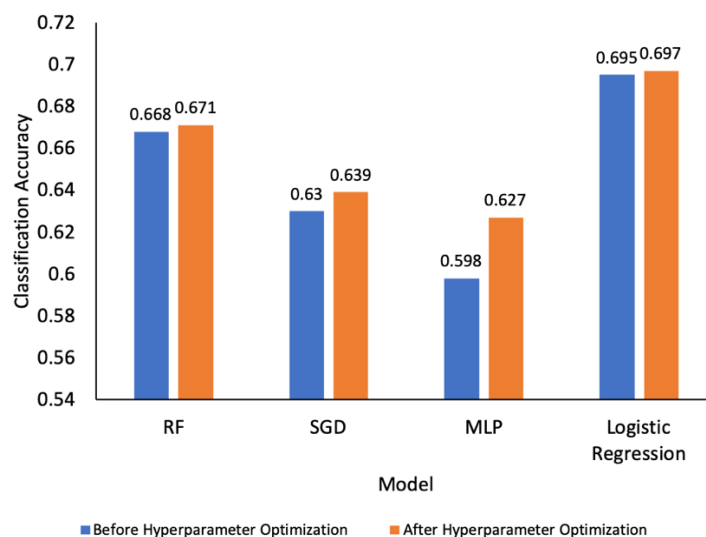


Figure 7. Classification accuracy of Random Forest, MultiLayer Perceptron, Stochastic Gradient Descent, and Logistic Regression Classifier models before and after hyperparameter optimization.

Since optimizing model hyperparameters only provided a marginal increase in classification accuracy, we explored analysis of atomic velocities at time steps beyond the transition state. We analyzed velocities at time steps 3, 10, 50, 100, 125, 130, 137, 150, 175, 200, and 250 fs using the same data set construction and cross validation procedures described above. Figure 8 plots the classification accuracy for Random Forest, MultiLayer Perceptron, Stochastic Gradient Descent, and Logistic Regression Classifier models at these time steps.

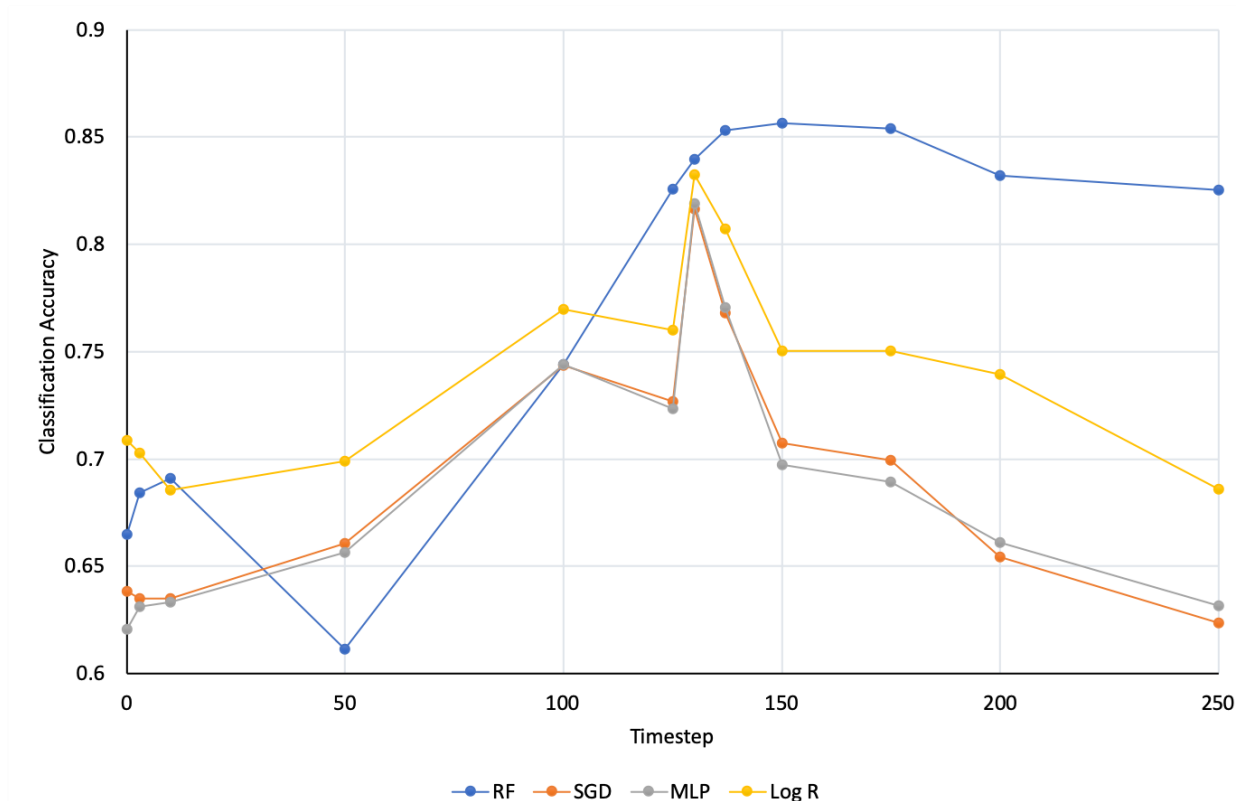


Figure 8. Classification accuracy of Random Forest, MultiLayer Perceptron, Stochastic Gradient Descent, and Logistic Regression Classifier models versus trajectory time step (fs).

For two out of four models, predictive accuracy increases by 50 fs, although the increase is only modest (~5%). However, after 100 fs, the accuracy increases significantly, and for all models accuracy is near 75% at this time step. The maximum accuracy occurs at 130 fs for three out of four models, and then decreases between this time step and 200 fs. At 130 fs, all four models attain between 82% and 84% predictive accuracy. For the random forest model, this rises above 86% at time step 150.

This steady increase in accuracy from 50-130 fs is the result of class 1 and class 2 trajectories finally separating during this time span, which can be seen as the red and blue trajectory line bands diverge in Figures 2 and 4. At this point, many trajectories approach exo product formation. However, the approximately 85% accuracy is larger than the ratio of products that have

a fully formed C7-C9 bond length. As previously mentioned, at 130 fs, 127 of the ballistic exo trajectories have a C7-C9 bond length of $<1.75 \text{ \AA}$ (also see Figure 3). Prior to time step 130, these trajectories are steadily approaching product formation and the overall velocity noise decreases, as can be seen by the near-linear increase in classification accuracy during this period (Figure 6). The decrease in accuracy after 130 fs is likely due to exo product formation and product IVR that causes the trajectory atomic velocities to become chaotic, which results in noise for Class 1 prediction.

To understand the improved prediction accuracy from 0 to 130 fs, we used the Random Forest model fit to the atomic velocities at time step 130 to analyze feature weightings assigned by the permutation feature selection implicit in tree-based algorithms. These weightings show the relative significance of each atomic velocity in predicting Class 1 versus Class 2 selectivity. Figures 9 plots the x-, y-, and z-axis coordinate atomic velocity contributions to the classification model. As defined in Figure 1, the z-axis coordinates of the atomic velocities correspond with the reaction coordinate of the transition state with motion perpendicular to the plane of the developing 1,3-diradical. The x-axis corresponds with the reaction coordinate direction of ring closure step.

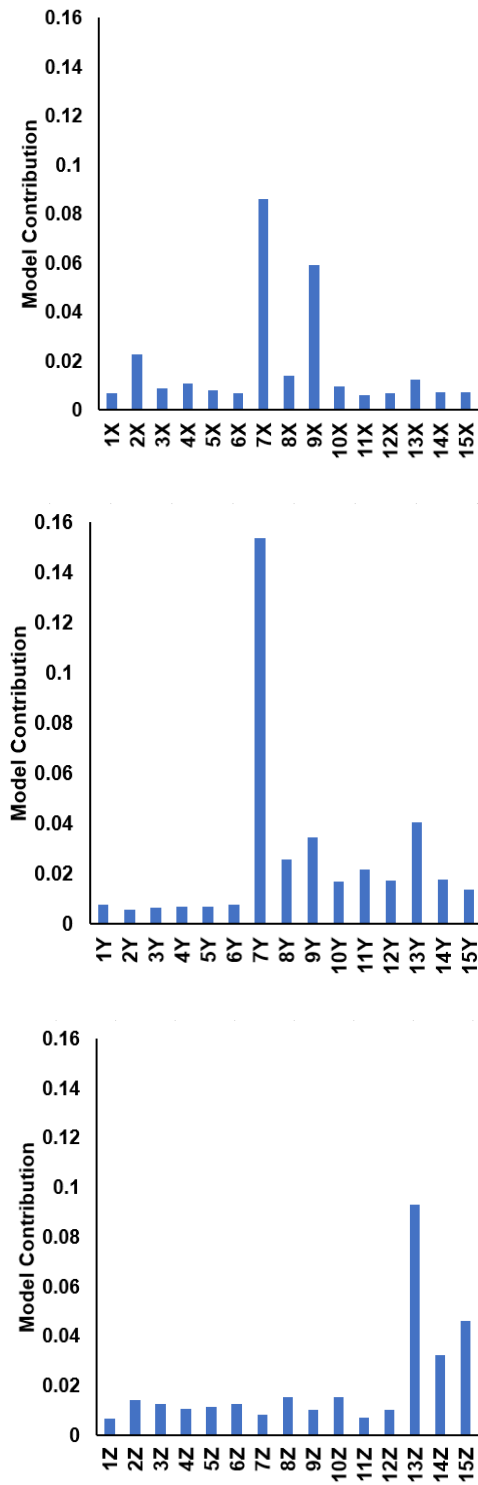
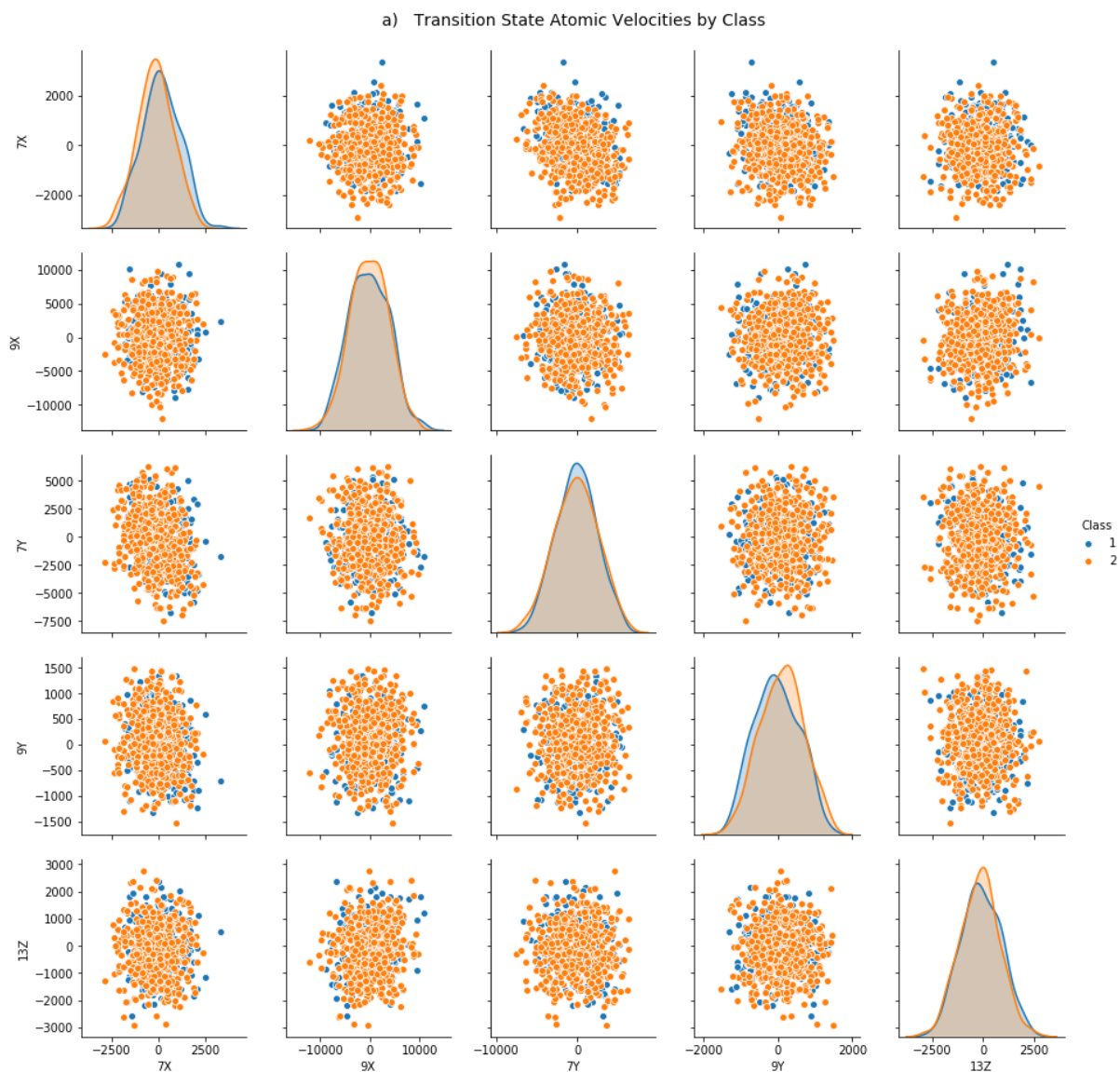


Figure 9. Classification model contributions for random forest plotted in relative feature importance, weighted to 1.0, from Cartesian atomic velocities at time step 130 fs. 1X-15X, 1Y-15Y, and 1Z-15Z are atomic velocities in the x-, y-, and z-coordinates as defined in Figure 1.

Consistent with the reaction coordinate traversing **TS2** for bicyclo product formation, the top plot of Figure 9 reveals that the most influential parameters in the x-axis are the velocities of the bond forming carbons C7 and C9. However, it is clear from the bottom two plots in Figure 9 that the y and z-axis velocities provide the most distinction of features for classification. The large weighting of 7Y is related to the movement of the bond forming carbons. 13Z, 14Z, and 15Z correspond to the methylene group that determines exo or endo product formation. Importantly, identifying the importance of y and z velocities for atoms 13, 14, and 15 demonstrates that machine learning, to a reasonable extent, can begin to identify important chemical groups that provide correlative prediction. However, it is important to note that this is different than showing class 1 versus class 2 to be a simple mapping of the mode 6 bending vibrational excitation shown in Figure 6. We did not group the CH₂ atoms together as a single, mass-weighted parameter because the classification algorithms used, such as the MultiLayer Perceptron, are capable of analyzing the impact of highly correlated parameters.

To visualize the impact of carbon atoms 7, 9, and 13 velocities for prediction of class 1 and class 2 trajectories, Figure 10 illustrates the distributions of the relevant vector components of these atomic velocities. Figure 10a shows these velocities at the transition state, and Figure 10b at 130 fs. The plots along the diagonal show the distributions of the corresponding parameter in Class 1 and Class 2 trajectories. The off-diagonal plots show the parameter indicated at the left of the row plotted along the y-axis, and the parameter indicated at the bottom of the column plotted along the x-axis. From these plots, it can be seen that at the transition state the motion of the ring closure and the methylene group is significantly less informative of trajectory outcome than at time step 130. At the transition state, Figure 10a, the distributions of Class 1 and Class 2 are heavily overlapped, making it difficult to predict, even with machine learning algorithms, the result of a

trajectory. This velocity analysis is highly consistent with Figures 2 and 4 that showed the heavily overlapping methylene bridge position between the transition state and about 100 fs. A time step 130 (Figure 10b), however, the distribution of Class 1 atomic velocities substantially segregates from that of Class 2, enabling distinction between the two classes. Again, this is also consistent with Figures 2 and 4 that show separation of the methylene bridge position 100-130 fs.



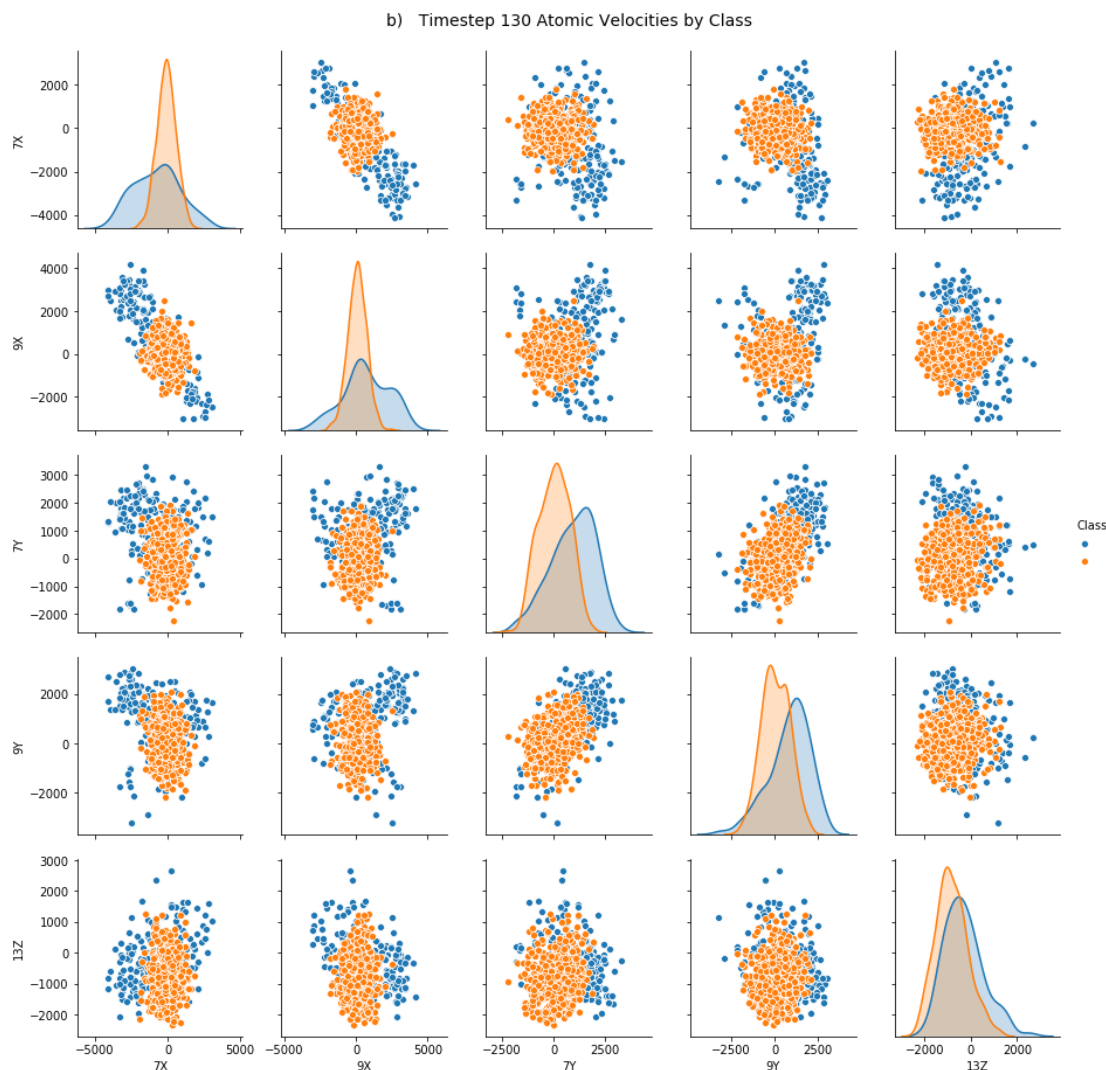


Figure 10. a) Pairplots showing the atomic velocity distributions at the transition state. b) Pairplots showing the atomic velocity distributions at 130 fs.

With random forest we also analyzed the second trajectory set that was initialized and propagated using Gaussian. This data set provided us with transition-state geometry features that could be used for class prediction. We used x, y, and z component atomic positions of each trajectory at time zero. We also included key two-atom distances, angles, and dihedral angles. The two-atom distances were: C7-N11, C9-N12, C7-C9, N11-C13, and N12-C13. The angles included were: C7-C13-C9, C7-N11-N12, and C9-N12-N11. The dihedrals included were: H14-C13-C7-

H8, H15-C13-C7-H8, H14-C13-C9-H10, H15-C13-C9-H10, C1-C7-C13-H14, C1-C7-C13-H15, C4-C9-C13-H14, and C4-C9-C13-H15. In addition, we also harvested and analyzed the transition-state velocities from this data set.

Interestingly, although not hyperparameter optimized, this second trajectory data set resulted in a slightly boost in prediction performance with $\sim 72\%$ accuracy using atomic positions, two-atom distances, angles, and dihedral angles. Using only position or only two-atom distances, angles, and dihedral angles both gave $\sim 71\%$ classification accuracy. From this data set, the use of the initial velocity features showed a decrease in accuracy compared to the DynSuite data set, likely because only part of the total vibrational energy is initiated as kinetic energy.

Similar to the DynSuite data set and analysis of time steps beyond the transition state, analysis of this second trajectory set at later time steps showed an increase in accuracy. We note that Gaussian does not propagate trajectories in a fixed unit time, and therefore, we used a very small approximate time step ~ 0.25 fs. Between 0-50 fs the prediction accuracy remains at about 72% accuracy, but at 100 fs, using the two-atom distances, angles, and dihedral angles the accuracy increases to 87% and goes up to 95% accuracy at ~ 130 fs.

Chaos Analysis

While Figure 10 demonstrates that there is very little distinction of atomic velocities between class 1 and class 2 trajectories at the transition state, and it is perhaps remarkable that machine learning algorithms can achieve a correlative model to nearly 70% accuracy, it is possible that the machine learning algorithms are inhibited from improved prediction with initial velocities at the transition state due to the large number of possible starting configurations. Additionally, it is also possible that machine learning cannot have improved prediction because very small changes

of these starting configurations result in different trajectory outcomes.⁴⁷ One hallmark of a chaotic system is that it shows significant sensitivity to small changes in initial conditions. In general, a chaotic system is difficult for a machine learning classifier to accurately predict outcomes.⁴⁸ From a chemical standpoint, chaos is perhaps more likely for very flat and ridge shape potential energy surface areas. To a certain extent, this chemical reaction has both, easily at the early time steps with descent from transition state **TS1**. For example, along the C_s symmetrical pathway **TS1** directly leads to a C_{2v} transition state for carbon radical pyramidalization.⁴⁵ The directly connected transition states require a potential energy surface bifurcation with a valley-ridge inflection point. Also, the energy surface surrounding the 1,3-diradical intermediate, the transition state for C7 and C9 pyramidalization, and **TS2** is extremely flat.³⁴ This flat energy surface is also a possible reason for why the velocity and geometry features are heavily overlapped between 0-100 fs.

In order to test the response of these quasiclassical trajectories to small perturbations in initial conditions we ran 11 trajectories with identical velocities generated from all positive vibrational normal modes. These 11 trajectories differed by a scaling of the kinetic energies inserted into the transition-state vibrational mode, ranging from 0.9 and 1.1. Figure 11 plots the root mean squared error (RMSE) of atomic position for each of these scaled trajectories relative to the unscaled trajectory. This plot demonstrates that even with extremely small alterations to initial conditions the trajectories deviate significantly within 200 fs and to a very large extent at 500 fs, with RMSE values between 0.1 and 0.3. This suggests that the chaos of these quasiclassical trajectories makes it inherently difficult for traditional machine learning algorithms to predict class 1 versus class 2 trajectories from initial conditions.

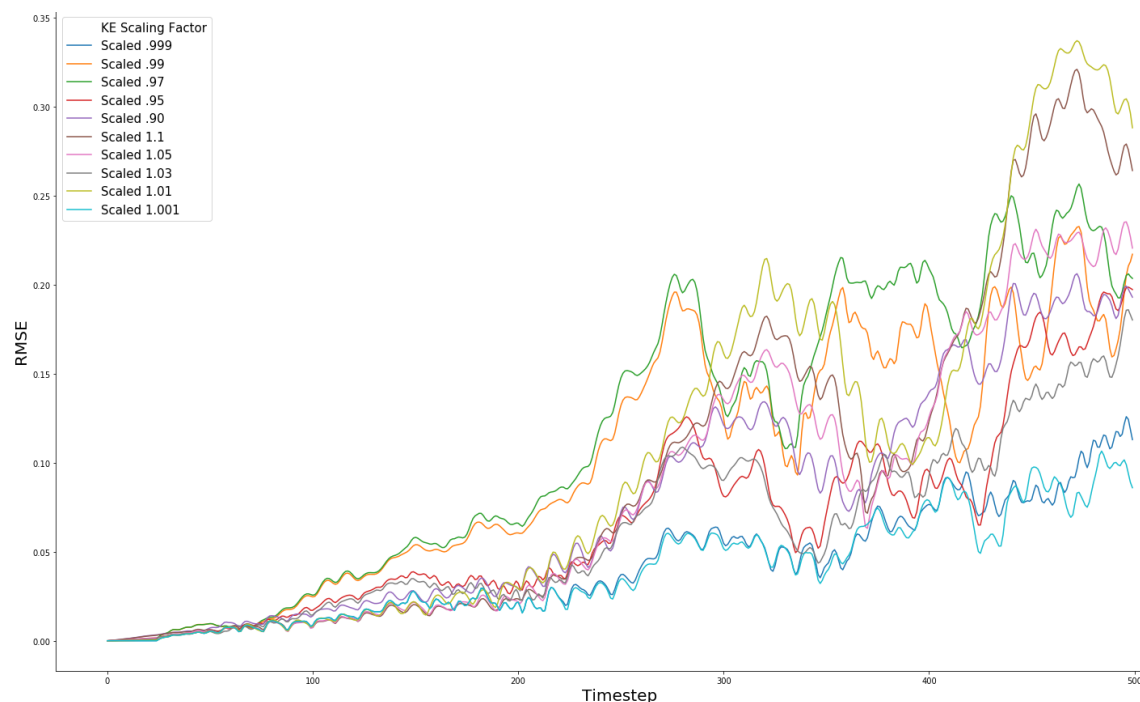


Figure 11. Plot of atomic position RMSE versus time for ten trajectories that differ by scaling the transition-state vibrational mode kinetic energy (KE).

Conclusions

For the diazabicyclo **1**, UM06-2X quasiclassical trajectories were able to replicate the experimental nonstatistical preference of the product **3** versus the endo product **4** starting at the concerted N₂ ejection transition state **TS1**. The trajectories indicate that the **3:4** ratio result from the relative amounts of ballistic exo-type trajectories versus trajectories that lead to the 1,3-diradical intermediate **2**, which lead to equal numbers of exo and endo products. These quasiclassical trajectories provided a set of vibrational, velocity, and geometric features for machine learning analysis with the goal of predicting exo versus 1,3-diradical outcomes.

Perhaps expected after visualization of the similarity of the trajectories at the transition state and in the first 100 fs of the trajectories (Figures 2, 4, and 10), using supervised classification algorithms to analyze vibrational mode energies only provided prediction of class 1 versus class 2 trajectories with approximately 60% accuracy. Additionally, Figure 6 showed that the frequency

of transition-state excitation of mode 6, which is the methylene group bending mode, is statistically indistinguishable between class 1 and class 2 trajectories. However, despite significant overlap of class 1 and class 2 trajectories, the use of transition-state atomic velocities did provide nearly 70% accuracy, and accuracy up to ~85% was achieved at 130 fs past the transition state. We also analyzed a second set to trajectories with different starting atomic positions. This resulted in a slight boost in prediction performance at the transition state with ~72% accuracy using atomic positions, two-atom distances, angles, and dihedral angles. Also, at ~100 fs and ~130 fs the accuracy is boosted to 87% and 95%, respectively.

As one possible reason for machine learning methods not providing higher prediction accuracy based on transition-state features, we analyzed atomic position RMSE values from a series of nearly identical trajectories that differed only by a scaling of the kinetic energies resulting from the transition-state vibrational mode. This revealed significant chaos of the trajectories. We are currently examining recurrent neural networks to see if they provide higher accuracy classification prediction at very early time steps near the transition state. Also, we note that it is possible that trajectories based on extremely accurate multi-reference methods might provide less overlapping features at the transition state and lead to higher machine learning prediction accuracy.

ASSOCIATED CONTENT

Supporting Information

The following files are available free of charge. DynSuite trajectory coordinates, data sets, and example Jupyter Notebooks of model, training, and testing.

AUTHOR INFORMATION

Corresponding Author

*dhe@chem.byu.edu

Author Contributions

The manuscript was written through contributions by N. R., S. L. P., and D. H. E. All authors contributed to generation and analysis of data. All authors have given approval to the final version of the manuscript. [‡]These authors contributed equally.

Funding Sources

The machine learning portion of this work was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Catalysis Science Program, under Award # DE-SC0018329. For the dynamics simulations in this work, D.H.E acknowledges the United States National Science Foundation Chemical Structure, Dynamics, and Mechanisms B (CSDM-B) Program for support under award CHE 1952420. N. R., S. L. P., M. S. T., B. O. G., R. S. H., and J. R. J. thank the BYU Department of Chemistry and Biochemistry for undergraduate research awards.

Notes

Any additional relevant notes should be placed here.

ACKNOWLEDGMENT

We thank the BYU Office of Scientific Computing and the Fulton Supercomputer Lab. We thank Professor Barry Carpenter for helpful discussions and the suggestion to analyze trajectory chaos.

We thank Professor William Hase for helpful discussion about dynamics trajectories. We thank Spencer Yu for generating preliminary recurrent neural network results.

ABBREVIATIONS

DFT, density functional theory; RMSE, root mean squared error.

REFERENCES

-
1. Hare, S. R.; Tantillo, D. J. Post-transition state bifurcations gain momentum – current state of the field. *Pure Appl. Chem.* **2017**, *89*, 679–698.
 2. Doubleday, C. Unusual Potential Energy Surfaces and Nonstatistical Dynamic Effects in Applied Theoretical Organic Chemistry; Ed. Tantillo, D. J. New Jersey: World Scientific 2018.
 3. Bekele, T.; Christian, C. F.; Lipton, M. A.; Singleton, D. A. “Concerted” Transition State, Stepwise Mechanism. Dynamics Effects in C2-C6 Enyne Allene Cyclizations. *J. Am. Chem. Soc.* **2005**, *127*, 9216–9223.
 4. Oyola, Y.; Singleton, D. A. Dynamics and the Failure of Transition State Theory in Alkene Hydroboration. *J. Am. Chem. Soc.* **2009**, *131*, 3130–3131.
 5. Doubleday, C.; Li, G.; Hase, W. L. Dynamics of the biradical mediating vinylcyclopropane-cyclopentene rearrangement. *Phys. Chem. Chem. Phys.* **2002**, *4*, 304–312.
 6. Bailey, J. O.; Singleton, D. A. Failure and Redemption of Statistical and Nonstatistical Rate Theories in the Hydroboration of Alkenes. *J. Am. Chem. Soc.* **2017**, *139*, 15710–15723.
 7. Aziz, H. R.; Singleton, D. A. Concert along the Edge: Dynamics and the Nature of the Border between General and Specific Acid-Base Catalysis. *J. Am. Chem. Soc.* **2017**, *139*, 5965–5972.
 8. Biswas, B.; Singleton, D. A. Controlling Selectivity by Controlling the Path of Trajectories. *J. Am. Chem. Soc.* **2015**, *137*, 14244–14247.

-
9. Chen, Z.; Nieves-Quinones, Y.; Waas, J. R.; Singleton, D. A. Isotope Effects, Dynamic Matching, and Solvent Dynamics in a Wittig Reaction. Betaines as Bypassed Intermediates. *J. Am. Chem. Soc.* **2014**, *136*, 13122–13125.
 10. Biswas, B.; Collins, S. C.; Singleton, D. A. Dynamics and a Unified Understanding of Competitive [2,3]- and [1,2]-Sigmatropic Rearrangements Based on a Study of Ammonium Ylides. *J. Am. Chem. Soc.* **2014**, *136*, 3740–3743.
 11. Bogle, X. S.; Singleton, D. A. Dynamic Origin of the Stereoselectivity of a Nucleophilic Substitution Reaction. *Org. Lett.* **2012**, *14*, 2528–2531.
 12. Wang, Z.; Hirschi, J. S.; Singleton, D. A. Recrossing and Dynamic Matching Effects on Selectivity in a Diels-Alder Reaction. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 9156–9159.
 13. Hare, S. R.; Li, A.; Tantillo, D. J. Post-transition state bifurcations induce dynamical detours in Pummerer-like reactions. *Chem. Sci.* **2018**, *9*, 8937–8945.
 14. Pemberton, R. P.; Tantillo, D. J. Lifetimes of carbocations encountered along reaction coordinates for terpene formation. *Chem. Sci.* **2014**, *5*, 3301–3308.
 15. Hong, Y. J.; Tantillo, D. J. Biosynthetic consequences of multiple sequential post-transition-state bifurcations. *Nat. Chem.* **2014**, *6*, 104–111.
 16. Siebert, M. R.; Manikandan, P.; Sun, R.; Tantillo, D. J.; Hase, W. L. Gas-Phase Chemical Dynamics Simulations on the Bifurcating Pathway of the Pimaradienyl Cation Rearrangement: Role of Enzymatic Steering in Abietic Acid Biosynthesis. *J. Chem. Theory Compu.* **2012**, *8*, 1212–1222.
 17. Siebert, M. R.; Zhang, J.; Addepalli, S. V.; Tantillo, D. J.; Hase, W. L. The need for enzymatic steering in abietic acid biosynthesis: Gas-phase chemical dynamics simulations of carbocation

rearrangements on a bifurcating potential energy surface. *J. Am. Chem. Soc.* **2011**, *133*, 8335–8343.

18. Doubleday C.; Suhrada C. P; Houk K. N. Dynamics of the degenerate rearrangement of bicyclo[3.1.0]hex-2-ene. *J. Am. Chem. Soc.* **2006**, *128*, 90–94.

19. Doubleday, C.; Nendel, M; Houk, K. N.; Thweatt, D.; Page, M. Direct Dynamics Quasiclassical Trajectory Study of the Stereochemistry of the Vinylcyclopropane-Cyclopentene Rearrangement. *J. Am. Chem. Soc.* **1999**, *121*, 4720–4721.

20. Yu, P.; Chen, T. Q.; Yang, Z.; He, C. Q.; Patel, A.; Lam, Y.-h.; Liu, C.-Y.; Houk, K. N. Mechanisms and Origins of Periselectivity of the Ambimodal [6 + 4] Cycloadditions of Tropone to Dimethylfulvene. *J. Am. Chem. Soc.* **2017**, *139*, 8251–8258.

21. Grayson, M. N.; Yang, Z.; Houk, K. N. Chronology of CH $\cdots$ O Hydrogen Bonding from Molecular Dynamics Studies of the Phosphoric Acid-Catalyzed Allylboration of Benzaldehyde. *J. Am. Chem. Soc.* **2017**, *139*, 7717–7720.

22. Xue, X.-S.; Jamieson, C. S.; Garcia-Borras, M.; Dong, X.; Yang, Z.; Houk, K. N. Ambimodal Trispericyclic Transition State and Dynamic Control of Periselectivity. *J. Am. Chem. Soc.* **2019**, *141*, 1217–1221.

23. Lourderaj, U.; Park, K.; Hase, W. L. Classical trajectory simulations of post-transition state dynamics. *Int. Rev. Phys. Chem.* **2008**, *27*, 361–403.

24. Lopez, J. G.; Vayner, G.; Lourderaj, U.; Addepalli, S. V.; Kato, S.; de Jong, W. A.; Windus, T. L.; Hase, W. A. A Direct Dynamics Trajectory Study of F $^-$ + CH $_3$ OOH Reactive Collisions Reveals a Major Non-IRC Reaction Path. *J. Am. Chem. Soc.* **2007**, *129*, 9976–9985.

-
25. Xie, Jing; Otto, Rico; Mikosch, Jochen; Zhang, Jiaxu; Wester, Roland; Hase, William L. Identification of Atomic-Level Mechanisms for Gas-Phase $X^- + CH_3Y$ S_N2 Reactions by Combined Experiments and Simulations. *Acc. Chem. Res.* **2014**, *47*, 2960–2969.
26. Manikandan, P.; Zhang, J.; Hase, W. L. Chemical Dynamics Simulations of $X^- + CH_3Y \rightarrow XCH_3 + Y^-$ Gas-Phase S_N2 Nucleophilic Substitution Reactions. Nonstatistical Dynamics and Nontraditional Reaction Mechanisms. *J. Phys. Chem. A* **2012**, *116*, 3061–3080.
27. Spezia, R.; Martínez-Nuñez, E.; Vazquez, S.; Hase, W. L. Perspective: chemical dynamics simulations of non-statistical reaction dynamics. *Phil. Trans. R. Soc. A* **2017**, 375:20170035.
28. Lourderaj, U.; Park, K.; Hase, W. L. Classical trajectory simulations of post-transition state dynamics. *Int. Rev. Phys. Chem.* **2009**, *113*, 2236–2253.
29. Paranjothy, M.; Sun, R.; Zhuang, Y.; Hase, W. L. Direct chemical dynamics simulations: coupling of classical and quasiclassical trajectories with electronic structure theory. Wiley Interdisciplinary Reviews: Computational Molecular Science, **2013**, *3*, 296–316.
30. Häse, F.; Galván, I. F.; Aspuru-Guzik, A.; Lindh, R.; Vacher, M. How machine learning can assist the interpretation of ab initio molecular dynamics simulations and conceptual understanding of chemistry. *Chem. Sci.* **2019**, *10*, 2298–2307.
31. Roth, W. R.; Martin, M. Zur Stereochemie des thermischen und photochemischen Zerfalls von 2,3-Diaza-bicyclo[2.2.1]hepten-(2). *Justus Liebigs Ann. Chem.* **1967**, *702*, 1–7.
32. Roth, W. R.; Martin, M. Zur stereochemie der 1,2-cycloaddition an das bicyclo[2.1.0]system. *Tetrahedron Lett.* **1967**, 4695–4698.
33. Allred, E. L.; Smith, R. L. Thermolysis of exo- and endo-5-Methoxy-2, 3-diazabicyclo [2.2.1]-2-heptene. *J. Am. Chem. Soc.* **1967**, *89*, 7133–7134.

-
34. Lyons, B. A.; Pfeifer, J. Peterson, T. H.; Carpenter, B. K. Dynamic models for the thermal deazetization of 2,3-diazabicyclo[2.2.1]hept-2-ene. *J. Am. Chem. Soc.* **1993**, *115*, 2427–2437.
35. Sorescu, D. C.; Thompson, D. L.; Raff, L. M. Molecular dynamics studies of the thermal decomposition of 2,3-diazabicyclo(2.2.1)hept-2-ene. *J. Chem. Phys.* **1995**, *20*, 7910–7924.
36. Reyes, M. B.; Carpenter, B. K. Mechanism of Thermal Deazetization of 2,3-Diazabicyclo[2.2.1]hept-2-ene and Its Reaction Dynamics in Supercritical Fluids. *J. Am. Chem. Soc.* **2000**, *122*, 10163–10176.
37. 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, 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; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision B.01. Wallingford CT, 2009.
38. b) Gaussian 16, Revision B.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini,

F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2016.

39. Carlsen, R.; Wohlgemuth, N.; Carlson, L.; Ess, D. H. Dynamical Mechanism May Avoid High-Oxidation State Ir(V)-H Intermediate and Coordination Complex in Alkane and Arene C-H Activation by Cationic Ir(III) Phosphine. *J. Am. Chem. Soc.* **2018**, *140*, 11039–11045.

40. Malpathak, S.; Ma, X.; Hase, W. L. Addressing an instability in unrestricted density functional theory direct dynamics simulations. *J. Comput. Chem.* **2019**, *40*, 933–936.

41. Khuong, K. S.; Houk, K. N. One-Bond, Two-Bond, and Three-Bond Mechanisms in Thermal Deazetizations of 2,3-Diazabicyclo[2.2.2]oct-2-enes, *trans*-Azomethane, and 2,3-Diazabicyclo[2.2.1]hept-2-ene. *J. Am. Chem. Soc.* **2003**, *125*, 14867–14883.

42. Garcia-Meseguer, R.; Carpenter, B. K. Re-Evaluating the Transition State for Reactions in Solution. *Eur. J. Org. Chem.* **2019**, 254–266.

43. Carpenter, Barry K. Dynamic behavior of organic reactive intermediates. *Angew. Chem. Int. Ed. Engl.* **1999**, *37*, 3341–3350.

44. Carpenter, Barry K. Intramolecular dynamics for the organic chemist. *Acc. Chem. Res.* **1992**, *25*, 520–528.

45. Carpenter, Barry K. Energy Disposition in Reactive Intermediates. *Chem. Rev.* **2013**, *113*, 7265–7286.
46. <http://scikit-learn.org>. Accessed on 1/1/2019.
47. Lebender, D.; Schneider, F. W. Neural Nets and the Local Predictor Method Used to Predict the Time Series of Chemical Reactions. *J. Phys. Chem.* **1993**, *97*, 8764–8769.
48. Carpenter, B. K.; Ezra, G. S.; Farantos, S. C.; Kramer, Z. C.; Wiggins, S. Empirical Classification of Trajectory Data: An Opportunity for the Use of Machine Learning in Molecular Dynamics. *J. Phys. Chem. B* **2018**, *13*, 3230–3241.

Table of Contents Graphic

